Max Planck Institut für Biochemie
Abteilung Strukturforschung
Biologische NMR Arbeitsgruppe

Functional and Structural Studies of the FGFR1 Oncogene Partner Protein and Biochemical Investigations of the Retinoblastoma Protein and its Binding Partners

Aleksandra Mikołajka

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigte Dissertation.

Vorsitzender: Univ.-Prof. Dr. St. J. Glaser

Prüfer der Dissertation: 1. Priv.-Doz. Dr. N. Budisa

2. Univ.-Prof. Dr. J. Buchner

Die Dissertation wurde am 22.11.2006 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 05.02.2007 angenommen.

PUBLICATIONS

Smialowski P., Martin-Galiano A., **Mikolajka A**., Girschick T., Holak T.A., Frishman D. (2006). **Protein solubility: sequence based prediction and experimental verification** *Bioinformatics*, submitted.

Mikolajka A., Yan X, Popowicz G., Smialowski P., Nigg E.A., Holak T.A. (2006). **Structure of the N-terminal domain of the FOP (FGFR1OP) protein and the mechanism of its dimerization and centrosomal localization**.
J Mol Biol **359**: 863-75.

Singh M., Krajewski M., **Mikolajka A.**, Holak T.A. (2005). **Molecular determinants for the complex formation between the retinoblastoma protein and LXCXE sequences**.
J Biol Chem **280**: 37868-76.

Smialowski P., Singh M., **Mikolajka A.**, Majumdar S., Joy J.K., Nalabothula N., Krajewski M., Degenkolbe R., Bernard H.U., Holak T.A. (2005). **NMR and mass spectrometry studies of putative interactions of cell cycle proteins pRb and CDK6 with cell differentiation proteins MyoD and ID-2**.
Biochim Biophys Acta **1750**: 48-60.

ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
Å	Ångstrom
AP	alkaline phosphatase
APS	ammoniumperoxodisulfate
bHLH	basic helix-loop-helix
BSA	bovine serum albumin
CAP350	centrosome associated protein 350 kDa
CDK	cyclin-dependent kinase
Da	Dalton (g mol^{-1})
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EB1	end-binding protein 1
EDTA	ethylenediaminetetraacetic acid
E. coli	Esherichia coli
FGFR1	fibroblast growth factor receptor 1
HAT	histone acetyltransferase
HDAC	histone deacetylase
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl-β-thiogalactopyranoside
ITC	isothermal titration calorimetry
K_A	association constant
K_D	dissociation constant
kDa	kilodalton
LB	Luria-Broth medium
LisH	LIS1 homology motif
MAD	multiwavelength anomalous diffraction
MAPK	the mitogen-activated protein kinase
MIR	multiple isomorphous replacement

MR	molecular replacemen
MTs	microtubule
MW	molecular weight
NC	nitrocelulose
NiNTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PAI	plasminogen activator inhibitor
PBS	phosphate buffered saline
PCM	pericentriolar material
PEG	polyethylene glycol
PMSF	phenylmethylsulfonylflouoride
ppm	parts per million
RMSD	root mean square deviation
SDS	sodium dodecyl sulphate
SIR	single isomorphous replacement
STAT	signal transducer and activator of transcription
TEMED	N,N,N',N'- tetramethylethylendiamine
Tris	tris(hydroxymethyl)aminomethane

Amino acids and nucleotides are abbreviated according to either one or three letter IUPAC code.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 The cytoskeleton

Dynamic remodeling of the cytoskeleton is responsible for many cell behaviors such as cell migration, its shape and mechanical strength, as well as the organelles and chromosomes segregation. All these processes have to be precisely regulated in response to external and internal signals (Dormann and Weijer 2006). The cytoskeleton of higher-eukaryotic cells is made up of three kinds of protein filaments: actin filaments (also called microfilaments), which average 8 nm in diameter, intermediate filaments - 10 nm in diameter, and microtubules - 25 nm, as well as a growing number of associated proteins. Within the digestive epithelia, members of these three protein families are actins, keratins and tubulins, respectively (Ku et al. 1999). The relative motion between filaments are induced by motor proteins (Kruse and Julicher 2006). One of the most important cytoskeletal activities is the mitotic spindle formation, which is regulated by a microtubule network (Nigg 2006). MTs dynamics are controlled by coordinated interactions with actin cytoskeleton and proteins that form adhesion sites, as well as by kinases and phosphatases via signaling cascades. There are two major groups of microtubule motors: kinesins (most of these move toward the plus end of the microtubules) and dyneins (which move toward the minus end). In animal cells, the microtubules originate at the centrosome.

1.1.1 The centrosome

The centrosome is the major microtubule-organizing center of animal cells. It influences not only the formation of the bipolar mitotic spindle but also determines the cell shape and polarity in interphase (Loffler et al. 2006). In addition, the centrosomes are also required for cytokinesis, segregation of signaling molecules, and developing some of the neurons (Andersen et al. 2003). Each mammalian somatic cell typically contains one centrosome, which is duplicated in coordination with DNA replication; thus the proper centrosomal cycle is tightly connected with cell cycle regulation. Both processes require cyclin-dependent kinase 2 (CDK2), in association with cyclins E and A, polo-like kinase 4 Plk4, as well as phosphorylated retinoblastoma protein pRb (Meraldi et al. 1999; Bettencourt-Dias et al. 2005; Habedanck et al. 2005). Numerical and structural centrosome aberrations have been implicated in cancer. Abnormalities in the centrosomal number were present in

mammalian cells with deregulation of BRCA1, BRCA2 and mitotic kinase Aurora A genes. Also the lack of the p53 tumor suppressor protein was connected with the centrosome anomaly (Nigg 2006).

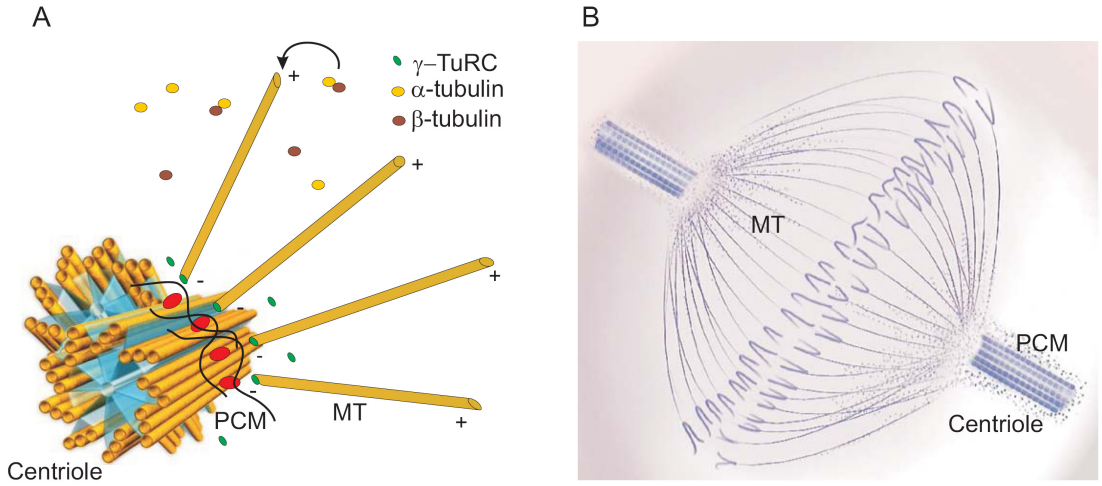


Figure 1.1. A) The centrosome is composed of two centrioles surrounded by the amorphous pericentriolar material (PCM). The primary function of the pericentriolar material is to nucleate microtubules (MTs) from γ-tubulin ring complexes (γ-TuRCs). Microtubules are polymers of α- and β-tubulin dimers growing away from the γ-TuRCs with their plus ends (+) expanding to the cytoplasm. B) Schematic view of the mitotic spindle. The spindle assumes a characteristic form as some microtubules are bound by chromosomal domains called kinetochores. The kinetochore of one sister chromatid attaches to the microtubules extending from one pole whilst second sister kinetochore attaches to those extending from the other.

A single centrosome is composed of two centrioles, the cylindrical arrays of nine triplets of microtubules that are surrounded by the amorphous pericentriolar material (PCM). The primary function of the pericentriolar material is to nucleate microtubules from so-called γ-tubulin ring complexes (γ-TuRCs) (Wiese and Zheng 2000). As mitosis proceeds microtubules grow out from γ-TuRCs by the polymerization of tubulin dimers (α- and β-tubulins) with their plus ends expanding toward the metaphase plate. The process is powered by the hydrolysis of GTP. Subsequent to nucleation MTs are released, so to remain associated with centrosomes; they need to be captured by centrosomal MT anchoring activities at their minus end (Figure 1.1) (Dammermann et al. 2003). A number of proteins, including ninein, $p150^{Glued}$, EB1, CAP350, and FOP have been implicated in centrosomal MT anchoring, but the process is far from understood (Yan et al. 2005).

1.1.1.1 Fibroblast growth factor receptor 1 oncogene partner (FOP)

The fibroblast growth factor receptor 1 (FGFR1) oncogene partner (FOP) is ubiquitously expressed in various tissues and organs and has a putative role as a regulator of normal proliferation and differentiation of the erythroid lineage. FOP is a hydrophilic protein, which contains, in its N- and C-termini, several leucine rich regions with consensus sequences L-X_2-L-X_{3-5}-L-X_{3-5}-L (in some of them the leucine is substituted by either a valine or an isoleucine) (Popovici et al. 1999). Because these motifs differ from the rest of known proteins possessing leucine consensus repeats (Kobe and Deisenhofer 1994; Kajava 1998), FOP may belong to a novel family of the leucine-rich proteins (Popovici et al. 1999). However, the precise function of the full-length protein is not clear. FOP was found to be coexpressed with PLK1 and CDC25A, which may indicate functional correlations to these proteins (Jensen et al. 2004; von Mering et al. 2005). The CDC25A protein (M-phase inducer phosphatase 1) regulates both the G1 to S and the G2 to M phase transitions by activating cyclin-dependent kinases (Zou et al. 2001). A *Drosophila* CDC25 homologue was reported to be required for the completion of the daughter centriole assembly (Vidwans et al. 1999) and dysregulation of CDC25A transcription by the human papillomavirus type-16 E7 oncoprotein can induce a centrosome duplication error (Duensing and Munger 2003). PLK1 regulates many aspects of mitosis, including centrosome maturation, chromosome cohesion, as well as mitotic entry and exit (Donaldson et al. 2001).

FOP is involved in stem cells myeloproliferative disorder (MPD), a disorder which is characterized by uncontrolled proliferation of progenitor cells (Aguiar et al. 1995; Macdonald et al. 1995). MPD is caused by gene translocations that produce protein fusions between FGFR1 and, for example, FIM/ZNF 198 (zinc finger 198), (Popovici et al. 1998; Reiter et al. 1998; Xiao et al. 1998) CEP1 (centrosomal protein 1) (Guasch et al. 2000), BCR (Demiroglu et al. 2001; Fioretos et al. 2001), or FOP (Popovici et al. 1999; Delaval et al. 2005 a). These translocations are associated with such diseases as lymphoma, myeloid hyperplasia, and eosinophilia (Vizmanos et al. 2004). In each case, the N-terminal protein-protein interaction domain of the partner is fused to the tyrosine kinase domain of FGFR1. All resulting fusion proteins have constitutive kinase activity caused by dimerization of the FGFR1 partner (Ollendorff et al. 1999; Xiao et al. 2000; Guasch et al. 2001). The FOP-FGFR1 fusion protein was identified through the translocation between the chromosomal region 6q27 that encodes the FOP protein and 8p11-12, which encodes FGFR1. FOP-FGFR1 is thought to promote cell proliferation and protect cells from apoptosis through the

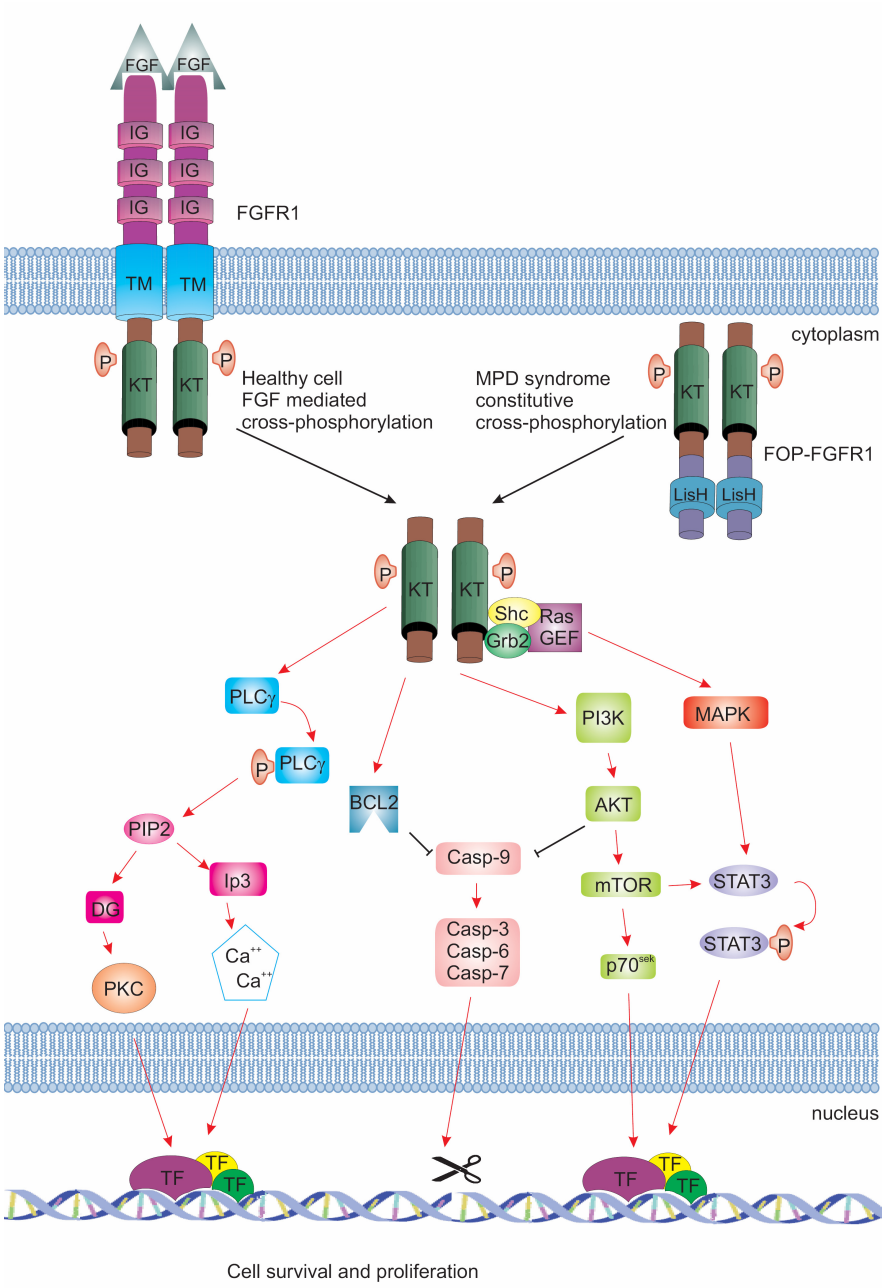


Figure 1.2. The fusion protein composed of the leucine-rich N-terminal region of FOP and the catalytic domain of FGFR1 is capable of protecting cells from apoptosis by using the same effectors as the wild-type FGFR1. Phosphorylation of STAT1 and of STAT3, but not STAT5, was observed in cells expressing FOP-FGFR1. FOP-FGFR1 is also responsible for phosphorylation of phosphatidylinositol 3 (PI3)-kinase and AKT, as well as for induction of the survival signals, including BCL2 overexpression and inactivation of caspase-9. The FOP-FGFR1-expressing cells show constitutive phosphorylation of the positive regulator of translation the p70S6 kinase; that is inhibited by the PI3-kinase and mTOR (mammalian target of rapamycin) inhibitors. The mitogen-activated protein kinase (MAPK) is also activated in FOP-FGFR1-expressing cells and confers the cytokine-independent survival to hematopoietic cells. PLC-γ associated with FOP-FGFR1 is constitutively tyrosine phosphorylated. Phosphorylated PLC-γ hydrolyzes phosphatidylinositoldiphosphate PIP_2 into diacylglycerol (DG) and inositoltriphosphate (IP3). DG is

necessary for further activation of protein kinase C (PKC) while IP3 leads to the release of intracellular calcium. KT: the tyrosine kinase domain; TM: a transmembrane domain; IG: an immunoglobulin-like domain; FGF: fibroblast growth factor; LisH: the LIS1 homology motif; TF: a transcription factor.

same signaling pathways that are utilized by wild type FGFR1 (Figure 1.2), which involve phospholipase C γ (PLC-γ) (Foehr et al. 2001), the mitogen-activated protein kinase (MAPK) (Cross et al. 2002), phosphatidylinositol 3-kinase (PI3K), AKT (Chang et al. 2003), and STAT1 and STAT3 kinases (Popovici et al. 1998; Guasch et al. 2001; Browaeys-Poly et al. 2005; Delaval et al. 2005 b). FOP-FGFR1 is also indirectly responsible for inhibition of caspase-9, which is a prominent effector of apoptosis (Cardone et al. 1998).

Both the FOP-FGFR1 fusion kinase and FOP alone, but not FGFR1, were found to localize to the centrosome (Andersen et al. 2003; Yan et al. 2005). These results suggest that the N-terminal fragment of FOP not only provides dimerization but also controls the localization of the active kinase (Delaval et al. 2005 b). The FOP-FGFR1 substrates at the centrosome could either be signaling molecules, recruited and phosphorylated at the centrosome (e.g., STAT, PLCγ), or intrinsic centrosomal proteins. Constitutively activated signaling pathways at this organelle could disrupt either centrosome duplication and structure or cellular processes associated with the centrosome, notably the regulation of cell shape, polarity, motility and cell cycle progression (Arlot-Bonnemains and Prigent 2002; Metge et al. 2004; Delaval et al. 2005 a; Delaval et al. 2005 b). It is also noteworthy that the N-termini of FOP show regional similarities with tonneau proteins, which are associated with the cortical cytoskeleton identified in *Arabidopsis thaliana*. This relationship suggests that FOP and FOP-related proteins play conserved roles in cell organization in both animals and plants (Camilleri et al. 2002; Delaval et al. 2005 b).

The N-terminal part of FOP (residues 69-102) harbors a LIS1 homology (LisH) motif. LisH-like domains are found in over 100 proteins, such as, for example, LIS1, muskelin, Nopp140, and katanin p60 (Emes and Ponting 2001). LisH motifs are thought to contribute to the regulation of microtubule dynamics, either by mediating dimerization, or else by binding to the cytoplasmic dynein heavy chain, or microtubules directly (Emes and Ponting 2001; Umeda et al. 2003). Importantly, the LisH motif has been found in a number of gene products that are associated with human diseases, including Miller-Dieker lissencephaly (Pancoast et al. 2005), Treacher Collins (Dixon et al. 2004), oral–facial–digital type 1 (Ferrante et al. 2001), the contiguous syndrome ocular albinism with late

onset sensorineural deafness syndromes (Bassi et al. 1999), and the stem cell myeloproliferative disorder (MPD) (Popovici et al. 1999). Although the LisH motif is often found together with other protein domains, such as WD repeats, SPRY, Kelch, AAA ATPase, RasGEF, or HEAT, in the FOP protein the LisH is the only identifiable signature (Emes and Ponting 2001).

1.2 Cell differentiation and cell cycle

In multicellular organisms cell proliferation, differentiation and death are strictly regulated to maintain homeostasis. Disruption of this equilibrium by loss of cell cycle control may eventually lead into tumor development (Pucci et al. 2000). The cell cycle is an alteration of two main processes: the S - synthesis phase where DNA is synthesized, and M - mitosis phase where the cell and its contents are divided equally into two daughter cells. The periods between these processes are called gap periods (G phases). During G_1 phase cell is growing and the chromosomes are prepared for replication. G_2 phase is a preparation for mitosis. G_0 is a temporary or permanent state of the cell cycle exit that occurs always in G_1 phase (Sandal 2002).

The two most important protein families responsible for cell cycle progression are cyclin-dependent kinases (Cdks) and cyclins that are switched on and off in different stages of the cell cycle (Harbour and Dean 2000; Giacinti and Giordano 2006). G_1 cyclins D and E mediate progression through the G1/S phases. Cyclin D is coupled with Cdk4 or Cdk6 and interacts with the AB pocket of the retinoblastoma protein (pRb) (Dowdy et al. 1993). Cdk4/6 and Cdk2 phosphorylate pRb that releases the E2F protein from the pRb-E2F complex, which permits the transcription of genes necessary for S-phase (Cobrinik 2005). The mechanism of conformational changes in the pRb structure after phosphorylation is not precisely known but the main hypothesis suggest the C-terminal domain of pRb to be a first target for phosphorylation by cyclin D/Cdk4 or Cdk6. This process facilitates the phosphorylation of the AB pocket region by cyclin E (A)/Cdk2 (Giacinti and Giordano 2006). Cyclins A and B coupled with Cdk1 and Cdk2 mediate the progression through the S/G2/M phases. A mitogenic transcription activation of D-type cyclins has been shown to be induced, for instance, by c-Myc, AP-1, and NF-κB transcription factors (Ekholm and Reed 2000). Physiological signals directly induce expression of different types of cyclins D, so proliferation of various cell types can be modulated by different mitogens (Matsushime et al. 1991). The Cdk activity is also regulated by the Cdk-inhibitory subunits

(CKIs) that in mammalian cell are divided into two classes: Cip/Kip and Ink4 (Harper and Elledge 1996). The Cip/Kip family includes p21, p27 and p57 and predominantly inhibits the Cdks of the G1 to S phase transition. The Ink4 inhibitors (inhibitors of Cdk4) contain p15, p16, p18 and p19 (Figure 1.3). All of them provide a tissue specific mechanism by which the cell cycle can response to extra and intracellular signals (Pucci et al. 2000).

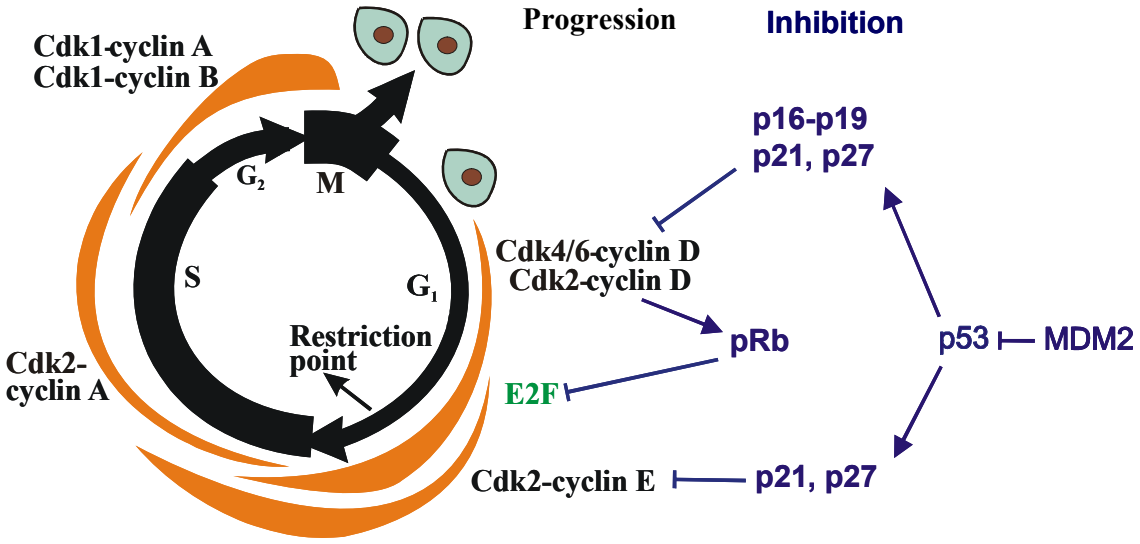


Figure 1.3. The cell cycle is precisely regulated by cyclins that are synthesized and degraded at specific points during the cell cycle, as well as Cdk's (cyclin dependent kinases) and Cdk's inhibitors. The lines thickness corresponds to the level of activity of given kinases.

1.2.1 The retinoblastoma protein, pRb

The retinoblastoma protein (pRb) is a tumor suppressor which plays a pivotal role in the negative control of the cell cycle. pRb is responsible for a major G_1 checkpoint, blocking the S-phase entry and cell growth through the regulation of E2F-responsive genes or through binding to proteins that modify chromatin structure (Figure 1.4) (Cobrinik 2005; Giacinti and Giordano 2006). There are two mechanisms by which pRb regulates the E2F-mediated gene transcription. Hypophosphorylated pRb can directly bind to the transactivation domain of E2F, which leads to E2F deactivation, or complex of both proteins can bind to the promoter of genes responsible for the G1/S phase transition (Giacinti and Giordano 2006). In addition, the retinoblastoma protein can regulate cell cycle by a E2F independent mechanism. pRb was found to inhibit Cdk activity and G1/S progression by increasing the expression and stability of p27 (Ji et al. 2004). It was also

reported that pRb is responsible for suppressing Ras signaling and the increased expression of genes responsible for differentiation (Lee et al. 1999). This is in agreement with reports illustrating importance of the retinoblastoma protein in activation of the myogenic program. pRb recruits the deacetylase HDAC1 to repress transcription of certain E2F-dependent genes in cells induced to the growth arrest (Brehm et al. 1998; Luo et al. 1998; Magnaghi-Jaulin et al. 1998). Interaction between hypophosphorylated pRb-HDAC1 coincides with disassembling of the MyoD-HDAC1 complex. Because MyoD, together with HDAC1, mediate the repression of a muscle-specific gene expression, disruption of this complex induces transcriptional activation of muscle-restricted genes, and cellular differentiation of skeletal myoblasts (Puri et al. 2001).

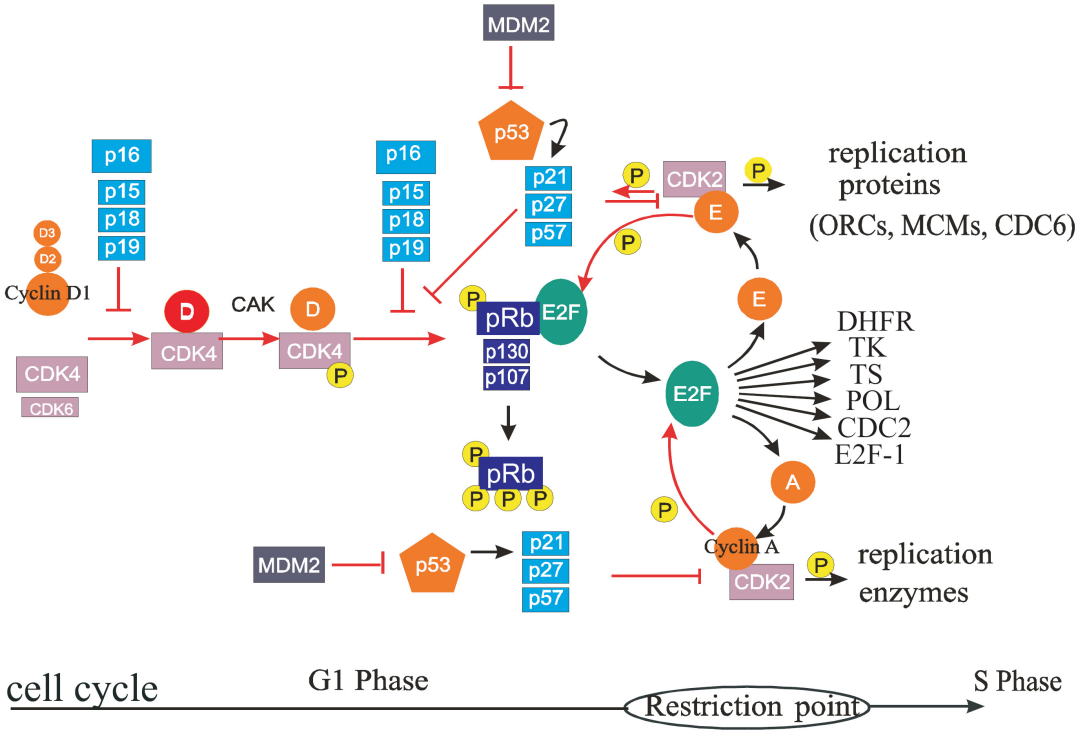


Figure 1.4. Network of protein – protein interactions on the onset of the G_1 phase of the cell cycle serves as a precise switch. If a cell is stimulated by myogenic signals from the environment it allows passing through the G1 restriction point, which causes, for example, the expression of the proteins crucial for the DNA synthesis in the S phase. Functions and modes of action of selected proteins involved in the cell cycle regulation are described in the text.

The mechanism by which pRb influences apoptosis is not well defined. E2F has been shown to mediate apoptosis following DNA damage, by inducing expression of the pro-apoptotic factor Apaf-1 (Blattner et al. 1999; Moroni et al. 2001). E2F cannot induce

apoptosis when pRb is co-expressed and possibly the anti-apoptotic effect of pRb is mediated through the inhibition of E2F (Fan et al. 1996; Pucci et al. 2000). p53 and pRb/E2F may be directly linked in cell proliferation and apoptosis (Fan et al. 1996; Pucci et al. 2000). Activated p53 causes a G1 arrest by inducing p21, followed by an inhibition of cyclin/CDK (Sandal 2002). In these conditions, pRb is not hyperphosphorylated and cells do not progress through the cell cycle (Harbour et al. 1999). In contrast, free E2F directly induces p53 transcription, thus connecting the pRb/E2F pathway to the p53-dependent apoptosis (Hiebert et al. 1995). Each of both tumor suppressors (p53 and pRb) may thus be able to compensate for the loss of the other (King and Cidlowski 1998).

1.2.2 Viral oncoproteins

Three DNA tumor virus-transforming proteins, including SV40 T antigen, adenovirus E1A, and human papilloma virus E7 are involved in processes of cell transformation and oncogenesis (Felsani et al. 2006). They bind to unphosphorylated pRb and reverse its growth suppressive function (DeCaprio et al. 1988; Whyte et al. 1988). The E7 protein is the major oncogenic protein produced by cervical cancer-associated HPV16 (Enzenauer et al. 1998). Human papilloma viruses (HPV) are a family of DNA viruses that can infect, prolong the life, and transform epithelial cells, particularly in the skin, head, neck and, cervix (Morozov et al. 1997; Enzenauer et al. 1998). The natural host of SV40 is the *Rhesus macaque*, however, SV40 tumor T antigen sequences and protein have been detected in human cancers and found to be associated with pRb and p53 (Borger and DeCaprio 2006; Caracciolo et al. 2006). The E1A protein is central to the regulation of adenoviral gene expression and viral replication by inactivating the function of pRB and transactivating late gene promoters (Li et al. 2005). All of these viral oncoproteins have two conserved regions in their sequences called CR1 and CR2, which are the regions of interaction with the pRb. The LXCXE sequence, which is thought to be a primary binding motif for the interaction with pRb, is located within the CR2 region of E7 and E1A oncoproteins, and in the CR1 region of the T antigen (Morris and Dyson 2001; Miura et al. 2003).

1.2.3 Histone deacetylase HDAC1

Transcriptional regulation requires not only the network of transcription factors but also chromatin remodeling (Cheung et al. 2005). Histone modifications, such as

phosphorylation, acetylation or metylation, change chromatin structure, which increases its accessibility that is required not only for transcription activity but also for DNA repair (Cheung et al. 2005). Deacetylation of histones by histone deacetylases (HDACs) catalyzes the removal of acetate from modified lysine residues located in the N-terminal tail region of the core histones H2A, H2B, H3 and H4. This process is correlated with the transcriptional repression (Luo et al. 1998). There are three classes of HDACs; class 1 containing HDAC1, 2, 3 and 8, class 2: HDACs 4, 5, 6, 7, and 9. The third class includes the NAD-dependent group of HDACs (Gray and Ekstrom 2001; Khochbin et al. 2001). pRb binding has been detected with a number of these proteins suggesting that pRb can represses S phase genes through the recruitment of an enzyme which modifies DNA structure, for example the histone deacetylase HDAC1 (Takaki et al. 2004). The other pocket proteins, p107 and p130, have also been shown to bind HDAC1 (Ferreira et al. 1998). pRb physically interacts with HDAC1 to cooperatively repress the E2F-mediated gene expression (Brehm et al. 1998; Luo et al. 1998; Magnaghi-Jaulin et al. 1998). The small pocket of pRb was found to be necessary and sufficient for HDAC1 binding and transcriptional repression. Further analysis of pRb/HDAC interactions demonstrated that endogenous HDAC1, 2, and 3 bound pRb *in vitro* and coimmunoprecipitated with the endogenous, hypophosphorylated pRb from H1299 cell extracts *in vivo* (Lai et al. 1999; Nicolas et al. 2000). HDAC1 and 2, unlike HDAC3, possess the LXCXE motif that was found in many pRb binding proteins. Thus it was suggested that HDACs can also interact with pRb by other linker proteins, such as RBP1 (Lai et al. 1999). It was also found that HDAC1 in complex with MDM2 can influence the tumor suppressor protein p53 stability by deacetylation its lysines residues (Ito et al. 2002).

1.2.4 Plasminogen activator inhibitor-2

The plasminogen activator inhibitor-2 (PAI-2) is a serine protease inhibitor that belongs to the ovalbumin-like serpins (ov-serpins) branch of the serpin superfamily (Remold-O'Donnell 1993). It is expressed in activated monocytes and macrophages, differentiating keratinocytes, and many tumors. PAI-2 was first described as an inhibitor of the urokinase-type plasminogen activator (uPA), an extracellular serine proteinase that cleaves plasminogen to plasmin (Kruithof et al. 1995). Plasmin facilitates cell migration and invasion by degrading fibrin and activating metalloproteinases. This plasminogen-dependent proteolysis occurs not only in inflammatory responses, but also in tissue

remodeling, fibrin deposition and lysis, wound healing, and tumor metastasis (Vassalli et al. 1991). Most serpins usually undergo a major structural change from the stressed state to an extremely stable form (the relaxed state) in order to trap the protease in a nonfunctional covalent serpin-protease complex. The target protease cuts the accessible reactive center loop (RCL) on the serpin to produce a covalent acyl intermediate. The RCL then inserts into β-sheet A to form the relaxed state. Attached to the RCL, the target protease moves to the base of the serpin, where it is partially unfolded (Huntington et al. 2000; Jankova et al. 2001). Serpins act as "suicide inhibitors" that react only once with their cognate protease, by this action both serpin and protease are deactivated (Lawrence et al. 1995).

PAI-2 has a unique insertion of 30 residues between helices hC and hD (CD loop), which is the site of glycosylation and the site of cross-linking to the cell membrane; it was also found to bind cytoplasmic proteins (Jensen et al. 1996; Jankova et al. 2001). Similar to some other ov-serpins, PAI-2 lacks a typical amino-terminal secretion sequence which is connected to additional, intercellular functions, distinct from protease inhibition (Darnell et al. 2003). It was reported that PAI-2 is induced in monocytes/macrophages in response to TNF-α where it protect cells from the TNF-α-mediated apoptosis. This can suggest that the physiological role of intracellular PAI-2 in inflammatory macrophages is the protection from the cytotoxic effects of their own TNF-α (Dickinson et al. 1995). PAI-2 was also induced in macrophages in response to viral RNA. It can prevent rapid viral infection by the induction of the interferon (IFN)-α/β production, which stimulates cells for the transcription of antiviral genes (Antalis et al. 1998). PAI-2 was thus postulated to modulate resistance to apoptosis and to regulate the transcription as an intracellular regulator of the signal transduction pathway (Antalis et al. 1998; Darnell et al. 2003). Recent data suggest one more intracellular role of PAI-2 as a pRb binding partner. Immunofluorescence, combined with *in vitro* experiments, indicated that PAI-2 and pRb colocalize in the nucleus where PAI-2 protects the retinoblastoma degradation mediated by the E7 oncoprotein, mostly by a direct interaction with the C-terminal domain of pRb. Full-length PAI-2, but not RLS and CD mutants, was required for the pRb and p53 restoration in HeLa cells (Darnell et al. 2003).

1.2.5 Helix-loop-helix (HLH) protein family

The proteins of the HLH family are crucial in coordinating the differentiation processes with the cell cycle. The HLH family includes both positive and negative regulators of cell

differentiation: basic HLH (bHLH)-type transcription factors and Id proteins, respectively. Because of the large number of HLH proteins that have been described, a classification scheme has been based upon their tissue distribution, dimerization capabilities, and DNA-binding specificities (Massari and Murre 2000). The class A basic HLH (bHLH), also known as the E proteins, such as those encoded by differentially spliced transcripts from the E2A (E12, E47 and E2-5/ITF1 proteins), E2-2/ITF2, and HEB/HTF genes, are ubiquitously expressed (Murre et al. 1989; Henthorn et al. 1990; Massari and Murre 2000). The class B bHLH proteins, which include members such as MyoD, myogenin, NeuroD/BETA2, and TAL, show a tissue-restricted pattern of expression. Differentiation in many different lineages is regulated by specific subfamilies of the bHLH proteins. Analogous to the MyoD sub-family of bHLH proteins, the Neurod sub-family of bHLH proteins regulates neuronal specification and differentiation (Bertrand et al. 2002). The E-protein sub-family of bHLH proteins (Tcf3, Tcf4 and Tcf12) has a crucial role in lymphocyte differentiation (Engel and Murre 2001; Greenbaum and Zhuang 2002). Dimerization is essential for DNA binding and transcriptional activity *in vivo* (Lassar et al. 1991; Massari and Murre 2000) and in general, class B bHLH proteins form heterodimers with the class A bHLH proteins, although the latter can also operate as homodimers (Shen and Kadesch 1995). The E-protein family members function as heterodimer partners for many of the tissue-restricted bHLH proteins, such as MyoD and Neurod proteins (Massari and Murre 2000). The basic region of each protein is required for binding to DNA, commonly to a region that includes a specific sequence motif known as the E-box (CANNTG) (Ephrussi et al. 1985; Lassar et al. 1989; Tapscott 2005) or the related N-box (CACNAG) (Tietze et al. 1992). A group of HLH proteins that lack a basic region, including Id and emc, define the class D of the HLH proteins (Massari and Murre 2000).

1.2.5.1 Id-2

Id proteins are classified as negative regulators of cell differentiation and positive regulators of cell proliferation. These proteins promote cell proliferation by binding to transcription factors of the bHLH subclass of proteins and inhibit their ability to bind DNA (Mori et al. 2000). They have also been shown to stimulate the G_1/S transition in the cell cycle in several cell culture systems, implying their roles in cell proliferation (Norton et al. 1998). The functions of Id proteins are integrated with cell-cycle-regulatory pathways coordinated by cyclin-dependent kinases and the retinoblastoma protein.

Id-2 is able to reverse the antiproliferative effects of tumor suppressor proteins of the pRb family (pRb, p107, and p130), thus allowing cell-cycle progression (Sandal 2002). Although the HLH region is well conserved among Id family members, only Id–2, and not other Id proteins, was shown to interact directly with pRb, p107, p130 (Iavarone et al. 1994; Lasorella et al. 1996). An Id–2 construct lacking the HLH domain did not prevent progression through S phase caused by pRb. Hypophosphorylated pRb was able to bind to a region of Id-2 corresponding to only the HLH domain. However, pRb containing mutations within the E1A/large T-binding pocket did not bind Id-2 suggesting that the interaction requires an intact small pocket domain of pRb (Iavarone et al. 1994).

1.2.5.2 MyoD

MyoD is a main regulatory protein of skeletal muscle differentiation belonging to a group of muscle specific basic helix-loop-helix (bHLH) transcription factors (Tapscott 2005). In a complex regulatory network with myf-5, myogenin, and MRF4, MyoD plays a key role in the determination and differentiation of all skeletal muscle lineages (Weintraub et al. 1991; Arnold and Braun 1996; Arnold and Winter 1998). It can also induce skeletal muscle differentiation in cells from many different lineages (fibroblast, fat, melanoma, neuroblastoma, chondroblast, liver and retinal pigmented epithelial cell lines) (Davis et al. 1987; Weintraub et al. 1989; Tapscott 2005).

The binding and activity of MyoD are regulated by many factors to achieve a temporal patterning of gene expression. The model of the transcriptional activity of the myogenic bHLH proteins is that they activate gene transcription by binding to the E-boxes in the regulatory regions of genes that are expressed in the skeletal muscle. The presence of certain binding sites paired with an E-box can be responsible for promoter-specific activity of MyoD, depending on the availability of the cooperating transcription factors. In the case of MyoD the adjacent basic regions in an α-helical conformation contact the DNA. MyoD also contains the “myogenic code” which consists of three residues that are conserved in all of the myogenic bHLH proteins (MyoD, Myf5, MyoG and Mrf4)(Tapscott 2005). “Myogenic code” is necessary to activate the transcription of specific muscle genes by either interacting with co-factors or inducing conformational changes (Bengal et al. 1994; Ma et al. 1994), or both, but not affect directly DNA binding. It was reported that an interaction with Mef2 factors is mediated by this region (Molkentin et al. 1995). Altering the conformation of the bound protein in a manner that presents other regions for a co-

factor interaction has been suggested for the promoter specific activity of NF-κB (Leung et al. 2004). MyoD has also a single transcriptional activation domain (AD), and a histidine- and cysteine-rich (H/C) region that contains a tryptophan residue that is needed for MyoD to interact with the Pbx/Meis complex (Tapscott 2005).

Growth arrest required functional pRb (Schneider et al. 1994; Novitch et al. 1996). *In vitro* protein binding and immunoprecipitation studies indicated that both MyoD and myogenin bind to hypophosphorylated pRb directly through the bHLH domain (Gu et al. 1993). There has been however discrepancy among experimental reports about the MyoD and pRb interaction. There are publications reporting both *in vivo* and *in vitro* interactions (Gu et al. 1993), other showing that there is no interaction whatsoever (Zhang et al. 1999; Zhang et al. 1999; Puri et al. 2001). It was also shown (Iavarone et al. 1994) that GST-Id-2 and GST-MyoD bind with comparable efficiency to the wild type pRb and that there is no binding found for the pRb C706F mutant.

1.3 X-ray protein crystallography

X-ray crystallography provides the direct images of protein molecules with detailed information about their structure at atomic level, their activity, binding with substrates, ligands and effectors as well as conformational changes. Three dimensional structures can show relationship between widely separated systems, without amino-acids sequence similarities that can give rise to protein taxonomy. NMR structures are determined in solution that gives a superposition of many possible protein conformations but protein size is a main limitation of the method (Yu 1999). In comparison with NMR, x-ray crystallography gives a more static description of the molecular structure but there are no limits in the size of the molecule to be analyzed.

1.3.1 The unit cell, space group and symmetry

A crystal is a regular, tree dimensional repeating array where the fundamental building block is a unit cell. A **unit cell** can be described by three vectors a, b and c which are related by angels α, β, and γ. The crystal lattice can be built up by translating the unit cell along each of the lattice vectors. The **asymmetric unit** that is always smaller than unit cell is sufficient to create the whole lattice if we use both translations and symmetry operators. The geometry of the unit cell together with the possible symmetry operations such as translations, rotations, mirrors and inversion centres, defines the **space group** of the

crystal. Macromolecules like proteins and DNA or RNA are chiral because they are composed of enantiomers (L-amino acids and D-sugars, respectively). Therefore, crystals from these molecules cannot possess an inversion centre or a mirror plane. As a consequence, the number of possible space groups is reduced from 230 to 65, which are distributed between 7 crystal systems: triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic. The combination of the 4 crystal lattices: primitive (P), body centred (I), face centred (F) and centred (C) with the mentioned 7 crystal systems allows a total of 14 Bravais lattices. Identification of the correct space group is essential for indexing of diffraction patterns, which is the first step in understanding the crystal structure.

1.3.2 The Bragg's law

Bragg´s law (Bragg and Bragg, 1913) interprets X-ray diffraction by a crystal lattice as a conjunct of reflections from different planes of atoms in the crystal (Figure 1.5). For a constructive interference, this can be written as follows:

$$\mathbf{2d\sin\theta = n\lambda}$$

where **d** is the distance between successive planes of atoms, **θ** is the angle of incidence of the X-rays that equals the angle of reflection, **n** is an integer, and **λ** is the wavelength of the X-rays (usually corresponding to the Cu Kα radiation = 1.5418 Å). Thus, as Bragg pointed out, X-ray diffraction can be regarded as the reflection of the beam of X-rays from the planes of points in the crystal lattice.

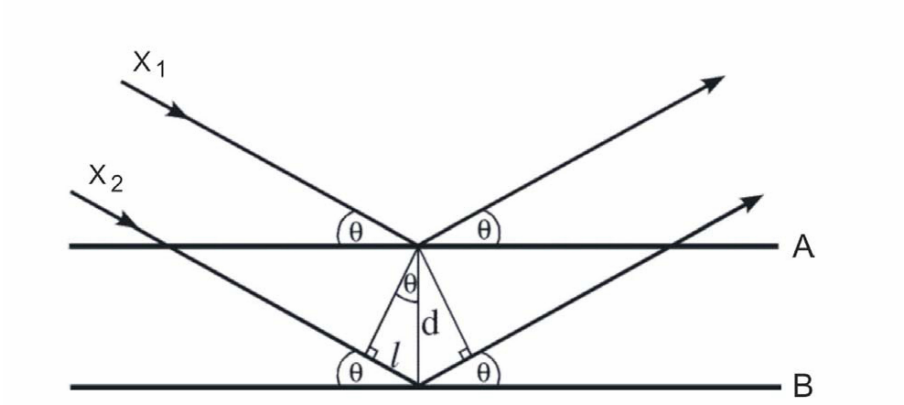


Figure 1.5. Scheme explaining Bragg's law. X-rays (X1, X2) that are reflected by lattice planes (A, B) with distance ***d*** have a difference in path length that is equal to ***2dsinθ***. A prerequisite for constructive interference is, that this difference in path is an integer multiple ***n*** of the used wavelength ***λ***.

1.3.3 Data collection and scaling

Data collection results in a (large) list of images. Each one representing a small wedge at the rotation of the crystal in the beam. Since the reflections are projected on a plane (alias the detector), they are distorted. Data processing has to reconstitute the original, undistorted 3–dimensional lattice of reflections. After data processing the data must be adjusted to a common scale because reflections with the same indices ought to have identical intensities. In practice there are a lot of factors that can modulate quality of the data set. Different regions of the detector can have different sensitivity; data set was collected from two different crystals or from two different detector positions; the beam can change from frame to frame and the crystal size and quality change as a result of exposition to X-ray beam. That is why the experimental differences in intensities necessitate the scaling of the data.

1.3.4 The phase problem

When waves are diffracted from a crystal, they give rise to diffraction spots. Each diffraction spot corresponds to a point in the reciprocal lattice and represents a wave with the amplitude and a relative phase. Knowledge of both, amplitudes and phases, allows the reconstitution of the electron density of the crystal. The amplitudes can be deduced from the intensities of the diffracted X-rays but the phases cannot be directly measured. This is known as the "phase problem". The following techniques are the most frequently used in determining phases:

- direct methods (small molecules and high resolution only)
- molecular replacement (MR)
- isomorphous replacement (SIR, MIR)
- anomalous dispersion (SAD, MAD)
- exploitation of radiation damage

Molecular replacement is the common method of solving the phase problem that depends on the availability of a sufficiently homologous model structure. Procedures for molecular replacement include self- and cross-rotation functions as well as translation searches which mean that the model structure must be placed in the correct orientation and position in the unknown unit cell. Three rotation angles are required to orient a molecule and three translational parameters are needed to place it in the unit cell. The usual strategy

is to look at these parts separately breaking six-dimensional search into two tree-dimensional ones.

Traditional molecular replacement methods are based on the properties of the Patterson function (map) of a protein crystal structure. The Patterson function is a vector map that is an inverse Fourier transform of the intensities (amplitudes squared), which requires only the measured data and can be computed without phase information. Peaks of the map are located at the positions of vectors between atoms in the unit cell. The vectors in the Patterson map can be divided into two categories. Intramolecular vectors (from one atom in the molecule to another atom in the same molecule) depend only on the orientation of the molecule, and not on its position in the cell, so these can be used in the rotation function. Intermolecular vectors depend both on the orientation of the molecule and on its position so, once the orientation is known, these can be exploited in the translation function. If the pairs of atoms belong to the same molecule, then the corresponding vectors are relatively short and their end-points are found not too far from the origin in the Patterson map; they are called self-Patterson vectors. If there were no intermolecular vectors (cross-Patterson vectors), this inner region of the Patterson map would be equal for the same molecule in different crystal structures, apart from a rotation difference. For homologous molecules it is not exactly equal but very similar. Therefore, the self-Patterson vectors can supply us with the rotation relationship between the known and the unknown molecular structures. The second step is to decide where the origin of the model molecule must be located to generate crystal structure of the measured molecule. A translation function can be used to find the axial position of any type of rotational symmetry operation. Initial model of the structure can be subsequently optimized by rigid body refinement. The target structure can be put through cycles of map calculation, model fitting and refinement, which help to reduce the bias introduced by the starting model.

Isomorphous replacement is one of the oldest methods of phasing for protein structures. It is based on the idea that introduction of a small molecule into a protein or nucleic acid crystal does not or hardly alter the structure of the macromolecule (this is what the term "isomorphous" refers to). If there are two crystals, one containing just the protein (native crystal) and one containing in addition bound heavy atoms (derivative crystal), diffraction data from both can be measured. Common heavy atoms are Hg (80e−), Pb (82e−), Au (79e−), Pt (78e−), or U (92e−). They can be incorporated by co-crystallization or by soaking. The differences in scattered intensities will largely reflect the scattering contribution of the heavy atoms, and these differences can be used to compute a Patterson

map. Because there are only a few heavy atoms, such a Patterson map will be relatively simple to deconvolute. The co-ordinates of the heavy metals can be derived also via direct methods. From the co-ordinates one can calculate structure factors - amplitude and phase for the heavy atoms.

In the next step Harker construction is used to solve the phases of a large structure from the phases and amplitudes of a small subset of the structure. If the heavy atom does not change the rest of the structure, then the structure factor for the derivative crystal ($\mathbf{F}_{PH}$) is equal to the sum of the protein structure factor ($\mathbf{F}_P$) and the heavy atom structure factor ($\mathbf{F}_H$).

$$\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_H$$

The representation of a structure factors is equivalent to thinking of a wave as a complex vector spinning around its axis. From the experiment we know only lengths of vectors $\mathbf{F}_P$ and $\mathbf{F}_{PH}$ from the measured reflection intensity $\mathbf{I}_P$ and $\mathbf{I}_{PH}$ but we would like to know the phase angle of vector $\mathbf{F}_P$. Additionally we know both length and angle of vector $\mathbf{F}_H$. According to previous equation:

$$\mathbf{F}_P = \mathbf{F}_{PH} - \mathbf{F}_H$$

We can solve this vector by representing the equation in the geometrical way.

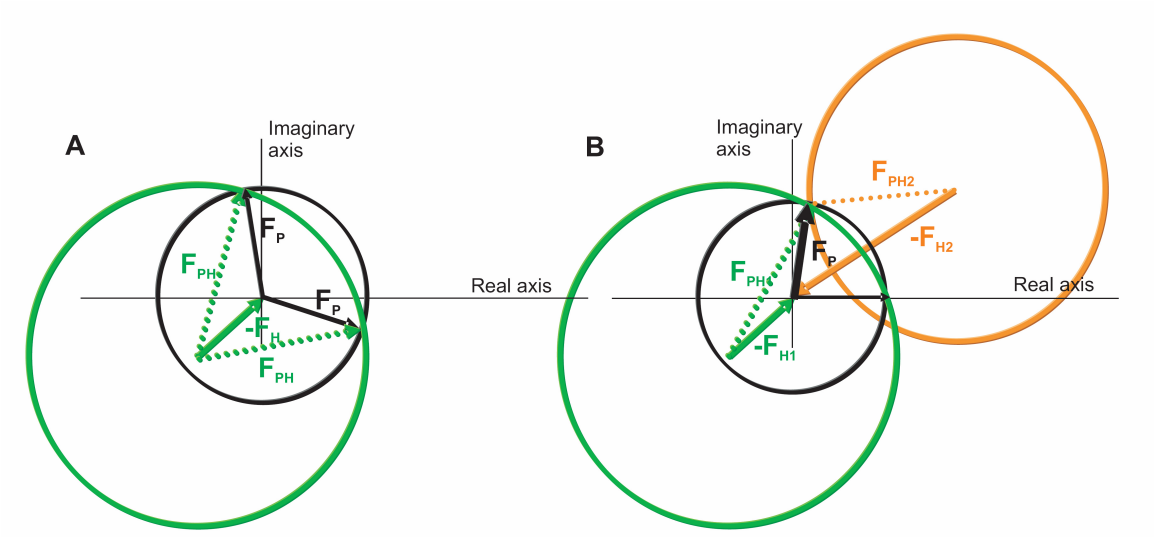


Figure 1.6. The Harker diagram for protein phase determination. **A)** Case where only one heavy atom is present. A circle of radius $|\mathbf{F}_P|$ representing the native protein is drawn (black) and from its centre the vector $-\mathbf{F}_H$. A second circle (in green) with radius $|\mathbf{F}_{PH}|$ is drawn at the endpoint of $\mathbf{F}_H$. The intersections of the two circles correspond to two equally probable protein phase angles. Both triangles fulfill the condition $\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_H$ **B.** Case where two different heavy atoms are present. In this situation it is easier to elucidate which is the correct solution because the drawn circles ideally intersect at the same position.

The two circles have two points of intersection from which one reads the two possible phases $\mathbf{F}_P$, which cannot be distinguished with a single isomorphous derivative. In principle, the twofold phase ambiguity can be removed by preparing a second derivative crystal with heavy atoms that bind at other sites. The information from the second derivative indicates that only one phase choice is consistent with all the observations. The need for multiple derivatives to obtain less ambiguous phase information is the reason for the term multiple isomorphous replacement (MIR).

A second means of obtaining phases from heavy-atom derivatives takes advantage of the heavy atom's capacity to absorb x-rays of special wavelengths. When the x-ray wavelength is near the heavy-atom absorption edge, a fraction of the radiation is absorbed by heavy atom and re-emitted with altered phase. As a result of this absorption Freidel's law does not hold, and the reflections hkl and –h-k-l are not equal in intensity. This inequality in symmetry related reflections is called **anomalous scattering or dispersion**. The atomic scattering factor (f) is the sum of "normal" atomic scattering factor f^0 and a complex "anomalous" correction having real (f') and imaginary (f'') components:

$$\mathbf{f(\lambda) = f_0 - \delta f'(\lambda) + \mathit{i} f''(\lambda)}$$

The real part of the atomic scattering factor ($f_0 - \delta f'$) is often written as f' so:

$$\mathbf{f = f' + \mathit{i} f''}$$

The vectors representing anomalous scattering contributions are $\Delta\mathbf{F}r$ (real) and $\Delta\mathbf{F}i$ (imaginary). The magnitude and anomalous scattering contributions $\Delta\mathbf{F}r$ and $\Delta\mathbf{F}i$ for a given element are constant, so these quantities can be looked up in tables of crystallographic information. The phases of $\Delta\mathbf{F}r$ and $\Delta\mathbf{F}i$ depend only upon the position of the heavy atom in the unit cell, so once the heavy atom is located by Patterson methods, the phases can be computed. Full knowledge of $\Delta\mathbf{F}r$ and $\Delta\mathbf{F}i$ allows establishing the phase of vector $\mathbf{F}_{PH}$ and subsequently $\mathbf{F}_P$.

Datasets of different wavelengths from a crystal that contain one or more atoms that give anomalous dispersion is called MAD (multi-wavelength anomalous dispersion). Data sets from a heavy atom derivative at different wavelengths are in many cases like those from distinct heavy atoms derivatives in MIR method, which exclude errors in phase determination. Data collected at the heavy atom absorption maximum, edge and at remote distance from the maximum all have distinct values for the real and imaginary

contributions of anomalous dispersion. In addition to wavelength dependent differences between Freidel pairs, individual reflections intensities vary slightly with wavelength (called dispersive differences). Both of these events contain phase information. One popular way to use MAD is to introduce selenomethionine in place of methionine residues in a protein. The selenium atoms (which replace the sulfur atoms) have a strong anomalous signal at wavelengths that can be obtained from synchrotron X-ray sources.

For more details about crystallization procedures please refer to (McPherson 1999) and for the basis of X-ray crystallography refer to (Blow 2002) and (Rhodes 2000).

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

All chemicals were of analytical grade, obtained from Merck (Germany) or Sigma Laboratories (USA), unless otherwise mentioned.

2.1.2 Enzymes

BSA 100 x Stock	New England BioLabs (USA)
Pfu Turbo DNA Polymerase	Stratagene (USA)
Dpn I	Stratagene (USA)
DAPzyme	Qiagen (USA)
Xa factor	Qiagen (USA)
Recombinant Enterokinase	Novagen (Canada)
Thrombin	Sigma (USA)

2.1.3 Kits and reagents

QIAquick PCR Purification Kit	Qiagen (Germany)
QIAprep Spin Miniprep Kit	Qiagen (Germany)
Quick – Change Mutagenesis Kit	Stratagene (USA)
Complete Protease Inhibitor Cocktail	Roche (Germany)

2.1.4 Oligonucleotides

Oligonucleotides used in cloning were provided by Metabion, or MWG (Germany).

2.1.4.1 Primers for FOP constructs cloning

FOP_s

5′ – GACGACGACAAGATGGCGGCGACGGCGGCCGCAGTGGTGGCC – 3′

FOP_as

5′ - GAGGAGAAGCCCGGTTTATGCAACATCATCTTCCAGATAATCCGC – 3′

FC2_s

5′- GACGACGACAAGATGGGTCGGGACCTGCTGGTGCAGACGCTGGAG -3′

FC2_as

5′- GAGGAGAAGCCCGGTTTATAGATCAAGTGCACCTTCCCCAGTGGT - 3′

FC3_s

5′ - GACGACGACAAGATGGGTAAAATTGGATCTTTTCTAAGCAACAGA – 3′

FC4_s

5′ - GACGACGACAAGATGGGTAAAAGTGGACTCAGCTCCCTGGCGGGA – 3′

FC2LS_s

5′ - GACGACGACAAGATGGCGGCGACGGCGGCCGCAGTGGTGGCCGAG – 3′

FC2S_as

5′ – GAGGAGAAGCCCGGTCTACCAGCGCCTGATCACTTCTAATAATAA – 3′

FC2L_as

5′ – GAGGAGAAGCCCGGTCTACCACTTACTTGGTGTTGTCTGTGCACT – 3′

2.1.4.2 Primers for FOP mutagenesis

FC2mut_s

5′ – ATGCATCACCACCATCACCACAAGAAATGGCGGGACCTGCTGGTG – 3′

FC2mut_as

5′ – CACCAGCAGGTCCCGCCATTTCTTGTGGTGATGGTGGTGATGCAT – 3′

LisH2_for

5′ – TCTTATACCATGGCGCATCACCACCACCATCACAAGAAAACTCCTTTAGTT – 3′

LisH2_rev

5′ – AACTAAAGGAGTTTTCTTGTGATGGTGGTGGTGATGCGCCATGGTATAAGA – 3′

FC2SXa_s

5′ – GAAAAAGTAGAGAACATTGAAGGCCGTAAAACTCCTTTAGTT – 3′

FC2SXa_as

5′ – AACTAAAGGAGTTTTACGGCCTTCAATGTTCTCTACTTTTTC – 3′

M1K_for

5′ – GAGAACATTGAAGGCCGTATGACTCCTTTAGTTAATGAG – 3′

M1K_rev

5′ – CTCATTAACTAAAGGAGTCATACGGCCTTCAATGTTCTC – 3′

M2K_for

5′ – TTAGTTAATGAGAGCCTGATGAAGTTTTTAAATACCAAA – 3′

M2K_rev

5′ – TTTGGTATTTAAAAACTTCATCAGGCTCTCATTAACTAA – 3′

M3K_for

5′ - AAAAAGTTTTAAATACCATGGACGGTCGTTTAGTGGCT - 3′

M3K_rev

5′ – AGCCACTAAACGACCGTCCATGGTATTTAAAAACTTTTT – 3′

M4L_for

5′ – TTTCTTCAGTTTTTTAACATGGACTTTACTTTGGCTGTT- 3′

M4L_rev

5′ – AACAGCCAAAGTAAAGTCCATGTTAAAAAACTGAAGAAA – 3′

M5L_for

5′ – ACTAGCACACTGCAAGGTATGGAAGGTCGAGAGAATTTA – 3′

M5L_rev

5′ – TAAATTCTCTCGACCTTCCATACCTTGCAGTGTGCTAGT – 3′

M6L_for

5′ – GGTACTGTGGGTGGACCCATGTTATTAGAAGTGATCAGG – 3′

M6L_rev

5′ – CCTGATCACTTCTAATAACATGGGTCCACCCACAGTACC – 3′

E97A_for

5′ - GCTGTTTTTCAACCTGCGACTAGCACACTGCAAGGT -3′

E97A_rev

5′ - ACCTTGCAGTGTGCTAGTCGCAGGTTGAAAAACAGC - 3′

T90A_for

5′ – TTTAACCTTGACTTTGCGTTGGCTGTTTTTCAACCT - 3′

T90A_rev

5′ - AGGTTGAAAAACAGCCAACGCAAAGTCAAGGTTAAA – 3′

TEdM_for

5′ - CTTGACTTTGCGTTGGCTGTTTTTCAACCTGCGACTAGCACA – 3′

TEdM_rev

5′ - TGTGCTAGTCGCAGGTTGAAAAACAGCCAACGCAAAGTCAAG – 3′

V74A_for

5′ - ACCAAAGACGGTCGTTTATTTGCTAGTCTTGTTGCA – 3′

V74A_rev

5′ - TGCAACAAGACTAGCAAATAAACGACCGTCTTTGGT – 3′

2.1.4.3 Primers for FOP-FGFR1 cloning

FR_as

5′ – GAGGAGAAGCCCGGTCATGTCAAGGCCACGATGCGGTCC – 3′

FR_s

5′ - GACGACGACAAGATGGCGGCGACGGCGGCCGCAGTGGTG – 3′

FOP_R_as

5′ – GAGGAGAAGCCCGGTCAGCGGCGTTTGAGTCCGCCATTGGCAAGC – 3′

2.1.4.4 Primers for FOP-FGFR1 mutagenesis

FRd_for

5′ - GGGCCAACCACTGGGGAAGGTGAGCTTCCCGAAGACCCTCGCTGG – 3′

FRd_rev

5′ – CCAGCGAGGGTCTTCGGGAAGCTCACCTTCCCCAGTGGTTGGCCC – 3′

2.1.4.5 Primers for CAP350 cloning

CAP_1_s

5′ - GACGACGACAAGATGGTGGATCATATCCTGGTTCAGGAG – 3′

CAP_2_s

5′ - GACGACGACAAGATGGGTCTGGATATTGAAAGCACTAGT – 3′

CAP_3_s

5′ - GACGACGACAAGATGGAAGAATTGGAGAAGCAACAGCAG -3′

CAP_as

5′ – GAGGAGAAGCCCGGTTCACACTAGTAGCATTCTCCCCTG – 3′

CAP4stop_for

5′ – GAGGAGGCACAGTGGTAGGTGAACTATGATGAGGATGAG – 3′

CAP4stop_rev

5′ – CTCATCCTCATCATAGTTCACCTACCACTGTGCCTCCTC – 3′

2.1.4.6 Primers for pRB cloning

RbCshort_s

5′ - GACGACGACAAGATGTACAAGTTTCCTAGTTCACCCTTACGG – 3′

RbClong_s

5′ - GACGACGACAAGATGTTGCAGTATGCTTCCACCAGGCCCCCT – 3′

RbC_as

5′ – GAGGAGAAGCCCGGTTCATTTCTCTTCCTTGTTTGAGGTATC – 3′

2.1.4.7 Primers for pRb mutagenesis

s_K713A

5′ – TGTTCCATGTATGGCATcTGtgcgGTGAAGAATATAGACCTT – 3′

as_K713A

5′ – AAGGTCTATATTCTTCACCGCACAGATGCCATACATGGAACA – 3′

s_C706F

5′ – TTGGACCAAATTATGATGTtTTCCATGTATGGCATA – 3′

as_C706F

5′ – TATGCCATACATGGAAaACATCATAATTTGGTCCAA – 3′

s_2mutCK

5′ – CAAATTATGATGTTTTCCATGTATGGCATCTGtGCCGTGAAGAAT – 3′

as_2mutCK

5′ – ATTCTTCACGGCaCAGATGCCATACATGGAAAACATCATAATTTG – 3′

2.1.4.8 Primers for E7 mutagenesis

s_30E72_T

5′ – GCTGAACCGGATCGTGCTTAAGGCTCTAACTCTCCTC – 3′

as_30E72_T

5′ – GAGGAGAGTTAGAGCCTTAAGCACGATCCGGTTCAGC – 3′

2.1.4.9 Primers for PAI2 cloning

PAI2_s

5′ – GACGACGACAAGATGGAGGATCTTTGTGTGGCAAACACACTC – 3′

PAI2_as

5′ – GAGGAGAAGCCCGGTTTAGGGTGAGGAAAATCTGCCGAAAAA – 3′

2.1.5 Strains and plasmids

2.1.5.1 Cloning strains

XL1-Blue	Stratagene (USA)
One Shot® TOP10	Invitrogen (Holland)

2.1.5.2 Protein expression strains

One Shot® BL21 Star™ (DE3)	Invitrogen (Holland)
One Shot® BL21 Star™ (DE3) pLysS	Invitrogen (Holland)
B834 (DE3)	Novagen (Canada)
U2OS	

2.1.6 Vectors

pET30 LIC/Xa	Novagen (Canada)
pET41 LIC/Ek	Novagen (Canada)
pET46 LIC/Ek	Novagen (Canada)
pACT2	Clontech (USA)
pFBT9	Clontech (USA)
pCMV	Stratagene (USA)

2.1.7 Antibiotics

Ampicillin, sodium salt	Sigma (USA)
Kanamycin, monosulfate	Sigma (USA)
Chloramphenicol	Sigma (USA)

2.1.8 Antibodies

His-tag H3-anti His probe	Santa Cruz Biotechnology, Inc., (USA)
pRb IF8-anti pRb MAb	Santa Cruz Biotechnology, Inc., (USA)

Goat anti-mouse Anti Goat	Santa Cruz Biotechnology, Inc., (USA)

2.1.9 Protein and nucleic acids markers

Mark12 Standard for SDS-PAGE	Invitrogen (Holland)
Prestained Protein Marker, Broad Range (6-175 kDa)	New England BioLabs (USA)
1 kb DNA-Leiter	Peqlab (Germany)

2.1.10 Peptides

Peptides were synthesed using solid phase, purified with C8 reverse phase chromatography and checked by mass spectroscopy.

2.1.11 Isotopically enriched chemicals

^{15}N-ammonium chloride, NH_4Cl 99.9%	(Campro Scientific Berlin, FRG)
^{15}N Lys, ^{15}N Leu	(Campro Scientific Berlin, FRG)

2.1.12 Cell growth media and stocks solutions

2.1.12.1 Stock solution of glucose

20 g of glucose were dissolved in distilled water to the final volume of 100 ml and autoclaved.

2.1.12.2 Stock solution of ampicillin

100 mg/ml of ampicillin were dissolved in distilled water. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

2.1.12.3 Stock solution of kanamycin

50 mg/ml of kanamycin were dissolved in distilled water. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

2.1.12.4 Stock solution of chloramfenicol

50 mg/ml of chloramfenicol were dissolved in izopropanol. The stock solution was sterile filtrated and stored in aliquots at -20°C until used.

2.1.12.5 Stock solution of IPTG

A sterile 1 M stock of IPTG in distilled water was prepared and stored in aliquots at -20°C until used.

2.1.12.6 Luria Bertani (LB)

For 1 liter LB medium:

10 g bacto tryptone
5 g bacto yeast extract
10 g sodium chloride

pH was adjusted to 7.0 with NaOH. For the preparation of agar plates the medium was supplemented with 15 g agar.

2.1.12.7 Minimal medium (MM) for uniform enrichment with ^{15}N

For 1 liter MM:

0.5 g NaCl
1.3 ml trace elements solution
1 g citric acid monohydrate
36 mg ferrous citrate
4.02 g KH_2PO_4
7.82 g K_2HPO_4 x 3H2O
1 ml Zn-EDTA solution
1 g NH_4Cl or $^{15}NH_4Cl$

pH was adjusted to 7.0 with NaOH, the mixture was autoclaved, upon cooling separately sterilized solutions were added:

25 ml glucose
560 µl thiamin, antibiotics
2 ml $MgSO_4$ stock

2.1.12.8 Defined medium for selective ^{15}N labeling of proteins (also for Sel-Met labeling)

For 1 litre of medium:

400 mg Ala, Gln, Glu, Arg, Gly
255 mg Asp, Met
125 mg cytosine, guanosine, uracil
100 mg Asn, Leu, His, Lys, Pro, Thr,
100 mg Try
400 mg Ile, Val
50 mg Phe, thymine, thymidine
1.6 g Ser
10 mg $CaCl_2$
2 g NaAc
10 g K_2HPO_4
1 g citric acid
1.3 ml trace element solution
36 mg ferrous citrate
1 ml Zn-EDTA solution
1 g NH_4Cl

pH was adjusted to 7.0 with NaOH, the mixture was autoclaved. To the cooled medium, separately sterilized solutions were added:

25 ml glucose
560 µl thiamin
antibiotics
2 ml 1 M $MgSO_4$
50 mg Cys, Trp, nicotinic acid
0.1 mg biotin
X mg ^{15}N-amino acid

Another portion of the ^{15}N-amino acid is added at the time of induction as well (the same amount as added before, 0.22 µm filtered).

Stock solutions

Glucose: 20% (w/v) in deionised H_2O

Ampicillin: 100 mg/ml

Kanamycin: 50 mg/ml

IPTG: 1 mM

Thiamin, 1%

$MgSO_4$, 1 M

Zn-EDTA solution: 5 mg/ml EDTA

8.4 mg/ml $Zn(Ac)_2$

Trace elements solution

2.5 g/l H_3BO_3

2.0 g/l $CoCl_2$ x H_2O

1.13 g/l $CuCl_2$ x H_2O

9.8 g/l $MnCl_2$ x $2H_2O$

2.0 g/l Na_2MoO_4 x $2H_2O$

pH lowered with citric acid or HCl.

2.1.12.9 Solution for making chemically competent *E. coli* cells

100 mM $MgCl_2$

100 mM $CaCl_2$, 15% glycerol

2.1.13 Buffer for DNA agarose gel electrophoresis

50 x TAE buffer (for 1 L)

40 mM Tris-acetate 242 g of Tris base

1 mM EDTA 100 ml of 0.5 M EDTA (pH 8.0)

Glacial acetic acid 57.1 ml

2.1.14 Protein purification buffers

2.1.14.1 Ion exchange and gel filtration chromatography buffers

Buffer A11

50 mM Na_2HPO_4

0.05% NaN_3

10 mM β-mercaptoethanol
pH 7.8

Buffer B1

50 mM Na_2HPO_4
1 M NaCl
10 mM β-mercaptoethanol
0.05% NaN_3
pH 7.8

PBS

140 mM NaCl
2.7 mM KCl
10 mM Na_2HPO_4
1.8 mM KH_2PO_4
0.05% NaN_3
pH 7.3

2.1.15 Buffers for Ni-NTA chromatography under native conditions

Binding buffer A0

50 mM NaH_2PO_4
300 mM NaCl
10 mM β-mercaptoethanol
pH 8.0

Binding buffer A1

50 mM NaH_2PO_4
300 mM NaCl
10 mM imidazole
10 mM β-mercaptoethanol
pH 8.0

Wash buffer B

50 mM NaH_2PO_4
300 mM NaCl
20 mM imidazole
10 mM β-mercaptoethanol
pH 8.0

Elution buffer C

50 mM NaH_2PO_4
300 mM NaCl
250 mM imidazole
10 mM β-mercaptoethanol
pH 8.0

2.1.16 Buffers for Ni-NTA chromatography under denaturing conditions

Binding buffer D

6 M guanidinium chloride
100 mM NaH_2PO_4
10 mM Tris
10 mM β-mercaptoethanol
pH 8.0

Wash buffer E

6 M guanidinium chloride
100 mM NaH_2PO_4
10 mM Tris
10 mM β-mercaptoethanol
pH 6.5

Elution buffer F

6 M guanidinium chloride
100 mM Na acetate
10 mM β-mercaptoethanol

pH 4.0

2.1.17 Protease buffers

Buffer Xa (Factor Xa cleavage buffer):

50 mM NaCl

1 mM CaC_{l2}

20 mM Tris-Cl

0.01% NaN_3

pH 6.5

Buffer Ek (Enterokinase cleavage buffer):

50 mM NaCl

2 mM $CaCl_2$

20 mM Tris-Cl

0.01% NaN_3

pH 7.4

Buffer DAPase (DAPase cleavage buffer):

150 mM NaCl

2 mM cystamine

20 mM Na_2HPO_4

0.05% NaN_3

10 mM β-mercaptoethanol

pH 7.0

Buffer TK (Thrombin cleavage buffer)

60 mM NaCl

60 mM KCl

50 mM Tris

2.5 mM $CaCl_2$

0.05% NaN_3

pH 8.0

2.1.18 Refolding Buffers

Refolding buffer RB

200 mM Arg
50 mM NaCl
1 mM EDTA
100 mM Tris
2 mM reduced glutathione
0.2 mM oxidized glutathione
0.05 NaN_3
pH 8.4

Dialysis buffer DB

250 mM Arg
300 mM NaCl
100 mM Tris
1 mM DTT
pH 8.5

2.1.19 Buffers for crystallization, NMR and ITC

FOP crystallization buffer

5 mM Tris
10 mM NaCl
0.05% NaN_3
pH 7.5

CAP2 crystallization buffer

5 mM Tris
50 mM NaCl
10 mM β-mercaptoethanol
0.05% NaN_3
pH 8.0

NMR and ITC buffer for pRb-AB

50 mM Na_2HPO_4

10 mM 2-mercaptoethanol

0.05% NaN_3

pH 7.4

2.1.20 Reagents and buffers for SDS-PAGE and Western blots

2.1.20.1 Buffers for the SDS-PAGE

Tricine SDS-PAGE

Anode buffer (+):

200 mM Tris pH 8.9

Cathode buffer (-):

100 mM Tris pH 8.25

100 mM tricine

0.1% SDS

Separation buffer:

1 M Tris pH 8.8

0.3% SDS

Stacking buffer:

1 M Tris pH 6.8

0.3% SDS

Separation acrylamide:

48% acrylamide

1.5% bis-acrylamide

Stacking acrylamide:

30% acrylamide

0.8% bis-acrylamide

Sample buffer (SB):

50 mM Tris-Cl, pH 6.8

100 mM DTT

2% SDS

0.1% bromophenol blue

10% glycerol

2.1.20.2 Protein visualization

Coomassie-blue solution:

0.1% Coomassie-blue

45% ethanol

10% acetic acid

Destaining solution:

5% ethanol

10% acetic acid

2.1.20.3 Reagents for the Western blots

Transfer buffer

10 mM CAPS, pH 11.0

20% Methanol

Alkaline phosphatase buffer

100 mM Tris pH 9.5

100 mM NaCl

5 mM $MgCl_2$

Wash buffer

10 mM Tris pH 8.0

150 mM NaCl

0.05% Tween20

2.1.21 Apparatus

2.1.21.1 ÄKTA explorer 10 purification system

ÄKTAexplorer 10 (Amersham Biosciences (Sweden)) is a complete fully automated liquid chromatography system consisting of a separation unit, fraction collector Frac-901 and a personal computer running UNICORN software v.3.1 to control the separation unit. The system enables to monitor: light absorption at 3 wavelengths simultaneously, in the range 190-700 nm; pH; conductivity; temperature; flow rate and pressure values. High performance pumps make possible to apply flow rates in range 0.01-10 ml and creating

gradients. Samples were applied with 10 or 50 ml superloop (Amersham Biosciences, Sweden), which was filled manually. UV absorption was monitored at wavelengths 255, 270 and 280. As the pH of buffers was known and constant, pH was not monitored. Runs were performed in cold cabinet, at 4°C.

2.1.21.2 Chromatography equipment, columns and media

ÄKTA explorer 10	Amersham Pharmacia (Sweden)
Peristaltic pump P-1	Amersham Pharmacia (Sweden)
Fraction collector RediFrac	Amersham Pharmacia (Sweden)
Recorder REC-1	Amersham Pharmacia (Sweden)
UV flow through detector UV-1	Amersham Pharmacia (Sweden)
BioLogic LP System	Biorad (USA)
HiLoad 26/60 Superdex S75pg	Amersham Pharmacia (Sweden)
HiLoad 10/30 Superdex S75pg	Amersham Pharmacia (Sweden)
Mono Q HR 5/5, 10/10	Amersham Pharmacia (Sweden)
Mono S HR 5/5, 10/10	Amersham Pharmacia (Sweden)
Ni-NTA-agarose	Qiagen (Germany)
SP sepharose	Amersham Pharmacia (Sweden)
GST Sepharose FF	Amersham Pharmacia (Sweden)

2.1.21.3 NMR spectrometers

600 MHz	Bruker (Germany)
750 MHz	Bruker (Germany)
600 MHz	Ultrashielded Bruker (Germany)
900 MHz	Ultrashielde Bruker (Germany)

2.1.21.4 Other apparatus

Autoclave Bachofer	Reutlingen (FRG)
Balances PE 1600	AE 163 Mettler (FRG)
BIO-RAD MINI-PROTEAN II gel apparatus	Bio-Rad Laboratories (UK)
Centrifuge 3K15	Sigma (FRG)
Centrifuge 5414	Eppendorf (FRG)
Centrifuge Avanti J-30I	Beckman (USA)

Centrifuge Microfuge R	Beckman (USA)
Centrifuges: J-6M/E and Avanti J-30I,	Beckman (UK)
Chambers for SDS PAGE and Western blotting	BioRad (FRG)
Ice machine Scotsman AF 30 Frimont	Bettolino di Pogliano (Italy)
Incubator shaker: Lab-Therm	Kühner (Switzerland)
Magnetic stirrer Heidolph M2000 Bachofer	Reutlingen (FRG)
MARresearch image plates, mar345 MARresearch	Hamburg (FRG)
PCR machine: Mastercycler personal	Eppendorf (FRG)
pH-meter pHM83 Radiometer	Copenhagen (Denmark)
Pipettes 2.5µl, 10µl, 20µl, 200µl, 1000µl	Eppendorf (FRG)
Quarz cuvettes	QS Hellma (FRG)
Shaker Adolf-Kühner	AG (Switzerland)
Sonicator: Sonifer 250	Branson (USA)
Spectrophotometer	Amersham Pharmacia (FRG)
Spectrophotometer	Ultraspec 3100 pro, Amersham Biosciences (Sweden)
Western blot apparatus (semidry)	MPI for Biochemistry
X-ray generator RU2000, 45kV, 120mA	Rigaku, Tokyo (Japan)

2.2 Methods

The most common protocols for procedures or formulations of solutions are taken directly or modified from the 'Molecular Cloning' (Sambrook 2001) or from manuals attached to the enzymes and chemicals.

2.2.1 DNA techniques

2.2.1.1 Isolation of the plasmids

All plasmids were isolated using the QIAprep Spin Miniprep Kit according to the attached protocol (QIAGEN).

2.2.1.2 Screening of positive colonies

Single colony was picked up from the plate and transfer to 10 ml of LB media containing appropriate antibiotic and grown overnight at 37°C. From 5 ml of this culture a plasmid

was isolated as described above, and then either double digested with restriction enzymes or PCR was performed using isolated plasmid as the template.

2.2.1.3 DNA sequencing

Clones, which were positive after screening of colonies and restriction assays, were given for sequencing to verify correctness of inserts in plasmids. Sequencing was performed by Medigenomix or MWG – Biotech (Germany).

2.2.1.4 Cloning to LIC vectors

PCR

Reaction was prepared as follows:

5 µl	10 x buffer
0.8 µl	dNTP 100mM
50 or 100 ng	template DNA
2.5 µl	sens_primer 100 ng/µl
2.5 µl	antisens_primer 100 ng/µl
up to 49 µl	water
1 µl	Pfu Turbo polymerase

PCR cycler was programmed depending on the primers properties, for example:

	Lid 105°C
1 cycle	95°C 2 min
30 cycles	95°C 1 min
	72°C 45 sec
1 cycle	72°C 10 min
	4°C hold

The PCR product has to be purified to remove dNTP and the template plasmid before proceeding to the T4 DNA polymerase digestion. PCR products were purified using a QIAquick PCR Purification Kit according to the manufacturer protocol (QIAGEN).

T4 DNA polymerase digestion

The following components were assembled in a sterile 1.5 ml microcentrifuge tube and kept on ice:

0.2 pmol	purified PCR product in up to 14.6 µl water
2 µl	10X T4 DNA polymerase buffer
2 µl	dGTP (pET30 Xa/LIC) or dATP (pET41 and 46 Ek/LIC) 25 mM
1 µl	DTT 100 mM
0.4 µl	T4 DNA Polymerase 2.5 U/µl (0.5 unit per 0.1 pmol PCR product) (LIC – qualified)
up to 20 µl	nuclease free water

The final concentration of insert is 0.01 pmol/µl. Concentration of the PCR product was calculated from ug/ul to molar concentration using the relation: number of bp in insert x 650 = pg/pmol. Reaction was incubated at 22°C for 30 min and subsequently inactivated by heating up to 75°C for 20 min.

2.2.1.5 Ligation

Ligation reaction was assembled with following components:

1 µl	the LIC vector (approximately 0.01 pmol of LIC vectors)
2 µl	(0.02 pmol) T4 DNA polymerase-treated LIC insert

The reaction was incubated at 22°C for 5 min and then, after addition of 1 µl of 25 mM EDTA, again 5 min at 22°C. Resulting 4 µl of the ligation reaction mixture was used for transformation of *E. coli* competent cells.

2.2.1.6 Gene synthesis

Primers for gene syntheses were designed to span desired protein sequences. They were optimized considering an *E. coli* codon usage and primers for the secondary structure calculation using NetPrime. These primers have to overlap at least 15 bp.

2.2.1.7 Site directed mutagenesis

Primers

Site directed mutagenesis requires primers which have:

1. a desired mutation and overlapping sequence
2. length of at least 25 – 45 bp
3. Tm bigger or equal to 78 °C (Tm = 81.5 + 0.41(%GC) - 675/N - % mismatch)
4. mutation positioned around the center of primer sequence and surrounded by at least 10 – 15 bp
5. optimally GC% around 40% or higher
6. FPLC or PAGE purity
7. at the ends at least 2 G or C
8. mutation with frequent codons

To check primers for the secondary structure, the software NetPrime from BioSoft was used (http://www.premierbiosoft.com/netprimer/netprimer.html). To increase the number used. To increase the number of positive clones after mutagenesis reaction, an optimized protocol was used (Wang and Malcolm 1999; Wang and Wilkinson 2001).

PCR

First step reaction

1 cycle	95°C 30 sec
2 cycles	95°C 30 sec
	55°C 60 sec
	68°C (1 min each 1kb of plasmid)

Second step reaction

1 cycle	95°C 60 sec
13 cycles	95°C 30 sec
	55°C 60 sec
	68°C 11 (1 min each 1kb of plasmid)
1 cycle	4°C 120 sec

Dpn I digestion

After completing second step the DNA template was degradated using 1 µl of the DpnI enzyme (10 U/µl) to 50 µl. Digestion at 37°C was incubated for 60 min. For a subsequent *E. coli* transformation 1 µl of the reaction mix was used.

2.2.2 Transformation of *E. coli*

2.2.2.1 Transformation by heat shock

1 µl of plasmid DNA were added to 50 µl of chemically competent *E. coli* cells. The mixture was incubated on ice for 30 min followed by a heat shock of 30 sec at 42°C, short cooling on ice, and addition of 250 µl LB medium. After 1 h of shaking at 37°C, 20-50 µl of the mixture were spread out on LB agar plates including selective antibiotics and incubated overnight at 37°C.

2.2.2.2 Transformation by electroporation

1µl of plasmid DNA was mixed together with the 40µl aliquot of electrocompetent *E. coli* and put between the electrodes of a 0.1 cm electroporation cuvette (Biorad, FRG). The cuvette was then put into the electroporator (Stratagene, FRG), and a pulse of 1660 V was applied. The value of the time constant was observed (usually 3-5 ms). The mixture was then washed out from between the electrodes with 1 ml of a sterile prewarmed (37°C) LB medium (without antibiotics), transferred to a sterile 1.5 ml tube and shaked (800 rpm) at 37°C. After 1 h 20 - 50 µl of cells mixture were streaked on a LB agar plate with an appropriate antibiotic.

2.2.3 Preparation of chemically competent cells

1. *E. coli* were grown in a 5 ml overnight culture.
2. 250 ml of LB medium were inoculated with a 5 ml overnight culture and incubated at 37°C with shaking to the $OD_{600} = 0.5$
3. The culture was transferred to the two chilled, sterile 500 ml centrifuge bottles and incubated on ice for 30 min, then centrifuged at 4000 rpm for 15 min at 0 - 4°C.
4. Cells were resuspended in 100 ml of 100 mM $MgCl_2$ by pipetting, then incubated on ice 30 min.
5. The culture was centrifuged at 4000 rpm for 15 min at 0 - 4°C.
6. The cell pellet was then resuspended in 10 ml cold, sterile 100 mM $CaCl_2$ with 15% glycerol and transferred to the chilled, sterile, small eppendorf tubes (100 µl/tube).

2.2.4 Protocol for electrocompetent cells

1. *E. coli* were streaked on an LB agar plate and incubated at 37°C overnight.

2. 50 ml of the LB medium in a 250 ml flask were inoculated with a single colony from the LB plate and incubated at 37°C with shaking (200 rpm) overnight.
3. 1 l of LB medium in a 3 l flask was inoculated with the 50 ml overnight culture. The culture was grown in a shaking (200 rpm) incubator at 37°C until the OD_{600} was between 0.5 - 0.6 (approximately 2 h).
4. The culture was transferred to the two chilled, sterile 500 ml centrifuge bottles and incubated on ice for 30 min. Thereafter centrifugation followed at 2000 G for 15 min at 0 - 4°C.
5. Supernatant was decanted, and bottles put back on ice. The cell pellet in each bottle was resuspended in approximately 500 ml of cold (0 - 4°C), sterile water, and subsequently centrifuged like before.
6. The cells in each bottle were washed again with 50 ml of cold sterile water, and centrifuged.
7. The cell pellet in each bottle was then resuspended in 20 ml cold, sterile 10% glycerol and transferred to a chilled, sterile, 50 ml centrifuge tube. Centrifugation followed at 4000 G for 15 minutes at 0 - 4°C.
8. The 10% glycerol was decanted and pellet resuspended for the second time in 1 ml cold sterile 10% glycerol.
9. Using a pre-chilled pipette the cell suspension was aliquoted (40 µl) to pre-chilled 1.5 ml tubes and frozen immediately in liquid nitrogen. The aliquots were kept at -80°C ready for use.

2.2.5 Protein expression

A number of parameters are significant to get the maximum yields of the protein, which include optimization of proper protein constructs, a type of culture media, temperature, induction OD, concentration of inducer (IPTG), and cell types. In this work proteins were cloned into pET vectors, which are IPTG inducible according to the protocol described in section 2.2.1.4.

2.2.5.1 *E. coli* expression in LB medium

1. 20 ml LB were inoculated with a fresh single bacterial colony and incubated overnight at 37°C with vigorous shaking (200 rpm) in a 50 ml falcon tube.

2. A pre-warmed 1 l LB medium in 3 l flask was inoculated with 20 ml of the overnight culture, supplemented with an appropriate antibiotic, and incubated at 37°C with shaking (150 rpm) until the OD_{600} reached the 0.7 - 0.9 value.
3. Induction by addition of 0.2 - 1 mM IPTG (final concentration) was usually used. The cells were then grown until the expected OD was reached.
4. Cells were pelleted by centrifugation at 4000 rpm in a Beckman centrifuge for 20 min at 4°C.
5. The cell pellet after the expression was used directly for protein purification or stored at -80°C for longer time or -20°C if to be used in 1-2 days.

2.2.5.2 *E. coli* expression in the minimal medium (MM)

1. 5 ml LB were inoculated with a single colony and shaked (200 rpm) overday in a 15 ml falcon tube at 37°C.
2. 100 ml MM were inoculated with 100 µl of the overday culture and shaked (150 rpm) overnight in a 250 ml flask at 37°C.
3. 1 l MM was inoculated with 100 ml of the overnight culture and shaked (150 rpm) in a 3 l flask.
4. Cells were induced with IPTG (final concentration 0.2 -1.0 mM) at OD_{600} = 0.7 - 0.9.
5. After an overnight incubation with shaking, the cells were pelleted by centrifugation at 4000 rpm at 4°C in a Beckman centrifuge for 20 min.
6. The cell pellet (after the expression) was used directly for the protein purification or stored at -80°C for longer time or -20°C if to be used in 1-2 days.

2.2.6 Protein purification and refolding

2.2.6.1 Protein purification under native conditions

Bacterial pellets were resuspended in buffer A0 or A1 (20 ml of buffer for 1 l colony). The cell suspension was sonicated 5 x 4 min using a macrotip (output control 7.5, 50%) and centrifuged at 20000 rpm for 20 min at 4°C. The obtained supernatant was incubated with a Ni-NTA slurry (Qiagen) equilibrated previously in buffer A0 or A1, for 2 h at 4 °C with gentle agitation. Next, the mixture was loaded onto an empty column and washed with buffer A0 or A1 and B. The protein was eluted with buffer C. All buffers are described in section 2.1.15.

2.2.6.2 Protein purification under denaturating conditions

In the case of insoluble proteins accumulated in inclusion bodies (IBs), bacterial pellets were first resuspended in buffer A0 or A1 (to remove soluble bacterial proteins) and sonicated. The supernatant was removed and inclusion bodies were resuspended in the binding buffer D. The solution was centrifuged at 25000 rpm for 30 min at 16°C. The obtained supernatant was incubated with a Ni-NTA slurry (Qiagen), equilibrated previously in buffer D, for 1 h at 20°C with gentle agitation. Next, the mixture was loaded onto an empty column and washed with buffer D and E. The protein was eluted with buffer F (section 2.1.16).

In the next step refolding of the protein was performed. There are strictly no set rules to achieve the best conditions for refolding proteins from inclusion bodies. During this work refolding was performed by both rapid dilution and dialysis. Many variables can be tried, like arginine, glutamic acid, reduced/oxidized glutathione ratio, urea, temperature, salt concentration, protease inhibitor, glycerol etc. For a rapid dilution technique, the protein sample was diluted in the refolding buffer RB (1:100) in ~150 µl aliquots, with 10-20 min intervals between aliquots. Thanks to the gradual dilution, the accumulation of high concentration of folding intermediates was avoided, which could otherwise lead to aggregation. The refolding mixture was left with stirring at 4°C for 12 h. For the dialysis method, 2 l of the dialysis buffer DB were prepared. The protein was diluted to the concentration 0.5 mg/ml and dialyzed overnight.

2.2.7 Protease digestion

For the pRb samples, the His–tag was removed prior to a pull–down assay experiments by a DAPzyme digestion in a DAPase buffer. The reaction was performed according to a manufacturer manual at room temperature overnight. The proteins were subsequently purified over NiNTA to remove His–tags or undigested proteins. For several of crystallization trials the His-tag was cleaved from CAP350 constructs by the recombinant enterokinase (the Ek buffer) according to the manufacturer manual, at room temperature for 24 h. For FOP constructs, the His-tag alone or the His-tag plus the N-terminal part of a protein, were cleaved with the Xa protease at room temperature for 2-3 days in buffer Xa (section 2.1.17). Subsequently, all the proteins were purified on the gel filtration HiLoad 26/60 Superdex S75pg Amersham Pharmacia (Sweden) column. The His-tag removal for all constructs was controlled by using SDS–PAGE and Western–blots.

2.2.8 Handling and storing of the proteins

All the operations involving the proteins at either level of purification were performed on ice or at 4°C, unless denaturating conditions were used. 2-mercaptoethanol was added to buffers during purification of the proteins containing Cys - to prevent disulfides bond formation. NaN_3 was added to all solutions at concentration 0.05%. The function of NaN_3 was to prevent bacterial and fungal growth in solutions stored for prolonged time. Purified proteins were stored at +4°C or at -80°C to further analyses.

2.2.9 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

To verify the purity of protein samples, SDS polyacrylamide gel electrophoresis was carried out. Because small proteins were analyzed, tricine gels were chosen (Schagger and von Jagow 1987).

Pouring gels

Separation gel:

1.675 ml H_2O

2.5 ml separation buffer

2.5 ml separation acrylamide

0.8 ml glycerol

25 µl APS

2.5 µl TEMED

Intermediate gel:

1.725 ml H_2O

1.25 ml separation buffer

0.75 ml separation acrylamide

12.5 µl APS

1.25 µl TEMED

Stacking gel:

2.575 ml H_2O

0.475 ml stacking buffer

0.625 ml stacking acrylamide

12.5 µl 0.5 M EDTA, pH 8.0

37.5 µl APS (stock 10%)

1.9 µl TEMED

The guanidinium HCl-free protein samples were prepared by mixing 20 µl of protein solution with 5 µl of sample buffer (SB) followed by 3 min incubation at 100°C. Due to the rapid precipitation of SDS in contact with guanidine, the samples to be examined by SDS-PAGE (after Ni-NTA chromatography under denaturing conditions) had to be prepared in a following fashion: 20 µl of the protein solution in a denaturing buffer was diluted with 400 µl 20% trichloroacetic acid (TCA). The sample was incubated for 5 min at room temperature followed by centrifugation for 5 min at 20 000 x g. Supernatant was discarded by suction, a precipitated protein pellet was washed once by vortexing with 400 µl ethanol. After centrifugation and ethanol removal, the protein pellet was resuspended in 20 µl of 2 x SB and the sample was boiled for 3 min.

2.2.10 Western immunoblotting

This method of protein detection was applied to confirm immunological identity of the expressed proteins and to determine presence of any degradation products or isoforms of the purified proteins.

The proteins separated by SDS-PAGE were blotted in a semi-dry blotter (home made). One 1 mm filter paper soaked in a blotting buffer was placed on the bottom anode plate of the apparatus. A nitrocellulose membrane (NC) (Amersham Bioscience), SDS-PAGE, and one filter paper soaked in blotting buffer, were placed on the top of it. The blotting was carried out at 200 mA for 1 h. The NC membrane was then immersed in 20 ml of a 5% milk powder solution in PBS-T and gently agitated for 1/2 h, than a respective primary antibody was added and the NC membrane was agitated for additional 1 h. The membrane was washed 3 x 15 min with 20 ml of washing buffer and incubated for 1 h with the secondary antibody – AP conjugated goat anti-mouse immunoglobulin (1:2000 dilution) (Santa Cruz) in a washing buffer. The NC membrane was washed 3 x 5 min with 20 ml of washing buffer and for visualization a substrate (SIGMA FAST BCIP/NBT) for AP was added for 10 min.

2.2.11 Determination of protein concentration

The concentration of proteins in solutions was estimated using the Bradford reagent (Biorad) (Bradford 1976). 5 µl of the protein solution (or 1 µl, if the protein solution is

very concentrated) to be measured were added to 495 µl of a Biorad-reagent working solution (working solution was prepared by 1:10 dilution of a Biorad-reagent stock solution in the PBS buffer or water, stored in the fridge). After thoroughly mixing the sample, the OD_{595} was measured. As a reference, a similar mixture was prepared with the 5 µl buffer instead of the protein solution. OD was subsequently converted into the protein concentration on the basis of a BSA calibration curve.

2.2.12 Pull down assays

The bait proteins with the tag were bound to the respective columns. A prey with an enzymatic cleaved tag was put over a column saturated with the bait. The complex was eluted with a buffer supplemented with 250 mM imidazole. Fractions obtained were tested on SDS–PAGE. For His-tagged proteins we used buffers described in section 2.1.15.

2.2.13 Mass spectrometry

Mass spectrometry was carried out on an ESI–MS API 165 Perkin–Elmer Sciex (Langen) coupled with HPLC (column: Macherey–Nagel EC 125/2 Nucleosil 300–5 C4 MPN; pump system: Microgradient System 140B/C Perkin Elmer; solvent A: water, 0.05 % TFA, B: MeCN, 0.05% TFA; gradient 10–95% B; the photodiode array: Agilent HP1100PDA; software: Masschrom, Biomultiview).

2.2.14 Binding tests on gel filtration column

A mixture of two proteins usually in 1:1 molar ratio was subjected to separation on the HiLoad 26/60 Superdex S75pg or HiLoad 10/30 Superdex S75pg Amersham Pharmacia columns. Subsequently, both proteins were loaded on the same column separately (the same concentration and volume of protein was achieved). We used the buffer 50 mM NaH_2PO_4, 300 mM NaCl, pH 7.8, 10 mM βME for following experiments: pRb + MyoD, pRb + Id-2, pRb + E7 (both full length and short), pRb + large T antigen, MyoD + Id-2. PBS buffer was used for pRb + PAI-2 all constructs, CAP350 + FOP. All fractions were checked on SDS–PAGE and after concentrating with mass spectroscopy.

2.2.15 PAI-2 activity tests

The PAI2 protein activity was tested for its ability to inhibit the urokinase type plasminogen activator (uPA), which in the presence of DNTB (5,5'-dithio-bis-2-

nitrobenzoic acid) performs enzymatic hydrolysis of the thioester substrate BLT - (*N*-α-benzyloxycarbonyl-L-lysine thiobenzyl). The reaction starts with a nucleophilic attack of uPA on a carbonyl group of lysine, which causes the BLT splitting. As the result of the reaction free thiol (TNB) groups are released from the DTNB substrate. The amount of them can be measured photometrically at a wavelength of 405 nm.

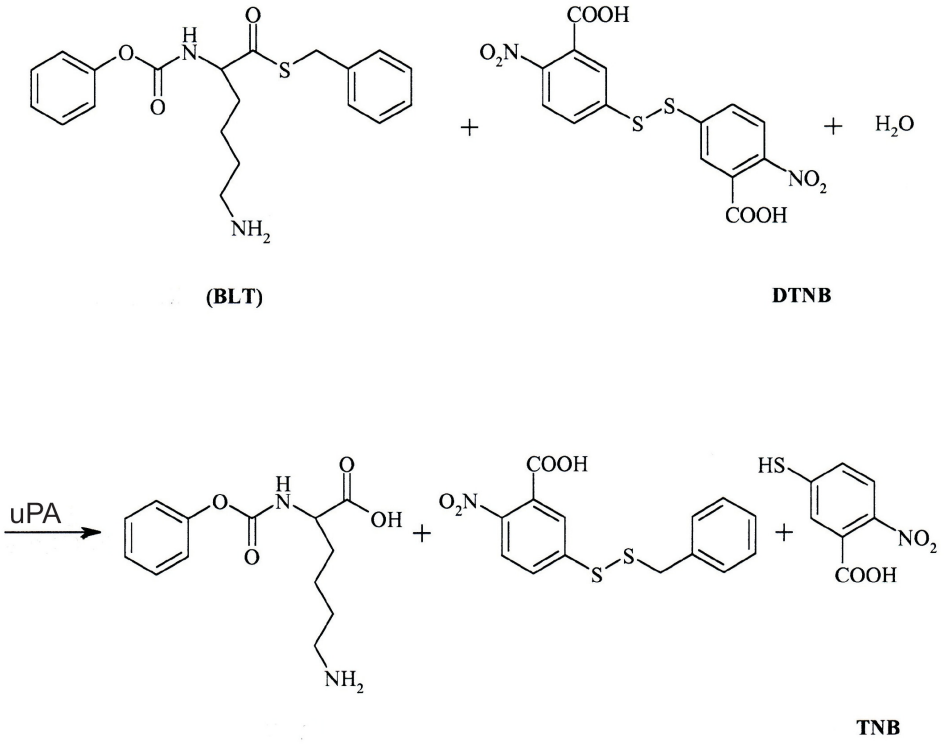


Figure 2.1. The serine protease uPA grabs the nucleophile of a carbonyl group of lysine and splits BLT - (*N*-α-benzyloxycarbonyl-L-lysine thiobenzyl) into two parts. As the result of the reaction a free, thiol (TNB) group is released from DTNB (5,5'-dithio-bis-2-nitrobenzoic acid) substrate.

Activity reaction mix

100 µl	BLT 3 mM in 30% DMSO
100 µl	DNTB 0.75 mM in 150 mM Tris pH 7.5
1 µl	uPA 1 mg/ml; 4 mg/ml
Up to 500 µl	150 mM Tris pH 7.5

An average activity of uPA was calculated using the formula:

$$\text{Activity} = (\text{dA/min}) / \mu\text{g Enzyme}$$

dA is an absorbance change in time.

2.2.16 Crystallization trials

Both, sitting drop and hanging drop vapor diffusion techniques were used for crystallization. The final purification of proteins was against 5 mM Tris pH 7.5, 10 mM NaCl, 0.05% NaN_3 for FOP constructs and 5 mM Tris pH 8.0, 50 mM NaCl, 10 mM β-mercaptoethanol, 0.05% NaN_3 for CAP2. Collected fractions were pooled and concentrated until the expected protein concentration was achieved (5-30 mg/ml). Subsequently, the samples were centrifuged to clear out small particles and precipitates. The crystallization procedure was performed as quickly as possible (usually directly after concentration of the proteins). To determine the preliminary crystallization conditions the Hampton Research Crystal Screens were used. 0.1 µl of a protein sample and 0.1 µl of the crystallization buffer were mixed in 96-well plates; the reservoirs were filled with 100 µl of buffer. Crystallization plates were set up at two temperatures: 4°C and room temperature. After crystals appeared the crystallization was performed on a larger scale in the best conditions of the initial screen. 1 µl of the protein sample and 1 µl of the crystallization buffer were mixed in 24-well plates. Reservoirs were filled with 500 µl of crystallization buffer.

2.2.17 Isothermal titration calorimetry (ITC)

A typical ITC experiment was carried out as follows: component 1 (ligand) in the syringe is titrated into the second component 2. Usually the ligand concentration is typically 10 to 20 times higher than the concentration of the component in the cell. The volume of the component 1 in the cell was 1430 µl and component 2 in the syringe 600 µl. All solutions were degassed prior to measurements. Heat generated by protein dilution was determined in separate experiments by injecting the protein solution into the PBS filled sample chamber. All data were corrected for the heat of protein dilution. The details of the experimental and injection parameters are described below.

Parameter	Value
Total number of injections	45
Volume of a single injection [µl]	5
Duration of an injection [sec]	10
Intervals between injections [sec]	600
Filter period [sec]	2

Equilibrium cell temperature [°C]	20-22
Initial delay [sec]	60
Reference power [μCal/sec]	15
Stirring speed [RPM]	270

For the pRb-AB and LXCXE peptides titrations, pRb-AB was component 2 in the cell, whereas ligands were the LXCXE peptides. The buffer system comprised: 50 mM Na_2HPO_4, 10 mM βME, pH 7.8, 0.01% NaN_3. The protein was extensively dialyzed against the buffer, whereas the peptides were dissolved in the final dialysis buffer, pH of peptide solutions were carefully checked and adjusted in some cases. Protein concentration was 20 - 50 μM whereas the peptide concentration was 200-400 μM. Experiments were carried out at 20 - 22°C. All data were corrected for the heat of protein dilution. Data were fitted using χ^2 minimization on a model assuming a single set of sites to calculate the binding affinity K_D. All steps of the data analysis were performed using the ORIGIN (V5.0) software provided by the manufacturer.

2.2.18 NMR spectroscopy

NMR spectra were recorded on a 600 MHz Bruker spectrometer equipped with a triple resonance probehead, processed using XWINNMR (Bruker), and analyzed with the program Sparky (Goddard, T.D., and Kneller, D.G., University of California, San Francisco). 1D-^{1}H spectra require protein concentrations as low as 0.05 mM (preferably in the phosphate buffer) whereas 2D or higher dimensional spectra require the protein to be labeled and require higher concentrations of the protein (0.2 mM to 1 mM). Typically, a 450 μl of the protein sample and 10 μl of D_2O was mixed and added to the NMR tube. All 1D-^{1}H NMR spectra were recorded with a time domain of 32 K complex points and a sweep-width of 10,000 Hz. The 2D ^{1}H-^{15}N-HSQC spectra were recorded with a time domain of 1 K complex data points with 128 complex increments with a sweep width of 8 kHz in the 1H dimension and 2 kHz in the ^{15}N dimension.

2.2.19 Immunofluorescence microscopy

Human osterosarcoma cells U2OS were grown at 37°C and 5% CO_2 in DMEM (Invitrogen-BRL), supplemented with 10% fetal calf serum and penicillin-streptomycin (100 IU/ml and 100 mg/ml, respectively). Immunofluorescence (IF) microscopy was

performed as described previously (Meraldi and Nigg 2002). U2OS cells were transfected with different FLAG-tagged FOP constructs for 24 h and costained with anti-γ-tubulin and anti-FLAG antibodies. Cells were grown on cover slips and fixed for 10 min in cold methanol. Primary antibodies were the rabbit anti-FLAG (1:1000, F7425, Sigma) and monoclonal anti-γ-tubulin (1:1000, T6557, Sigma). Secondary reagents were the Alexa-Fluor-555-conjugated goat anti-mouse and Alexa-Fluor-488-conjugated goat anti-rabbit (1:1000, Molecular Probes). DNA was stained with 4,6-diamidino-2-phenylindole (DAPI, 2 μg/ml). IF microscopy was performed using a Zeiss Axioplan II microscope with 40 x and 63 x oil immersion objectives. Photographs were taken using a Micromax 1300 x 1030 pixel CCD camera (Princeton Instruments) with Metavue software (Universal Imaging). Images were processed with the Adobe Photoshop software (Adobe Systems Inc.).

2.2.20 Yeast two hybrid system

Yeast two-hybrid assays typically use selection genes encoding proteins capable of synthesizing amino acids such as histidine, leucine and tryptophan. Two genes of the proteins that are supposed to interact are cloned into two plasmids. Only cells possessing both plasmids can survive on selective medium (–LW). If the proteins interact cells posses active transcription factor, for example GAL4. In this work we used three reporter genes– *ADE2*, *HIS3* and *lacZ* that are expressed from GAL2, GAL1, and MEL1 promoters. These three promoters are under control of GAL4. Cells expressing active form of GAL4 can survive on the (QDO) medium - lacking leucine, tryptophan, histidine, and adenine, with 2% (wt/vol) glucose as the carbon source (Chien et al. 1991; Fields and Sternglanz 1994).

2.2.20.1 Yeast transformation protocol

1. Inoculate the yeast strain containing the first plasmid into the appropriate volume of the appropriate SC-omission medium in a flask and incubate at 30°C overnight.
2. Determine the cell titer and calculate the volume of cells that yields 2.5 x 10^8 cells for each 50 ml of the YPAD culture needed.
3. Pour this culture volume into an appropriate sterile centrifuge tube and pellet the cells at 3000 x g for 5 min. Resuspend the cell pellet in the appropriate volume of pre-warmed (30°C) YPAD and transfer to another sterile culture flask.
4. Incubate at 30°C while shaking at 200 rpm for 3 to 4 h until the cell titer reaches 2 x 10^7 cells/ml.

5. Harvest the cells by centrifugation at 3000 x g for 5 min.
6. Wash the cell pellet via resuspension with 1/2 volume of sdd water and collection by centrifugation at 3000 x g for 5 min.
7. Resuspend the pellet in the appropriate volume of 100 mM sterile LiAc and transfer to an appropriate centrifuge tube. Incubate for 15 min at 30°C. Pellet the cells again by centrifugation and remove the supernatant.
8. Add the transformation mix to a seperate tube and mix thoroughly by vortexing (50% PEG, 1.0 M LiAc, SS-DNA (2 mg/ml), library plasmid DNA, sdd water).
9. Vigourously vortex the cell pellet until it is totally resuspended, which should take about 1 min. If you have problems getting the pellet resusupended let it sit for 5 min and then vortex.
10. Incubate the transformation mix at 30°C for 30 min.
11. Heat shock at 42°C for time indicated by table below with mixing by inversion for 15 sec after every 5 min.
12. Collect the cells by centrifugation as above. Gently resuspend the cell pellet in an appropriate volume of sdd water and plate onto selection agar.
13. Incubate the plates for 3 - 5 days at 30°C or until colonies appear.

3 RESULTS AND DISCUSSION

3.1 Centrosomal proteins - structural and functional studies

The aim of this part of my thesis was the structural and functional studies of two centrosomal proteins FOP and CAP350. Protein cloning, expression, purification, and functional experiments were performed in-house. The structural research was divided into two parts: data collection was performed on the MPG/GBF beamline BW6 at DESY, Hamburg, and the structure determination in-house. Functional experiments have determined the *in vitro* and *in vivo* binding between FOP and CAP350, as well as dimerization abilities and centrosomal localization of FOP.

3.1.1 Results

3.1.1.1 Plasmid constructions and protein expression

For the IF microscopy, the C-terminally flag-tagged FOP plasmids were constructed in mammalian pCMV expression vectors by standard procedures. The FOP cDNA has been previously described (Andersen et al. 2003). Human osteosarcoma cells-U2OS were transfected with different FLAG-tagged FOP constructs for 24 h and costained with anti-γ-tubulin and anti-FLAG antibodies.

For the yeast two-hybrid experiments, fragments of FOP were amplified by PCR and subcloned into pACT2 or the two-hybrid bait vector pFBT9, modified to encode kanamycin resistance; kindly provided by Dr. F. Barr. Transformed yeast samples were plated on synthetic media either lacking leucine and tryptophan (–LW) or lacking leucine, tryptophan, histidine, and adenine, with 2% (wt/vol) glucose as the carbon source (QDO). All results were confirmed by streaking several independent colonies on both selective (QDO) and nonselective (–LW) plates.

For *E. coli* expression, constructs of human FOP and CAP350 were cloned into the pET46 Ek/LIC as described in Materials and Methods section. Mutations were introduced by the Quick-Change Site Directed Mutagenesis Kit. To increase expression level of the FC2LisH construct (residues 54-134), I engineered a highly expressed construct (1-134) with a factor Xa cleavage site at position 54. Additionally, to facilitate the estimation of the protein concentration, Trp135 was added at the end of the polypeptide chain. The proteins were expressed in an *E. coli* strain BL21 STAR™ (DE3) in the LB medium with

100 µg/ml ampicillin by overnight induction with 1 mM IPTG at 37°C. CAP350 and FOP-FGFR1 constructs were induced with 0.5 mM IPTG at 25°C. Additionally, the constructs for the Se-Met labeling were expressed in a minimal medium enriched with unlabeled amino acids and supplemented with 100 mg/ml Se-Met (Muchmore et al. 1989). Overnight induction was carried out with 1 mM IPTG at 37°C in a methionine auxotroph *E. coli* B834 (DE3) (Novagen).

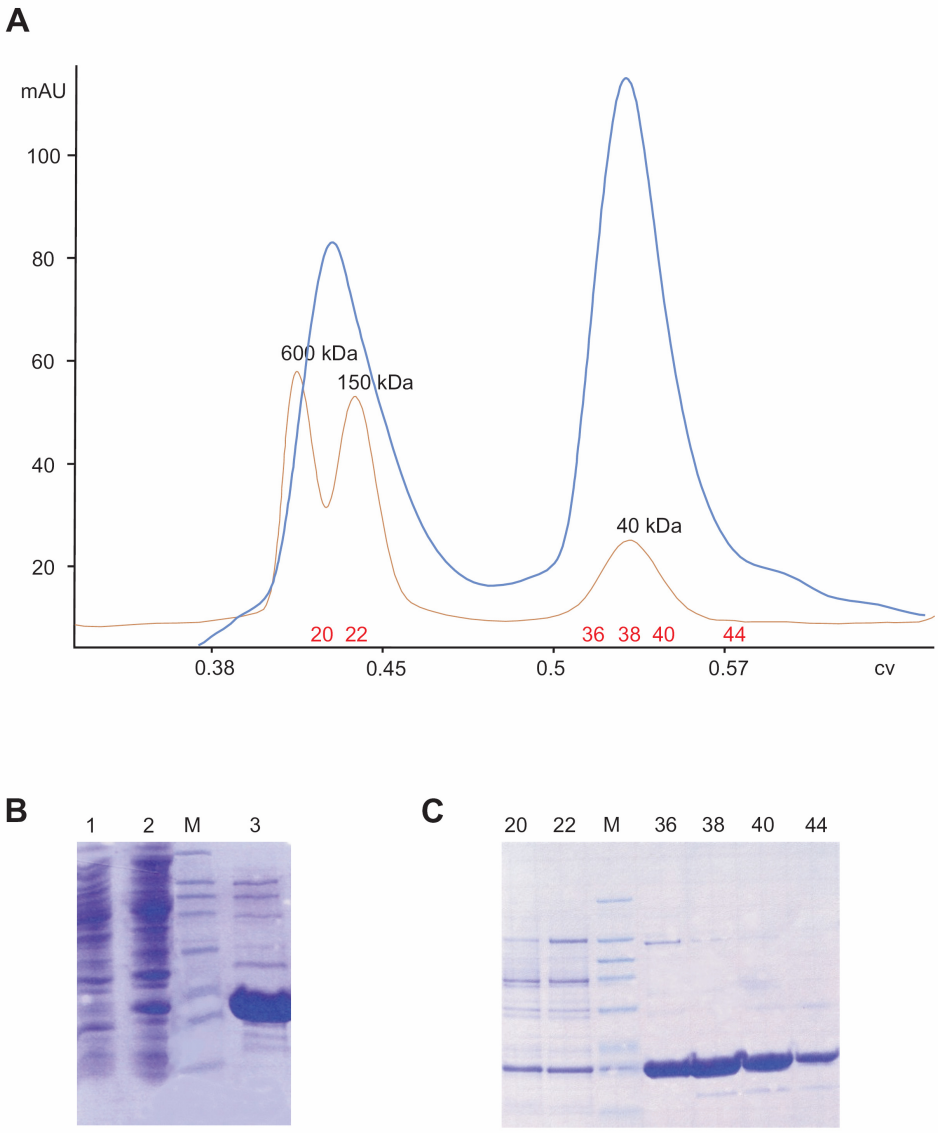


Figure 3.1.1. Purification and solubility analysis of the FC2S construct. Panel A) the gel filtration elution profile of the FC2S construct after a Ni-NTA column - line blue, line brown – protein markers, numbers in red – fractions taken for SDS analysis (the column volume 318.557 ml). Panel B) SDS PAGE assay of protein purification under native conditions; lane 1 – insoluble fraction, lane 2 – soluble fraction before loading on the Ni-NTA column, M – protein marker, lane 3 – the soluble fraction after purification on the Ni-NTA column. Panel C) SDS PAGE analysis of gel filtration fractions.

3.1.1.2 Protein purification strategies

The Ni-NTA metal-affinity chromatography was used as the first step of purification. All FOP constructs were soluble and purified under native conditions. CAP350 constructs (CAP3 and CAP4) were accumulated in inclusion bodies and were purified under denaturating conditions, followed by a dilution refolding method. CAP2 and FOP-FGFR1 constructs were partly soluble and were purified by using both strategies (Materials and Methods section 2.2.6). Refolding of proteins was performed by using a rapid dilution technique with buffer RB. Gel filtration chromatography was carried out as a next step of purification using the ÄKTA Explorer 10 chromatography system. It was impossible to purify C-terminal constructs of FOP (FC3 and FC4) and FOP-FGFR1 constructs (FR_H, FRH_del, FR_del) because of their low expression, instabilities or insolubility (Table 3.1.1). In order to get the highly expressed FC2LisH construct, the FC2S protein (Figure 3.1.1) was purified against the Xa protease buffer. Removal of the His-tag and the 53 aa N-terminal fragment with the Factor Xa protease was followed by the final purification against 5 mM Tris pH 7.5, 10 mM NaCl, 0.05% NaN_3.

Table 3.1.1. List of proteins and their constructs used in this work together with experiments performed for their analysis.

Name	Residues (from-to)	Tag	Length	Purification	Oligomerization	Folding 1D NMR	Crystallization trials
FOP_full length	1-399	15 aa His-tag from pET46 Ek/LIC vector	413	-	-	-	-
FC2	17-150	15 aa His-tag from pET46 Ek/LIC vector	151	+	dimer	partly folded	negative
FC3	225-399	15 aa His-tag from pET46 Ek/LIC vector	189	-	-	-	-
FC4	290-399	15 aa His-tag from pET46 Ek/LIC vector	125	-	-	-	-
FC2L	1-173	15 aa His-tag from pET46 Ek/LIC vector	188	+	tetramer	unfolded	-

FC2NC	17-134	15 aa His-tag from pET46 Ek/LIC vector	135	+	dimer	partly folded	-
FC2S	1-134	15 aa His-tag from pET46 Ek/LIC vector	153	+	dimer	partly folded	negative
FC2SHis Mut	1-134	10 aa His-tag MAHHHHHH KK	149	+	dimer	partly folded	-
FC2LisH	54-134	Tag removed by Xa digestion	82	+	dimer	folded	positive
CAP350_1	3059 - 3117	15 aa His-tag from pET46 Ek/LIC vector	74	+	dimer	-	-
CAP350_2 EkHis	2960- 3117	Tag removed by Enterokinase digestion	158	+	dimer	-	negative
CAP350_2	2960- 3117	15 aa His-tag from pET46 Ek/LIC vector	173	+	dimer	folded	negative
CAP350_3	2692- 3117	15 aa His-tag from pET46 Ek/LIC vector	441	+	aggreg.	-	-
CAP350_4	2960- 3074	15 aa His-tag from pET46 Ek/LIC vector	130	+	-	-	-
FR1	1-173 [a] 427-759 [b]	15 aa His-tag from pET46 Ek/LIC vector	520	+	aggreg.	-	-
FR_H	54-173 [a] 427-759 [b]	10 aa His-tag MAHHHHHH KK	462	-	-	-	-
FRH_del	54-163 [a] 462-759 [b]	10 aa His-tag MAHHHHHH KK	401	-	-	-	-
FR_del	1-163 [a] 462-759 [b]	15 aa His-tag from pET46 Ek/LIC vector	459	-	-	-	-

[a] range of FOP residues in FOP-FGFR1 construct

[b] range of FGFR1 residues in FOP-FGFR1 construct

3.1.1.3 *In vitro* binding tests between CAP320 and FOP

Mapping studies showed that the last 47 C-terminal amino acids of CAP350 are both necessary and sufficient for the binding of the N-terminal part of FOP. The structure of the CAP350 fragment is unknown, but its functional importance is confirmed by a very high degree of sequence conservation with the C-terminus of vertebrate CAP350 homologues (Yan et al. 2005). For *in vitro* binding experiments, I used the two best folded constructs of both proteins. Purity of them was estimated to be about 95% for CAP350 and 98% for FOP. Gel filtration analysis showed that single proteins migrated as dimers. Mixture of both proteins was separated on the same s75 Superdex column after 0.5 h of incubation. We did not notice any peak shift which indicates the lack of tight binding (Figure 3.1.2).

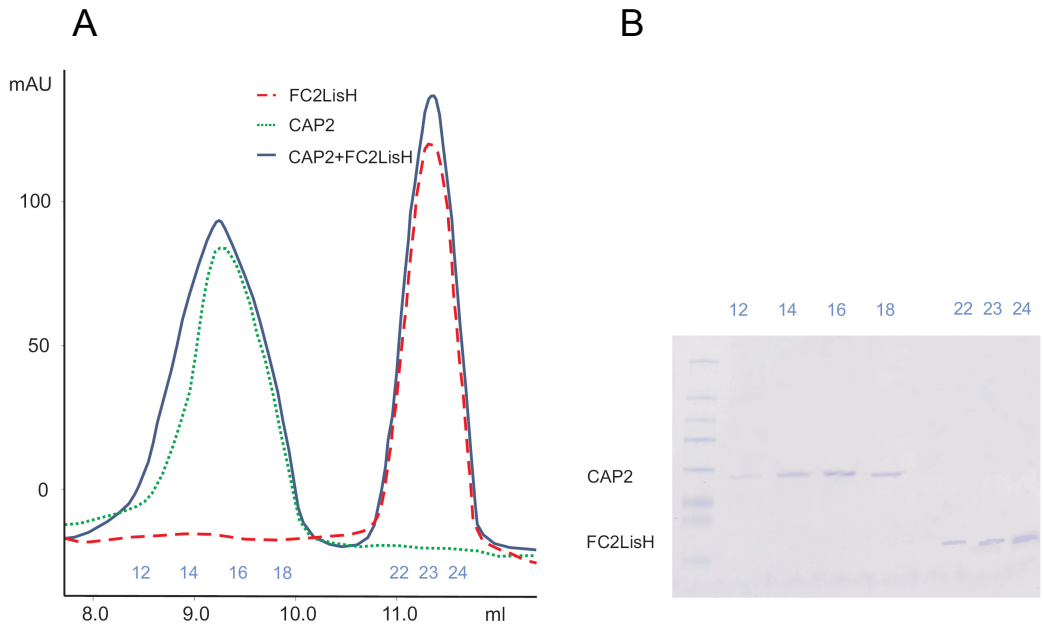


Figure 3.1.2. Panel A) - gel filtration analysis of complex formation between the C-terminal part of the CAP350 protein (CAP2 construct; dotted line) and the FC2LisH construct of FOP (dash line). We observed no complex formation between both proteins (straight line). Panel B) shows the SDS-PAGE analysis of consecutive chromatography fractions.

3.1.1.4 Optimizing the constructs of FOP and CAP350 for crystallization

Purified proteins were analyzed using a variety of methods. For their folding properties, they were checked by NMR spectroscopy (Figure 3.1.3) (Zhang and Forman-Kay 1995; Rehm et al. 2002), CD spectroscopy, and for oligomerization by gel filtration chromatography.

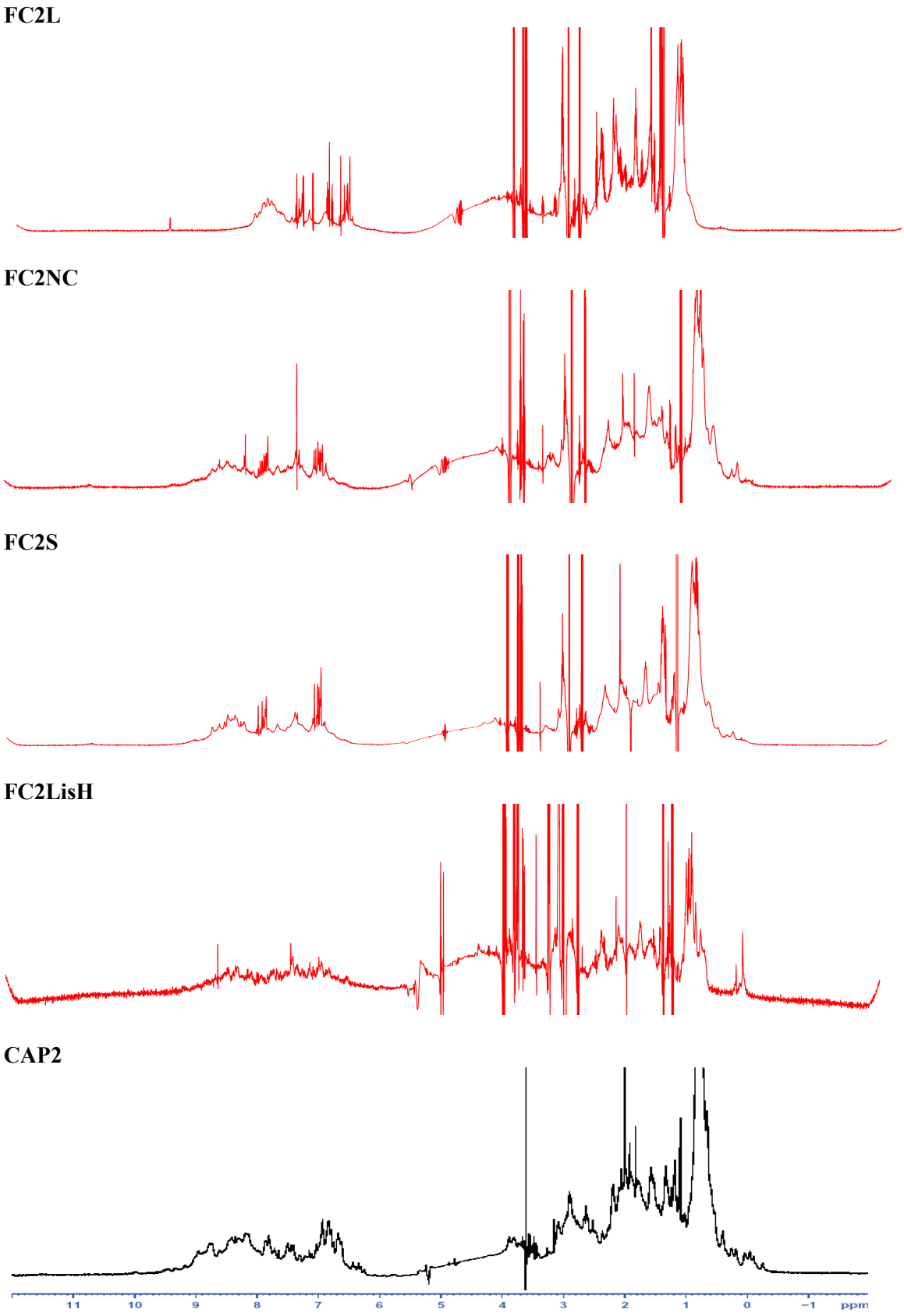


Figure 3.1.3. 1D-^{1}H spectra of various FOP (red) and CAP350 (black) constructs used in this work show the folding characteristics of these protein domains.

Small dispersion of NMR signals at approximately 8.3 ppm in the 1D proton spectrum indicates a protein to be unfolded as this is a region characteristic of backbone amides in a random coil configuration. On the other hand, large NMR signal dispersion beyond 8.3 ppm and in the aliphatic region of the spectrum, between 1.0 and −1.0 ppm, proves a protein to be folded.

3.1.1.5 Crystallization and data collection

Crystallization trials of the CAP350 construct, CAP2, were performed with both tagged and untagged proteins. A 15 aa His-tag was removed by recombinant enterokinase according to the manufactured manual. Although 1D-^{1}H spectra showed that the CAP2 constructs were well structured in solution, all my attempts to crystallize them were unsuccessful. I tried different crystallization screens, which included the Hampton screens: 1, 2, Lite, PEG/Ion, Index and SaltRx, as well as Nextal, as well as two temperatures 4°C and room temperature.

Crystallization of FOP was performed for the constructs: FC2, FC2S, which were classified as partly folded by 1D-^{1}H NMR spectra, and for the folded construct - FC2LisH. Crystallization attempts were carried out with the sitting and hanging drop vapor diffusion methods at 20°C and 4°C. I was successful in crystallizing the FC2LisH construct (Figure 3.1.4). The crystals appeared in both commercial Hampton screens, as well as in self-made solutions. The best crystals were obtained in 0.2 M NaCl; 0.1 M HEPES pH 7.5; 25% w/v PEG 3350 with the protein concentration of ~7 mg/ml at room temperature. The crystals belonged to the space group $P2_12_12$ and contained one dimer per an asymmetric unit. Unit cell dimensions were a = 47.7 Å, b = 92.1 Å, and c = 37.1 Å. Data were collected to 1.6 Å resolution at 1.05 Å wavelength (Figure 3.1.5). Molecular replacement (MR) and multiple isomorphous replacement (MIR) methods were not successful for finding phases. For molecular replacement, I tried to use the phases from the LIS1 protein, which has a similar dimerization domain to FOP. For multiple isomorphous replacement, the crystals were soaked with platinum, uranium and lead salts. Finally, I have made a Se-Met labeled constructs for MAD. Because the native FC2LisH construct did not possess methionine, I prepared six mutants, which leucine or lysine residues were replaced by methionine. From the total of six of them, only two crystallized. The crystals of the Se-Met labeled protein FC2LisH_M2 (K63M) grew in the same buffer as the native protein but in 4°C. The crystals of the FC2LisH_M5 (L104M) appeared in 0.2 M amonium sulfate; 0.1 M Bis-Tris

pH 5.5; 25% w/v PEG 3350 at room temperature (Figure 3.1.4). Because the FC2LisH_M5 crystals diffracted only to 8 Å resolution, the data were not used for further procedure. Finally, data were collected to 2.2 Å resolution at 0.9793 Å (peak), 0.9796 Å (the inflection point) and 0.9500 Å (high energy remote) for the FC2LisH_M2 construct. The unit cell dimensions were a = 47.7 Å, b = 94.3 Å, and c = 37.8 Å (Table 3.1.2).

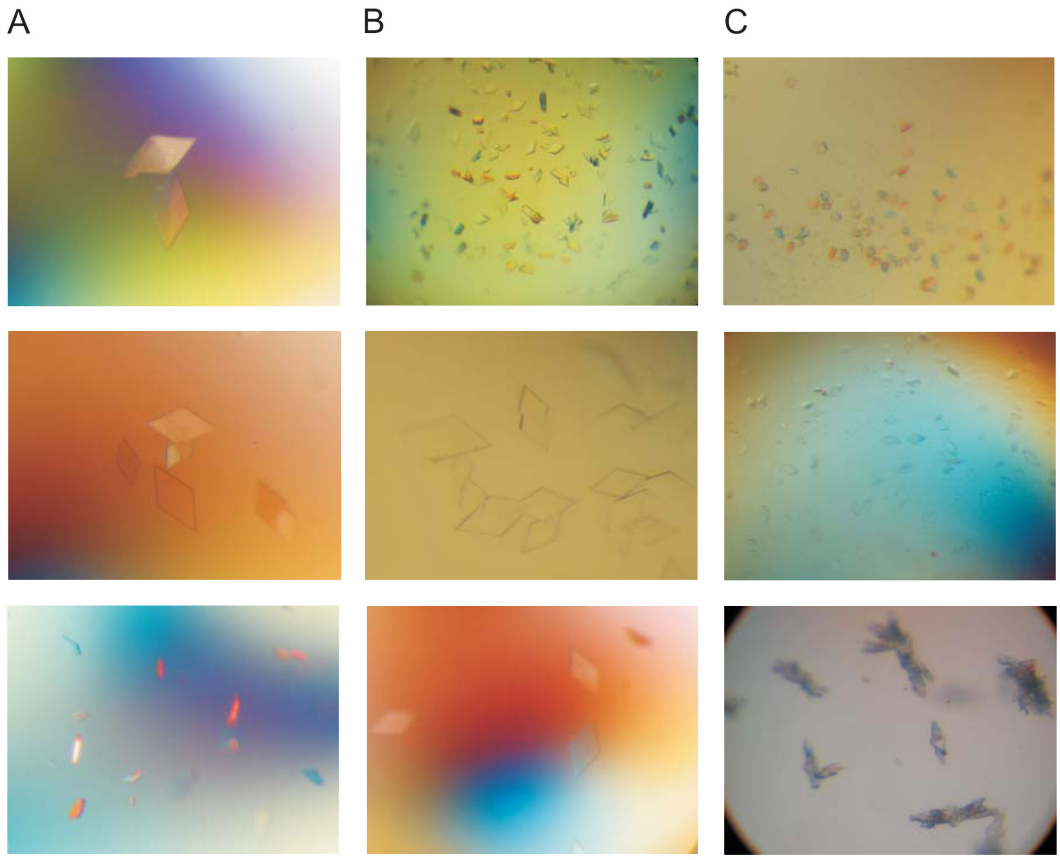


Figure 3.1.4. Crystals of the N-terminal FOP constructs (residues 54-134) growing in different conditions. Panel A shows FC2LisH (wt), panel B shows FC2LisH_M2 (K63M) mutant. Crystals of FC2LisH_M5 (L104M) are shown in panel C.

Collected data were indexed and integrated by MOSFLM (Powell 1999) and XDS (Kabsch 1993) programs. Scaling and merging were performed by SCALA (CCP4 1994) and XSCALE (Kabsch 1993) programs. PHENIX (Adams et al. 2004) and SHARP (Bricogne et al. 2003) programs were used to locate Se atoms and to calculate initial phases using the MAD data to 2.25 Å. ARP/wARP (Perrakis et al. 1999) was able to build 132 of the expected 164 residues. The initial model was improved and revised manually by using Xfit software (McRee 1999) and this partial model was used for calculation of a high

resolution map up to 1.6 Å from the wild-type dataset (Figure 3.1.5). Subsequently ARP/wARP was used to add solvent atoms and manual model building in Xfit was used to complete the model.

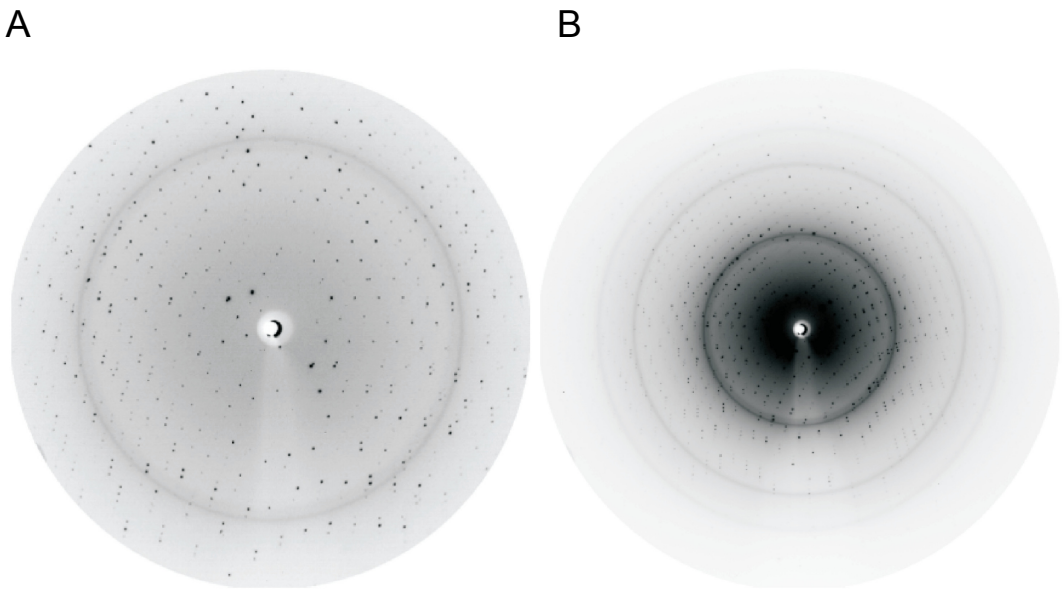


Figure 3.1.5. Representative diffraction patterns of the FOP construct FC2SLisH. The edge of the image plate corresponds to 3 Å in panel A and to 1.6 Å in panel B.

The structure was finally refined by the Refmac5 program (Murshudov et al. 1997) with results of 0.205 for the R-factor and 0.247 for R-free (Table 3.1.2). The final model contains 152 out of the 164 expected residues and 200 water molecules. Missing residues are located at the N-termini of both monomers; additionally residues 101, 102, and 103 were not visible in one chain. There were also no interpretable densities for solvent exposed side chains of residue Lys69 in both monomers.

Table 3.1.2. Data collection and refinement statistic for FC2LisH and for Se-Met crystals FC2LisH_M2.

	FC2LisH	**FC2LisH_M2**		
Crystal parameters				
Space group	$P2_12_12$	$P2_12_12$		
Unit cell (a,b,c) (Å)	47.8, 92.1, 37.1	47.7, 94.3, 37.8		
Data collection	Native	Peak	Inflection point	Remote
Resolution (Å)	47.8 – 1.6	47.7 –2.2	47.7 –2.2	47.7 –2.2

Wavelength (Å)	1.05	0.9793	0.9796	0.9500
No.of reflections	177567	75418	57897	61299
No.of unique reflections	21919	15079	14999	15806
Completeness (%)[a]	93.6 (86.4)	90.1 (39.4)	89.6 (37.8)	94.4 (77.9)
R_{merge} (%)[a]	3.5 (8.9)	2.0 (6.9)	2.2 (7.0)	2.5 (6.5)
I/σ(I)[a]	23.9 (11.9)	44.9 (14.8)	41.1 (14.3)	38.1 (14.7)
Phasing statistics				
Phasing power iso/ano		- / 1.3	1.03/ 0.15	1.3/ 1.2
FOM acentric/centric		0.48/ 0.39		
Solvent content		38.5		
No of Se atoms[b]		2		
Refinement				
Resolution (Å)	46.08 – 1.6			
No. of reflections	21205			
R-factor (%)	20.5			
R_{free} (%)	24.7			
Average B (Å^2)	16.1			
Rms bond lengths (Å)	0.010			
Rms bond angles (°)	1.2			
Asymmetric unit				
No. of protein molecules	2			
No. of residues/atoms	352/ 1399			
No. of solvent atoms	200			
RMSD of monomers (Å)	0.42			
Ramachandran statistics				
Most favoured regions (No./%)		123/ 95.3		
Additionally allowed regions (No./%)		6/ 4.7		
Generously allowed regions (No./%)		0/ 0		
Disallowed regions (No./%)		0/ 0		

[a] values in prentices are shown for the last resolution shell (1.7-1.6 Å for FC2LisH and 2.3-2.2 Å for FC2LisH_M2)

[b] number of atoms for an asymmetric unit

3.1.1.6 The molecular structure of the N-terminal domain of FOP

The N-terminal fragment of FOP forms a dimer in which each chain is composed of five α helices connected by irregular loops (Figure 3.1.6).

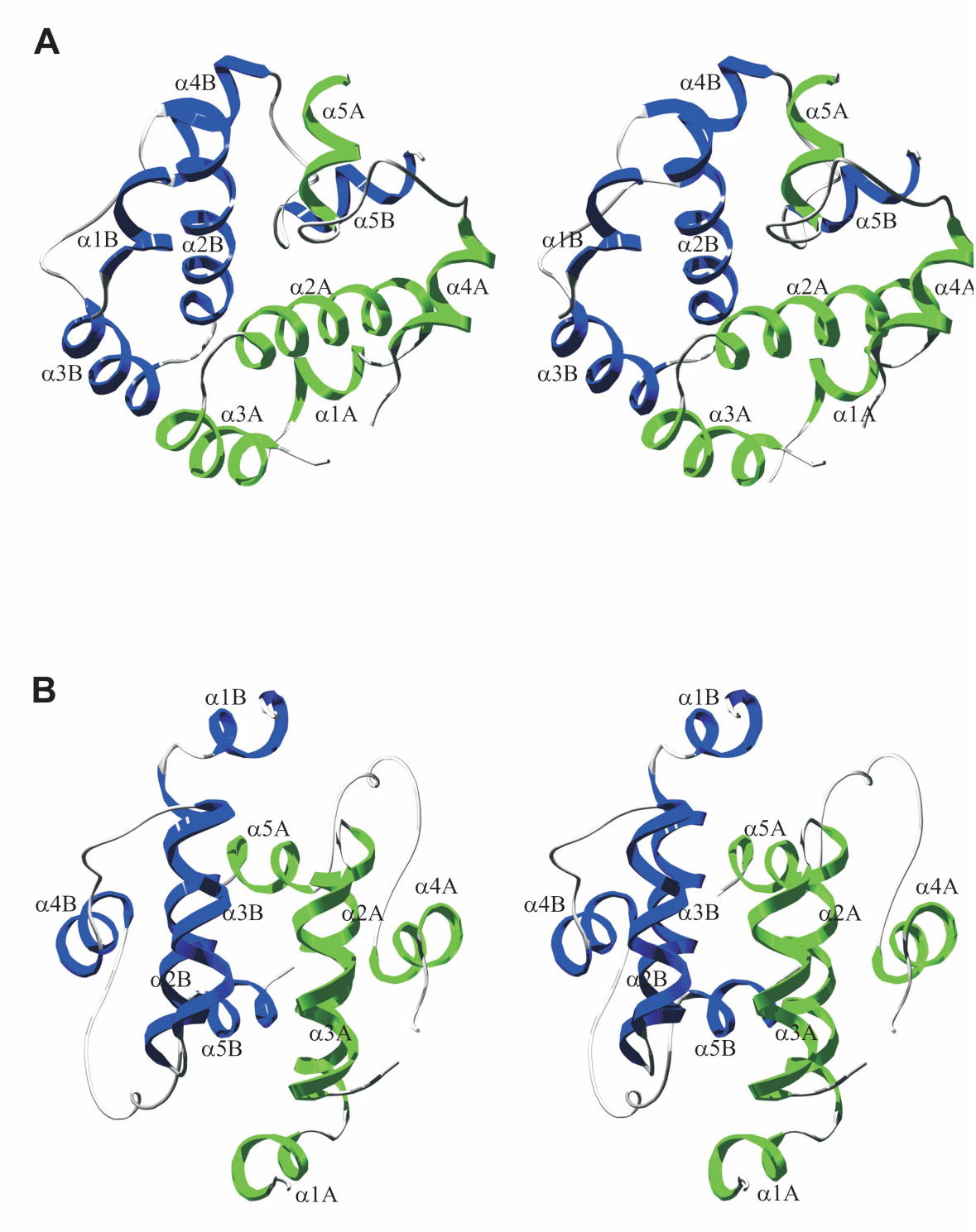


Figure 3.1.6. Stereo ribbon representations of the structure of the N-terminal FOP (54–134). Chain A is shown in green, and chain B is shown in blue. Helices are named a1, a2, a3, a4 and a5 from the N to C-terminal ends; helices a2 and a3 belong to the LisH domain.

Helix α1 is formed by residues 60 to 65 and is situated at 85º to helix α2, which is the longest helix in the molecule (16 amino acids, residues 69 to 84). Helix α3 contains residues 88-98 and is arranged approximately antiparallel to helix α2. Helices α3 and α4 are connected by an 8-residue loop (99-106), which is stabilized by hydrogen bonding interactions with Arg72 (helix α2), Gln95, Pro96, Thr98 (helix α3), Glu108, Asn109 and Leu110 (helix α4). Helix α4 is formed by residues 107-114. The last α5 helix (residues 128-134) runs antiparallel to α4 and is separated from it by a 13 amino acid loop. This long loop is stabilized by hydrogen bonds with Glu80 (helix α2), Arg107, Glu108, Ala111, Arg112 (helix α4), Leu128, Leu129, Glu130, Val131, Arg133 (helix α5), and also by internal hydrogen bonding. Electron density corresponding to the longest loop was well defined in comparison to that of the shorter α3-α4 loop. The shorter loop is not so effectively stabilized by hydrogen bonds and may constitute a flexible element of the structure. This solvent exposed loop, as well as hydrophilic residues of helix α4, are engaged in crystal contacts.

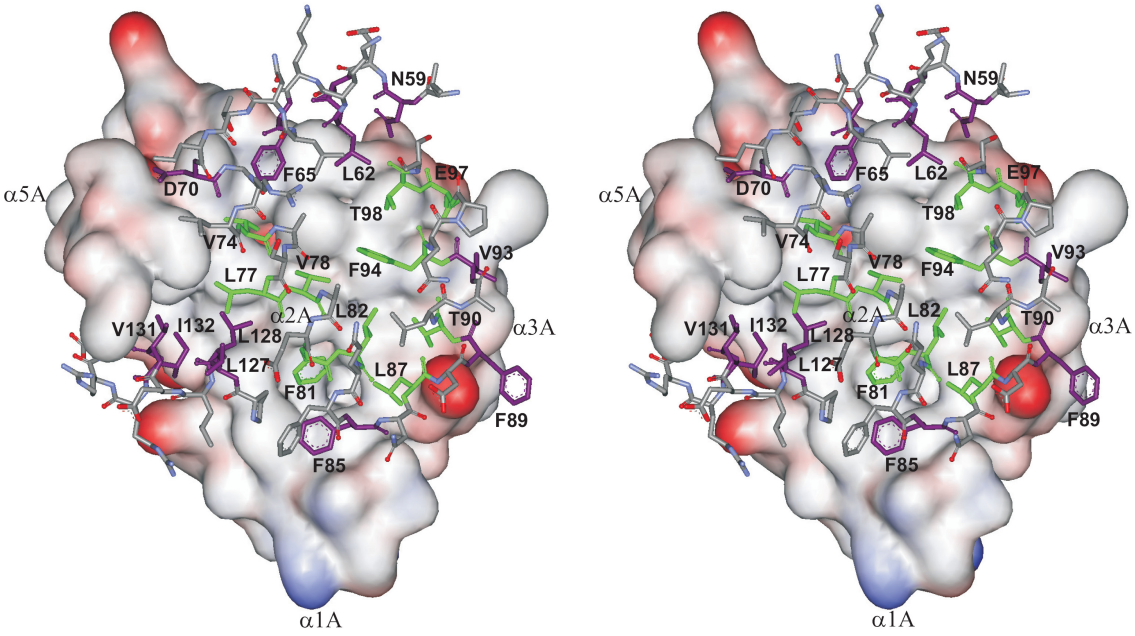


Figure 3.1.7. Dimerization of the N-terminal domain of FOP. One chain in each stereo representation is depicted as a surface plot (residues in red, negatively charged; residues in blue, positively charged; residues in white, neutral). Part of the second chain is shown in stick representation. Amino acid residues that stabilize the dimer are numbered and shown in magenta or in green if they are also conserved amino acids belonging to the LisH domain. The surface model corresponds to the monomer A of Figure 3.2 and is rotated relative to the position in Figure 3.2 (b) by 90º around the line that connects labels α5A–α1A (i.e. approximately the vertical axis of the molecule).

The middle part of the molecule possesses the LisH motif that is an α-helical homodimerization domain composed of two chains running antiparallel. Each chain of FOP-LisH consists of helix α2 (residues 69-84) and helix α3 (residues 88–98) separated by a short 4 amino acid kink. The chain is completed by the loop between residues Glu99–Tyr102.

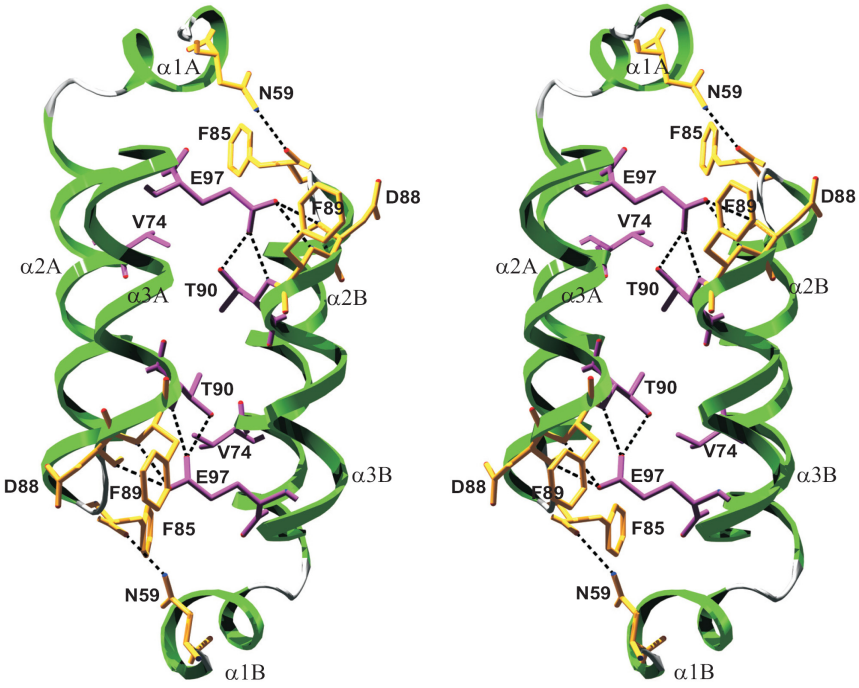


Figure 3.1.8. Stereo view of a hydrogen bonding network between two N-terminal FOP chains. Highly conserved amino acids of the LisH domain (Val74, Thr90 and Glu97) are shown in magenta, the other interacting amino acids are shown in yellow.

The tertiary structure of the monomer is maintained by hydrophobic forces created mainly by leucine residues. Hydrophobic core interactions are formed by Leu73 (α2) and Leu114 (α4), Leu77 (α2) and the two leucine residues Leu110 (α4) and Leu128 (α5). A hydrophobic part of the interface between two helices α2 and α3 is filled with Leu82, Leu87, and Leu91. Additionally helices α2 and α4 are stabilized by a hydrogen bond between Glu80 and Arg107. In the dimer, the monomer/monomer interface has hydrophobic character maintained by residues that form helices α1, α2 and α3. The surface between the chains is tightly packed with Leu, Val, and Phe residues, forming a hydrophobic core. Some of these nonpolar residues like, Val74, Leu77, Val78, Phe81, Leu82, Leu87, Leu91 and Phe94 belong to conserved amino acids characteristic for the

LisH domain (Figure 3.1.7). Second region of FOP responsible for homodimerization is located at the C-terminal part of the chain. Residues 126-133 form intermolecular contacts with Val74 and Leu77, the two conserved residues of the LisH domain, and with Phe65, Asp70, Leu128 and Val131. In addition to hydrophobic contacts, the dimer is stabilized by several hydrogen bonds created between Glu97, which is one of the most conserved residues of the LisH domain, and Thr90, Phe89, and Asp88 of the second chain of the dimer. Phe85 contributes to the formation of additional hydrogen bonds with Asn59 (helix α1) (Figure 3.1.8).

3.1.1.7 The LisH motif is necessary but not sufficient for FOP homodimerization

To identify further the determinant residues of FOP that are required for homodimerization, yeast two-hybrid experiments were carried out (Figure 3.1.9). These studies showed that the region containing residues 54 to 134 of FOP, including the LisH domain, was essential for dimerization. The LisH motif itself was not sufficient, as indicated by the inability of FOP(1-106) to dimerize. However, critical dimerization interactions did apparently occur through the LisH motif because a point mutation within this motif (V74F) abolished dimer formation (Figure 3.1.9). An amino acid sequence comparison of proteins containing LisH motifs revealed sequence conservations in which Leu82 and Glu97 are the most conserved residues. Analysis of the structure reveals that Glu97 is involved in intermolecular hydrogen bonding (Figure 3.1.8) and Leu82 is mainly important in maintaining the tertiary structure of the monomer. Mutations of these residues did not, however, disorder the protein structure and dimer formation. These results were confirmed by gel filtration experiments (Figure 3.1.10; Table 3.1.3). The wild-type FC2LisH construct and two Se-Met mutants, K64M and L104M, migrated as dimers. In case of the V74F, L87M and L127M constructs, mutations affected protein stability and folding, leading to aggregation. A single mutation in Leu127, located at the beginning of helix α5, also leads to protein destabilization, suggesting an important role for the C-terminal part of the construct in dimerization. In the case of conserved residues responsible for intermolecular hydrogen bond formation, substitution for alanine did not prevent dimer formation (the T90A mutant). The double mutant T90AE97A migrated mostly as a dimer with a fraction forming aggregates (Figure 3.1.10).

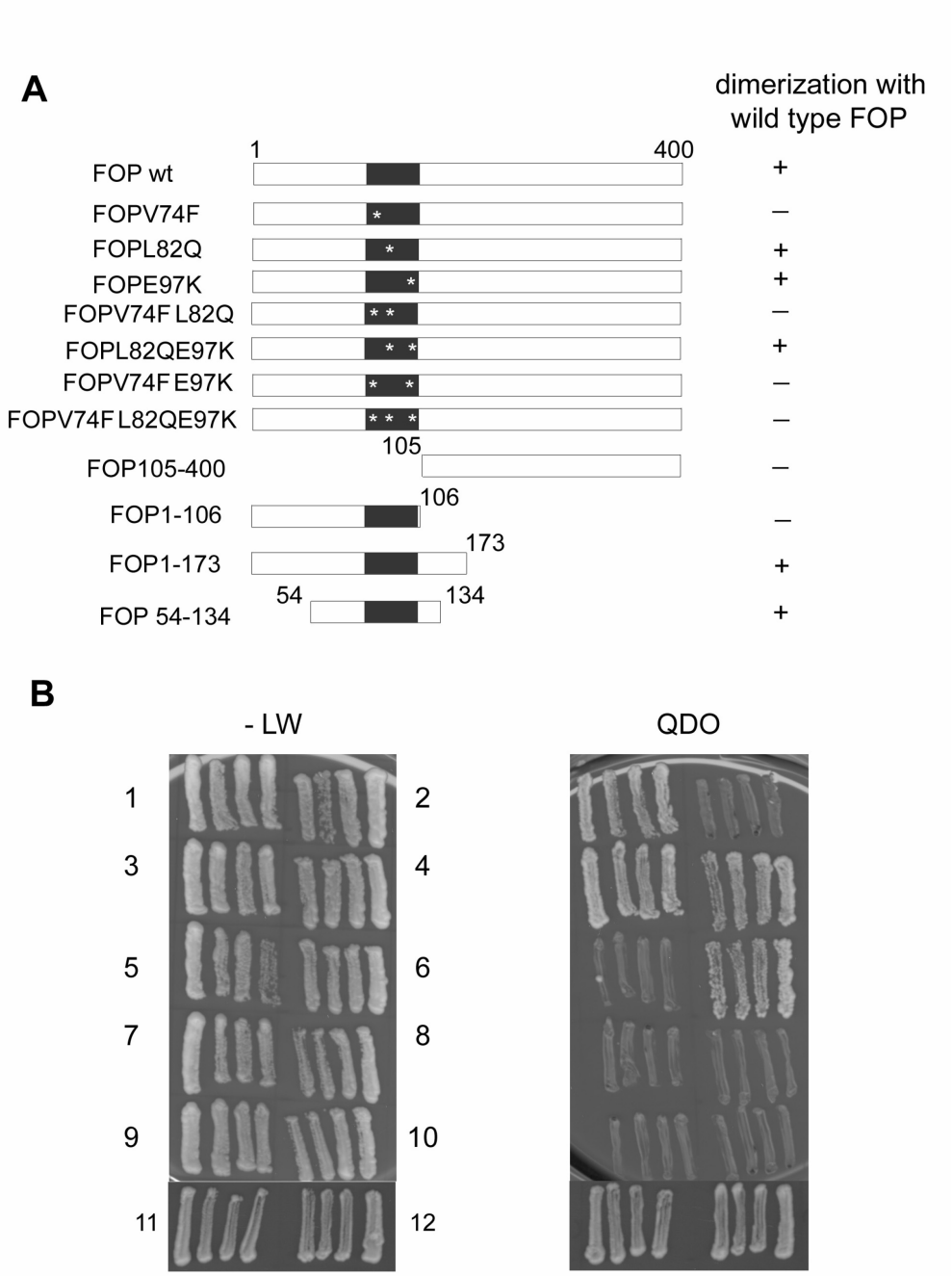


Figure 3.1.9. The dimerization domain mapping. A) Summary of the FOP homodimerization. The wild type and mutant FOP constructs are indicated schematically on the left-hand side; their ability to dimerize is shown on the right. B) Yeast two-hybrid analyses of the interactions between full-length wild-type FOP and various mutant FOP constructs. The right-hand panels show the results of growing transformed yeast on a selective medium (QDO), to test for dimerization, whereas the left panels show growth on a non-selective medium (KLW), to control for viability.

Table 3.1.3. List of mutants expressed in *E. coli* in order to verify mechanism of dimerization.

Name	Mutation	Expression	Oligomerization	Crystallization trials
FC2LisH_M1	K54M	-	-	-
FC2LisH_M2	K63M	+	dimer	positive
FC2LisH_M3	K69M	+	95 % oligomers 5 % dimer	negative
FC2LisH_M4	L87M	+	dimer	negative
FC2LisH_M5	L104M	+	dimer	positive
FC2LisH_M6	L127M	+	20 % dimer 80 % oligomers	negative
FC2LisH_V74F	V74F	+	oligomers	-
FC2LisH_T90A	T90A	+	dimer	-
FC2LisH_E97A	E97A	+	50% dimer 50% oligomers	-
FC2LisH_TEdM	T90A E97A	+	60 % dimer 40 % oligomers	-

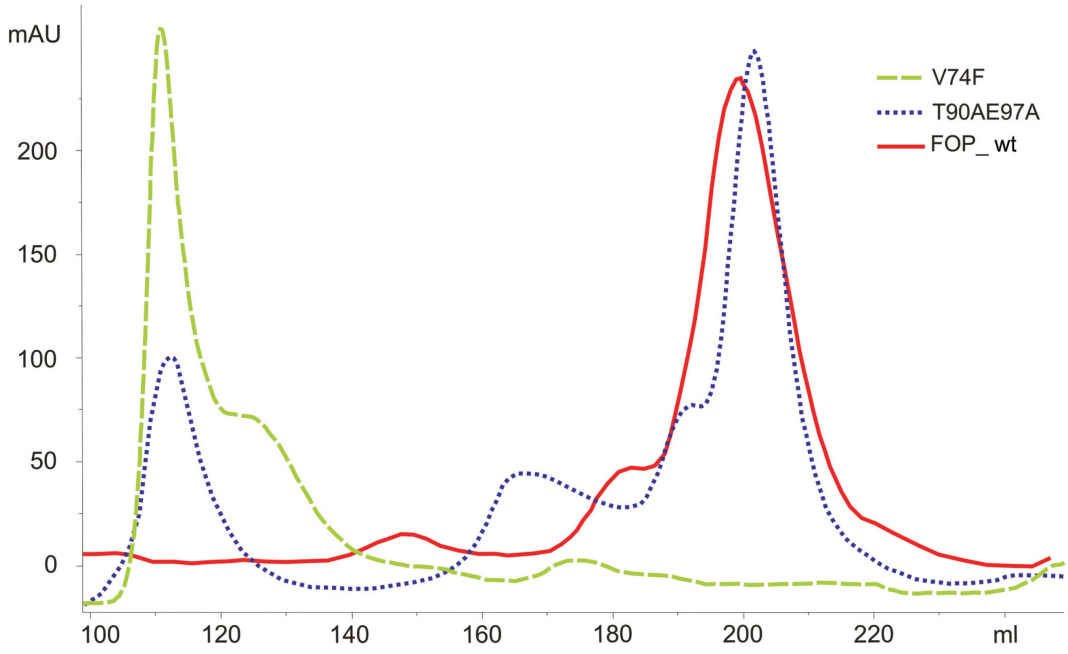


Figure 3.1.10. Size-exclusion chromatography elution profiles of the N-terminal FOP constructs. FOP-wild-type (FOP_wt) migrates as a dimer (continuous line), FOP (V74F) forms oligomers (dot–broken line), and FOP (T90AE97A) (dotted line) migrates partly as a dimer and partly as oligomers.

3.1.1.8 The N-terminal FOP domain is sufficient for centrosome localization

In a final series of experiments, we evaluated the relevance of the fragment containing residues 54 to 134 for centrosome localization of FOP. As shown in Figure 3.1.11, IF staining of U2OS cells with anti-FOP antibody revealed dot colocalizing with γ-tubulin, indicating that fragment indeed localized to the centrosome. The same was revealed for constructs expressing the full length FOP and FOP(1-173) (NMR shows unambiguously that the recombinant FOP(1-173) expressed in *E. coli* is unstructured. This fragment, however, when expressed in human cell lines might be folded or becomes folded upon binding to its target protein). Constructs FOP(1-106) and FOP(105-400) did not show centrosome localization.

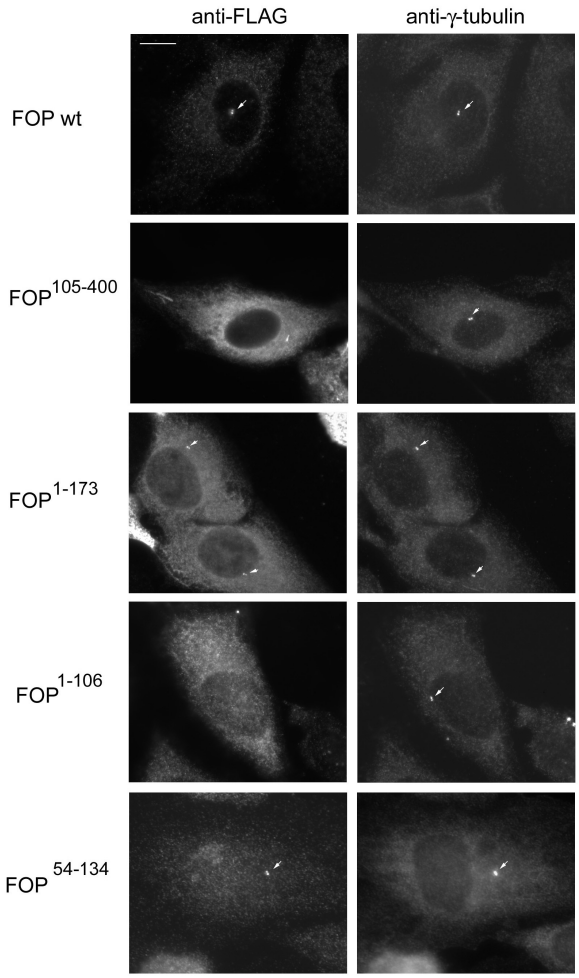


Figure 3.1.11. Centrosome localization of different FLAG-tagged FOP constructs expressed in U2OS cells. Full-length FOP (FOPwt) and N-terminal FOP constructs ((54–134) and (1– 173)) co-localized with γ-tubulin at the centrosome. However, constructs expressing FOP (1–106) and FOP (105–400) did not show centrosome localization.

3.1.2 Discussion

3.1.2.1 LisH motif as a part of dimerization domain

Experiments performed by Kim et al. suggest that the tertiary structure of the LisH motif is sustained predominantly by the dimerization and that the monomer do not have any stable helical fragments (Kim et al. 2004). This can explain why mutations of residues crucial for the dimerization lead to protein aggregation. These results can be readily rationalized by examination of the determined structure. Val74 and Leu87 belong to the conserved residues of the LisH domain and are engaged in the formation of a hydrophobic, tightly packed core and the dimerization interface (Figure 3.1.7). Val74 is situated in helix α2 and its side chain is located between two hydrophobic amino acids of neighboring chain Leu128 and Phe81. A much larger Phe residue does not fit into the narrow groove between these residues. In addition the side chain of Val74 makes intermolecular contact with two residues located at the C-terminal part of the chain, Leu129 and Ile132. Leu87 is situated in a short kink between two LisH helices with its side chain directing towards the core of the LisH domain. This residue is engaged in monomer stabilization and participates in the hydrophobic interface of the dimer. The dimerization surface formed by the LisH domain of FOP is 570 $Å^2$ per monomer; that is about one third of the dimerization interface of the whole N-terminal FOP molecule (1500 $Å^2$) and does not appear sufficient for dimer stabilization.

Our data indicate that the centrosomal localization of FOP depends on residues 54 to 134, but the role of the LisH domain in this process is presently not clear. In the case of the OFD1 protein, which was recently found to be a genuine core component of the centrosome (Andersen et al. 2003; Romio et al. 2004), removing the LisH domain did not affect centrosomal localization (Romio et al. 2004). The centrosome-binding domain of CEP1, which is like FOP associated with MPD disease, has the leucine zipper motif (Guasch et al. 2000) but does not possess the LisH domain. Leucine zippers are well-characterized motifs involved in protein-protein interactions, not only common in transcription factors (Landschulz et al. 1988) but also in centrosomal proteins, including ninein and CEP250 (Guasch et al. 2000). However, the leucine-rich repeats of FOP differ from other known leucine-rich consensus sequences and their role in centrosomal localization remains to be determined. In summary, we conclude that the LisH motif can

works as part of a larger functional and structural domain, which plays an important role in protein dimerization and localization.

3.1.2.2 Structural neighbors of FOP

DALI (Holm and Sander 1993) and VAST (Gibrat et al. 1996) servers were used to locate structural homologues of the FOP fragment. We found that the LisH domain (residues 69-102) and the C-terminal part of FOP (residues 126-134) show structural similarities to the N-LIS1 protein (Kim et al. 2004) The lowest RMSD (1.5 Å for matching residues) was obtained for the alignment of chain A of the LIS1 with chain B of the FOP (Figure 3.1.12). The rest of the FOP structure possesses no significant matches.

A

```
FOP          Hs 5901954     KDGRLVASLVAEFLQFFNLDFTLAVFQPETSTLQ    (69-102)
LIS1         Mm 7305363     RQRDELNRAIADYLRSNGYEEAYSVFKKEAELDM    (6-39)
TONNEAU1     At 11494364    PSGRLLSALICEYLDWAQLNHTLIVYQPESNLPK    (7-40)
OFD1         Hs 64654925    LLIGASNSLVADHLQRCGYEYSLSVFFPESGLAK    (69-102)
TBL1         Hs 5032159     ITSDEVNFLVYRYLQESGFSHSAFTFGIESHISQ    (54-87)
AAAATPASE    Lm 11071772    RRVKGVIVLVEQFLLEQGYHQTLHALQQESRISL    (38-71)
NOPP140      Hs 12643725    VVPSDLYPLVLGFLRDNQLSEVANKFAKATGATQ    (9-42)
NPAT         Hs 4505431     LLPSDVARLVLGYLQQENLISTCQTFILESSDLK    (2-35)
FLJI0805     Hs 7023065     IESSDVIRLIMQYLKENSLHRALATLQEETTVSL    (5-38)
```

B

```
FOP          Hs 5901954     KDGRLVASLVAEFLQFFNLDFTLAVFQPETSTLQ    (69-102)
LIS1         Mm 7305363     RQRDELNRAIADYLRSNGYEEAYSVFKKEAELDM    (6-39)
```

C

```
FOP  Hs        NESLKKFLNTKDGRLVASLVAEFLQFFNLDFTLAVFQPETSTLQ    (59-102)
LIS1 Mm        ------VLSQRQRDELNRAIADYLRSNGYEEAYSVFKKEAELDM    (2-39)

FOP  Hs        GLEGRENLARDLGIIEAEGTVGGPLLLEVIRRW----------------    (103-135)
LIS1 Mm        NEEL--------------DKKYAGLLEKKWTSVIRLQKKVMELESKLNEAKE    (40-77)
```

Figure 3.1.12. A) Alignment of LisH domains from different proteins, the most conserved amino acids are shown in dark blue, the rest of the conserved amino acids are shown in light blue. B) Alignment of LisH domains of FOP and LIS1 proteins, amino acid residues responsible for hydrogen bonding across the dimer interface are shown in red, residues involved in hydrophobic contacts between monomers are highlighted in green. C) Alignment of N-terminal parts of FOP and LIS1, residues identical in both proteins are highlighted in red, highly similar residues are shown in orange, residues with the structure homology are shown in bold.

The LisH domains of FOP (residues 69-102) and N-LIS1 protein (residues 6-39) display high structural similarity although sequence homology is low (about 17%). The RMSD for main chain atoms is equal to 1.21 Å for a monomer and 1.34 Å for the dimer. The biggest differences are observed at the N-terminus of the LisH domain (helix α2 of FOP) and the lowest values of RMSD are at the C-terminal regions of the LisH dimers. This can be accounted for by the difference in dimer stabilization. In the case of N-LIS1, the N-terminal helix of LisH domain is stabilized by both intermolecular hydrophobic forces and hydrogen bonds. In contrast, the corresponding helix of FOP is stabilized only by intermolecular hydrophobic interactions. In both proteins the C-terminal helices of LisH domains reveal similar patterns of interactions (Figure 3.1.12 B). Only three, the most conserved residues, Leu82, Phe94, and Glu97 of FOP-LisH are preserved in N-LIS1. The rest of the residues in helix α2 are replaced by different hydrophobic amino acids. In the case of α3, the differences between amino acids are more significant: small hydrophobic residues of LIS1 are replaced by much bigger hydrophilic amino acids in FOP (Figure 3.1.12). In addition to the LisH domain, the C-terminal part of FOP (residues 126-134) shows structural similarity to the first two turns of helix α4 of N-LIS1 (residues 49-57). Thermodynamic parameters obtained for the N-terminal LIS1 fragment indicated that the LisH motif accounts for only a part of free energy of the homodimerization (Mateja et al. 2006). A second site of dimerization was found right at the beginning of helix α4 (Figure 3.1.12 C) (Mateja et al. 2006), which corresponds to the C-terminal FOP residues 126-134. In N-LIS1, a single residue downstream of the LisH motif, Trp55, was found to contribute significantly to the free energy of dimerization. A complete set of the residues of the N-LIS1 49-57 fragment that contribute to the dimerization interface are highlighted by an underline Gly49-Leu-Leu-Glu-Lys-Lys-Trp-Thr-Ser57. The corresponding fragment in FOP is Pro126-Leu-Leu-Leu-Glu-Val-Ile-Arg-Arg134 and thus has a more hydrophobic character. Overall, however, we assume that the mechanism of dimerization is analogous in both proteins. The search for the α-helical fold of dimerization seen in these two proteins produced no significant structure matches and thus this dimerization layout can be considered to be a novel fold.

3.1.2.3 Role of the LisH domain in protein-protein interactions

FOP together with CAP350 was predicted to be a centrosomal protein that is involved in microtubule anchoring to subcellular structures (Yan et al. 2005). *In vivo*

studies showed that the C-terminal 47 amino acids of CAP350 were sufficient for the interaction with the N-terminal 175 residues of FOP, including the LisH domain. The LisH motif was required but not sufficient for binding of FOP to CAP350, as indicated by the inability of FOP1-106 to bind (Yan et al. 2005). These studies revealed a strict correlation with our results. Constructs FOP(1-106) and FOP(105-400) did not show centrosome localization opposite to full length FOP and two N-terminal constructs 1-173 and 54-134. The ability of FOP constructs to localize to the centrosome is connected with ability to interact with CAP350. However in our *in vitro* experiments we were unable to find tight binding between CAP350 best folded construct CAP2, spanning 158 C-terminal amino acids, and FOP(54-134). These data suggest that centrosomal localization of FOP can be independent to CAP350 binding and that 54-134 fragment of FOP can interact with CAP350 through some additional protein components. It was reported that a presence of CAP350 and FOP is required for the accumulation of MT plus-end binding protein EB1 at the centrosome where all three proteins cooperate in MT anchoring (Yan et al. 2005). However we cannot exclude the possibility that if the mechanism of centrosomal localization of FOP is mediated by CAP350 then, interactions require bigger parts of CAP350 and FOP. Yeast two hybrid system biding analysis of binding between EB1 and FOP demonstrated that FOP constructs 1-173 and 105-400 were not sufficient for this process indicating that some additional residues distinct from LisH domain are required for interaction (Yan et al. 2005).

Similar to FOP few other proteins with LisH domain like LIS1 (Sapir et al. 1999), plant tonneau and centrosomal protein RanBPM (Nakamura et al. 1998; Emes and Ponting 2001) were associated with microtubule binding but exact mechanism of interactions is not precisely described. It was found that the LIS1 protein interacts directly with mNudE in centrosome where both proteins take part in dynamic reorganization of microtubules (Feng et al. 2000). Two point mutations of LIS1 H149R and S169P completely abolish LIS1-mNudE interaction in the two-hybrid assay (Feng et al. 2000) indicating that LisH domain is not essential or at list not sufficient for binding. Both mutations are localized in region that is a part of the predicted WD40 repeats structure, a motif that is known to mediate protein–protein interactions. However FOP lacks any other domains important in mediating protein-protein interactions except LisH domain, what exclude similar mechanism of interaction. Lis1 can also interact with tubulin directly influences microtubule dynamics *in vitro* but fragment of protein sufficient for this process was not determined (Sapir et al. 1997).

The crystal structure of the N-terminal domain of the FOP protein described here provides the basis for a structural understanding of the dimerization mechanism involved in the constitutive FGFR1 activation. The future goal can be the verification of the interaction between the FOP and its binding partners. It would be also interesting to know if the N-terminal part of FOP is responsible for binding directly to microtubules. FOP and CAP350 were found to be the subject to mitotic phosphorylation (Yan et al. 2005) and exploration of the cell cycle regulation connected with these proteins can help to understand the role that centrosome plays in the normal cell progress as well as in human diseases.

3.2 Cell differentiation and cell cycle proteins – interaction studies

In this section a variety of biophysical and biochemical methods, like for example, NMR, mass spectrometry, pull-down assay, gel filtration chromatography and ITC, were used to study protein-protein interactions between pRb and the HLH proteins, MyoD and Id-2, as well as proteins containing the LXCXE motif.

3.2.1 Results

3.2.1.1 Cloning and plasmids

All Id-2, MyoD, and pRb (AB, ABC, and ABL) constructs were cloned into the pET30 Xa/LIC vector (Novagen) according to the manufacturer manual. Short His-tag, DAPzyme, factor Xa digestion sites and all mutations were introduced by Quick-Change Mutagenesis Kit (Stratagene). The HPV16 E7 full-length construct was cloned into pET8 vector. The HPV16 E7-short was cloned into pET30 Xa/LIC (Novagen) vector and spanned residues 21-40 (the same sequence like E71 peptide). C-Rb and PAI2 constructs were cloned into the pET46 Xa/LIC vector. All constructs used in this work were summarized in the Table 3.2.1.

3.2.1.2 Protein expression

All constructs were expressed in *E. coli* strains BL21 STAR (DE3) (Invitrogen). Cells were cultured according to standard procedures in LB, minimal media or a selectively labeled media. The temperature of pRb cultures were decreased to 18°C at OD_{600} = 0.4. The cultures were induced at OD_{600} = 0.7 with 1 mM IPTG and kept for 6 h at 37°C or for 12 h at 18°C (pRb). After this time cells were spun and pellets frozen.

Table 3.2.1. List of proteins used in this part of thesis.

Name	Residues range	Length aa	Vector	Tag
Id-2	2 -132	131	pET30 Xa/LIC	N-terminal 8 aa MHHHHHHK
MyoD	99 -173	74	pET30 Xa/LIC	N-terminal 8 aa MHHHHHHR
pRb-ABC	379-928	550	pET30 Xa/LIC	N–terminal 11 aa MHHHHHHIEGR
pRb-ABL	379 - 791	413	pET30 Xa/LIC	N–terminal 11 aa MHHHHHHIEGR
pRb-AB	379 - 578 642 - 791	350	pET30 Xa/LIC	N–terminal 11 aa MHHHHHHIEGR
C-Rb	790 - 928	139	pET46 Xa/LIC	N–terminal 15 aa MAHHHHHHVDDDDKM
E7 full length	1 - 98	98	pET8	C-terminal 6 aa HHHHHH
E7 short	21 - 40	20	pET30 Xa/LIC	N-terminal 43 aa, His-tag, S-tag plus Enterokinase cleavage site
PAI2 wt	1 - 415	415	pET46 Xa/LIC	N–terminal 15 aa MAHHHHHHVDDDDKM
PAI2 ΔCD	1 - 65 98 - 415	383	pET46 Xa/LIC	N–terminal 15 aa MAHHHHHHVDDDDKM
PAI2 R380A	1 - 415	415	pET46 Xa/LIC	N–terminal 15 aa MAHHHHHHVDDDDKM

3.2.1.3 Purification and protease digestion

The *E. coli* cells which overexpressed the pRb or PAI-2 constructs were lysed in a lysis buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10 mM βME 10 mM imidazole. Additionally the lysis buffer was supplemented with a protease inhibitor cocktail (one tablet of "Complete" without EDTA, Roche, per 50 ml lysate), sonicated and centrifuged. In the first step of purification supernatants were passed through Ni-NTA resin (Qiagen). The sample was loaded onto the column and washed with the lysis buffer supplemented with 20 mM imidazole. Elution was performed with the lysis buffer supplemented with 250 mM imidazole. The pRb-AB and pRb-ABL constructs were subjected to the MonoQ

(Amersham-Pharmacia) anionic exchange chromatography. Buffer A11 consisted of 50 mM NaH_2PO_4 pH 7.8, 10 mM βME, buffer B1 had additionally 1 M NaCl. PAI2 constructs were purified using S75 gel filtration chromatography in PBS buffer.

The *E. coli* cells, which overexpressed the Id-2, MyoD, C-Rb or E7-short constructs, were lysed in a lysis buffer (6 M GuaHCl, 50 mM Tris pH 8.0), sonicated and centrifuged. Supernatants were purified over Ni-NTA resin. Buffers consisted of 6 M GuaHCl, 50 mM Tris (wash pH 6.0 and pH 5.2; elution pH 4.5). Refolding was made by dialyzing a low concentrated protein sample to 50 mM NaH_2PO_4, 300 mM NaCl, pH 7.8, and 10 mM βME. The E7 full-length protein was a gift from Roland Degenkolbe.

Prior to a pull-down assay, a His-tag was removed from the pRb protein by Xa protease, MyoD or Id-2 samples by DAPzyme (Qiagen) enzymatic digestion, and the proteins were subsequently purified over Ni-NTA to remove His-tags or undigested proteins. The efficiency of His-tag removal was checked using SDS-PAGE and Western-blot.

Identities of all the constructs were checked on the DNA level by the DNA sequencing. All proteins were also checked by Western-blot, mass-spectrometry and/or the N-terminal Edman-sequencing. Purity of all proteins was estimated to be 95-98% from SDS-PAGE and mass spectrometry.

3.2.1.4 Activity tests of pRb, MyoD, Id-2 and PAI2 ΔCD

Folding of proteins was checked by 1D-^{1}H NMR spectra and peak dispersion in amide and aliphatic region was used as a factor. Recorded spectra revealed that PAI2 ΔCD, pRb-AB, pRb-ABL were very well structured in solution. pRb-ABC was indicated as partly unfolded. NMR spectroscopy of ^{15}N uniformly labeled MyoD sample showed that MyoD (MCNC) was able to bind with DNA (sequence 5'-TCAACAGCTGTTGA-3'). Additionally MyoD (MCNC) was checked for formation of a complex with Id-2 using both: affinity chromatography - pull down assay, and gel filtration chromatography. The pRb was functional as proven by its interaction with the HPV16 E7 full length, E7-short, peptides E7-1, E7-0, and the SV40 large T antigen peptide. Gel filtration chromatography of equimolar mixtures of E7 (full length), or E7-short with pRb-AB both showed the complex formation. In both cases proteins were eluted in one peak. Binding between pRb and large T antigen was demonstrated by gel filtration chromatography coupled with mass spectrometry.

3.2.1.5 Activity tests of PAI-2

The PAI-2 protein activity was measured photometrically in the reaction with the urokinase type plasminogen activator (uPA). The experiment was performed as described in the Materials and methods (section 2.2.15) for two uPA concentrations: 1 mg/ml for a blank test and for reaction 1, and 4 mg/ml for reaction 2. A blank test was performed without the PAI-2wt protein. In reactions 1 and 2, 1 µl of PAI-2wt concentrated to 2 mg/ml was inserted into a reaction mixture in time 0. The absorbance was measured every 10 min for 2h. All results are summarized in Table 3.2.2.

Table 3.2.2. The uPA activity test performed by using absorbance changes in the presence or in the absence of PAI-2 protein.

Time (min)	**Absorbance (405 nm)**			**uPA activity (dA_{405}/min)/µg enzyme**		
	Blank test	**Reaction 1**	**Reaction 2**	**Blank test**	**Reaction 1**	**Reaction 2**
0	0.5831	0.119	0.6563			
10	0.7201	0.1278	0.7016	13.7E-03	0.88E-03	1.13E-03
20	0.7701	0.1725	0.7351	9.35E-03	2.68E-03	0.98E-03
30	0.7959	0.1501	0.7771	7.09E-03	1.04E-03	1.01E-03
40	0.8377	0.1642	0.7814	6.37E-03	1.13E-03	0.78E-03
50	0.9145	0.1744	0.8049	6.63E-03	1.11E-03	0.74E-03
60	0.9629	0.1801	0.8495	6.33E-03	1.02E-03	0.80E-03
70	1.0075	0.1992	0.8573	6.06E-03	1.15E-03	0.71E-03
80	1.0618	0.2092	0.8825	5.98E-03	1.13E-03	0.70E-03
90	1.0540	0.2103	0.8872	5.23E-03	1.01E-03	0.64E-03
100	1.0795	0.2209	0.8850	4.96E-03	1.02E-03	0.57E-03
110	1.1830	0.2301	0.8764	5.45E-03	1.01E-03	0.50E-03
120	1.1328	0.2373	0.8568	4.58E-03	0.98E-03	0.41E-03
Average				**6.81E-03**	**1.18E-03**	**0.75E-03**

In the absence of PAI-2wt (blank test) the absorbance changing rate was faster than in reactions 1 and 2 performed in the presence of PAI-2. The calculated average uPA activity was about six times higher in the blank test than in reactions 1 and 2. These experiments indicate that PAI-2wt was active and had the inhibition capability.

3.2.1.6 Peptide synthesis

Peptides were synthesized using a solid phase, purified with the C8 reverse phase chromatography and checked by mass spectrometry (Table 3.2.2).

Table 3.2.2. List of peptides corresponding to fragments of different proteins used in this study. The first column shows the abbreviations of peptides used in this part; the second column shows the location of the peptide sequence in the full-length protein; third the LXCXE sequence in the peptides or protein. The Leu (Ile), Cys, and Glu residues in the LXCXE motif are shown in bold. ITC and NMR titration data are shown for a small pocket of pRb and the various peptides. K_D values were determined by NMR in case of fast exchange or weak binders.

Peptide name	**Corresponding protein**	**Sequence**	**ITC K_D µM**	**NMR K_D µM**
E7–0	HPV16 E7 (21–29)	D**L**Y**C**Y**E**QLN	0.07 0.21 ± 0.04 pRb(K713A)[a] 0.044 ± 0.02 pRb(C706F)[a]	tight binding
E7–1	HPV16 E7 (21–40)	D**L**Y**C**Y**E**QLNDSSEEEDEIG	-	tight binding
E7-0-I	HPV16 E7 (21–29)	D**I**Y**C**Y**E**QLN	0.32 ± 0.07	-
Large – T	SV40 large T antigen (102-110)	N**L**F**C**S**E**EMD	0.44 ± 0.06	tight binding
HDAC1 17mer	HDAC1 (411-427)	DKR**I**A**C**E**E**EFSDSEEEG	10.0 ± 3.0	40.0 [b]
HDAC1 20mer	HDAC1 (409-428)	SSDKR**I**A**C**E**E**EFSDSEEEGE	10.0 ± 3.0	-
HDAC1 9mer	HDAC1 (413-421)	R**I**A**C**E**E**EFS	20.0 ± 4.0	50.0 [b]
PAI2_p	PAI2 (158-174)	F**L**E**C**A**E**EARKKINSWVK	no binding	no binding

[a] pRb-AB mutants are used in these experiments

[b] Specific values refer to the complexes in fast exchange between binding components, the estimated error is 10%.

3.2.1.7 Interactions between pRb and the HLH domain of MyoD or Id-2

To observe interaction between the pocket region of pRb and the bHLH domain of MyoD, we used a pull down assay as a first attempt (Figure 3.2.1). A MyoD construct with His-tag

was the "bait" and pRb (ABL) without a His-tag was a "prey". The MyoD protein was bound to the Ni-NTA column and washed with a buffer without imidazole, in the next step pRb was passed through the column and washed after 15 min of the incubation. After elution with 250 mM imidazole only the presence of the MyoD protein was noticed on the SDS page. A reverse experiment, where MyoD was the "prey" and pRb was a "bait" gave the same results.

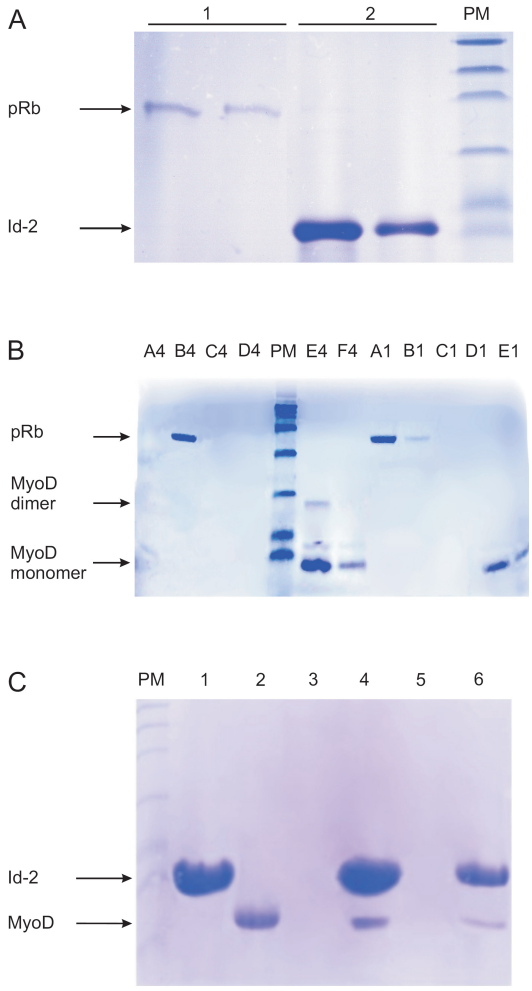
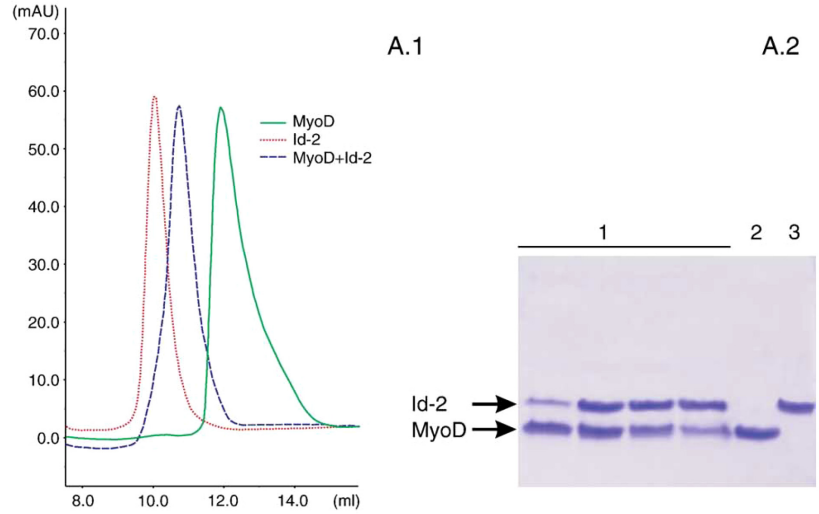
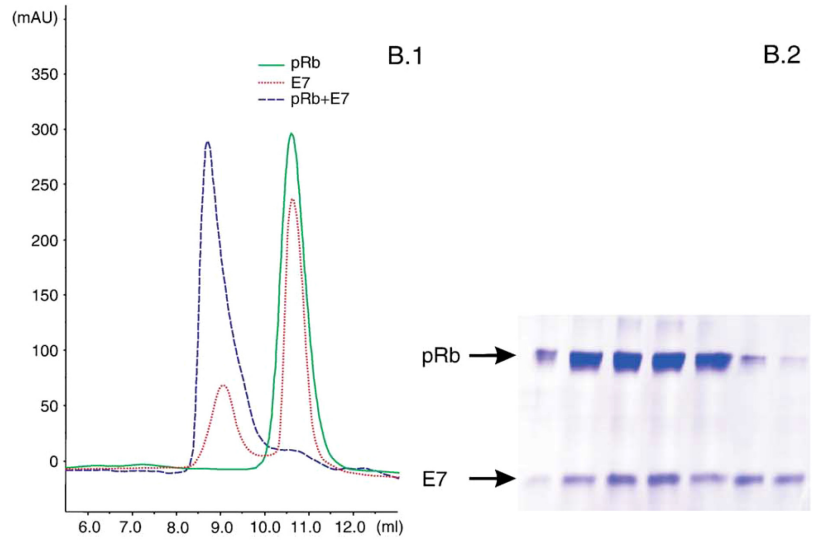
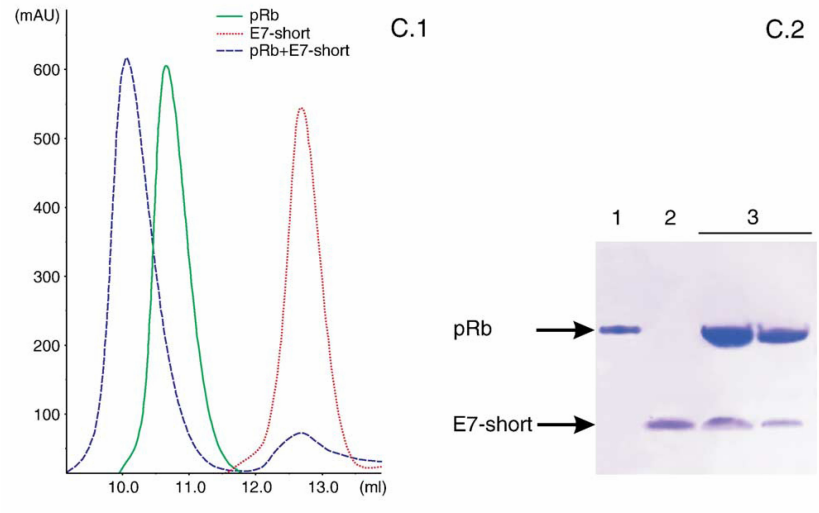


Figure 3.2.1. SDS-PAGE of affinity chromatography pull-down assays. (A) His-tagged Id-2 (the bait) was bound to the Ni-NTA column and then pRb (AB) (the prey) was passed through the column. 1-pRb (AB) flow-through; 2 – 250 mM imidazole elution fractions. (B) The His-tagged MyoD (the bait) was bound to the Ni-NTA column and then pRb (AB) (the prey) was passed through the column. A4, MyoD flow-through; B4 and C4, pRb flow-through; D4, E4, and F4, various fractions eluted with 250 mM imidazole. A1, B1 and C1, pRb flow-through (amount of MyoD decreased four times); D1 and E1, 250 mM imidazole elutions. (C) His-tagged Id-2 was bound to the Ni-NTA column and then MyoD (the prey) was passed through the column. 1—Id-2 alone, 2— MyoD alone, 3 and 5—MyoD flow-through, 4—fraction eluted with 250 mM imidazole, 6—fraction eluted with 250 mM imidazole (amount of Id-2 decreased four times). PM is the protein marker (6-175 kDa).


(mAU)
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
8.0
10.0
12.0
14.0
(ml)
MyoD
Id-2
MyoD+Id-2
A.1
A.2
1
2
3
Id-2
MyoD


(mAU)
350
300
250
200
150
100
50
0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
(ml)
pRb
E7
pRb+E7
B.1
B.2
pRb
E7


(mAU)
600
500
400
300
200
100
10.0
11.0
12.0
13.0
(ml)
pRb
E7-short
pRb+E7-short
C.1
C.2
1
2
3
pRb
E7-short

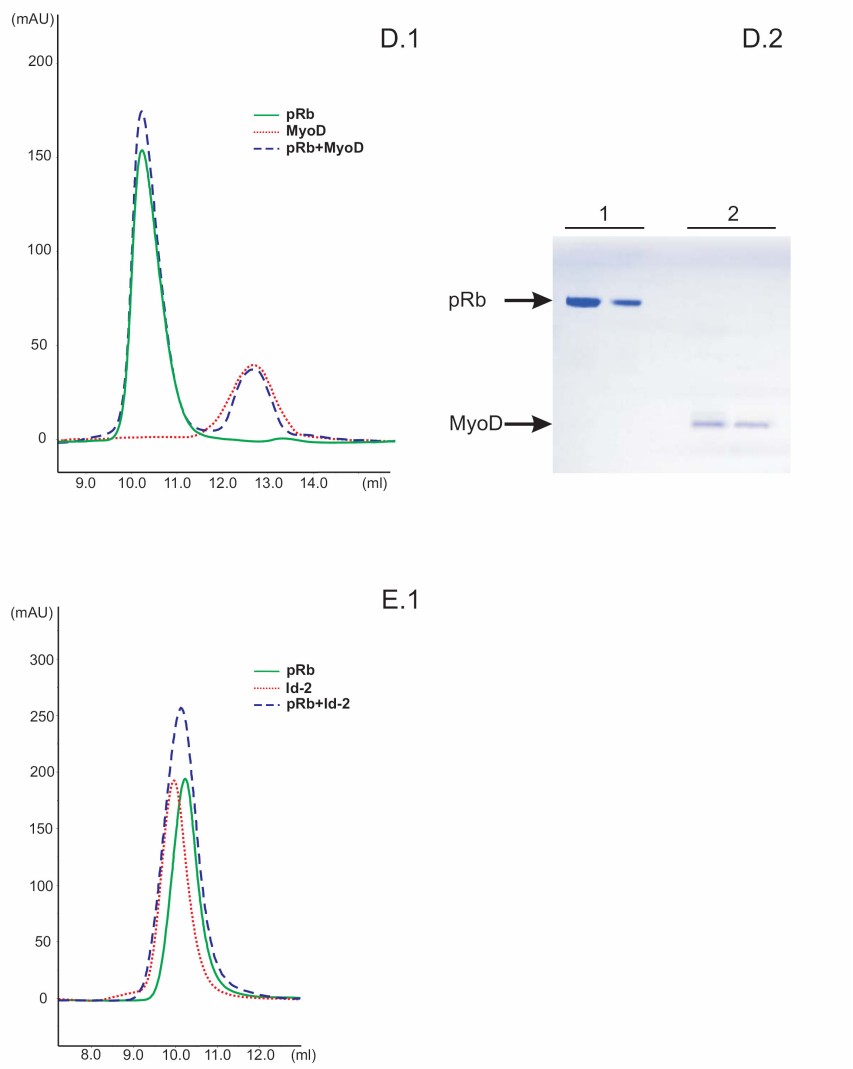


Figure 3.2.2. Monitoring of the protein–protein complexes formation by gel filtration chromatography. Panels A.1 to E.1 show chromatograms and panels A.2 to D.2 the SDS-PAGE analysis of consecutive chromatography fractions. Panel A.1 shows the MyoD - Id-2 complex (dash line), Id-2 alone as a dotted line (tetramer), and MyoD alone as a straight line (tetramer). Panel A.2 shows the SDS-PAGE analysis of chromatography fractions - 1 MyoD - Id-2 complex (consecutive fractions); 2 - MyoD alone; 3 - Id-2 alone. Panel B.1 shows the pRb (AB) - E7 (full-length) complex (dash line), E7 (full-length) alone (dimers and oligomers) as a dotted line, and pRb (AB) alone (straight line). Panel B.2 shows the SDS-PAGE analysis of chromatography fractions - all lanes, pRb (AB) - E7 complex. Panel C.1 shows the pRb (AB) - E7 (short) complex (dash line), the E7 (short) alone (dot line), and pRb (AB) alone (straight line). Panel C.2 shows the SDS-PAGE analysis of chromatography fractions: 1 - pRb (AB); 2, - E7 (short); 3 - the E7 (short) - pRb (AB) complex. Panel D.1 illustrates that no complex was formed between MyoD and pRb (AB) (dash line), pRb (AB) alone (straight line), and MyoD alone (dot line). Panel D.2 shows the results of SDS-PAGE analysis of chromatography fractions: 1 - pRb (AB); 2 - MyoD. Panel E.1 demonstrates that no complex was formed between Id-2 and pRb (AB) (dash line). Id-2 alone is shown as a dotted line (tetramer) and pRb (AB) alone as straight line. No SDS-PAGE analysis was performed because both proteins elute at the same column volume.

Using the gell filtration method I confirmed the binding between pRb and E7 protein and between MyoD and Id-2 (Figure 3.2.2, A.1 to C.1). The same method demonstrated the lack of interaction between pRb and MyoD, as well as pRb and Id-2 (Figure 3.2.2, D.1 and E.1). Gel filtration experiments were proved by mass spectroscopy. Mass spectra of concentrated fractions, corresponding to the pRb peak in gel filtration, showed no MyoD (Figure 3.2.3, B). As a positive experiment I noticed the presence of the SV40 large T antigen peptide bound to pRb after the sample from gel filtration was analyzed by mass spectrometry (Figure 3.2.3, A).

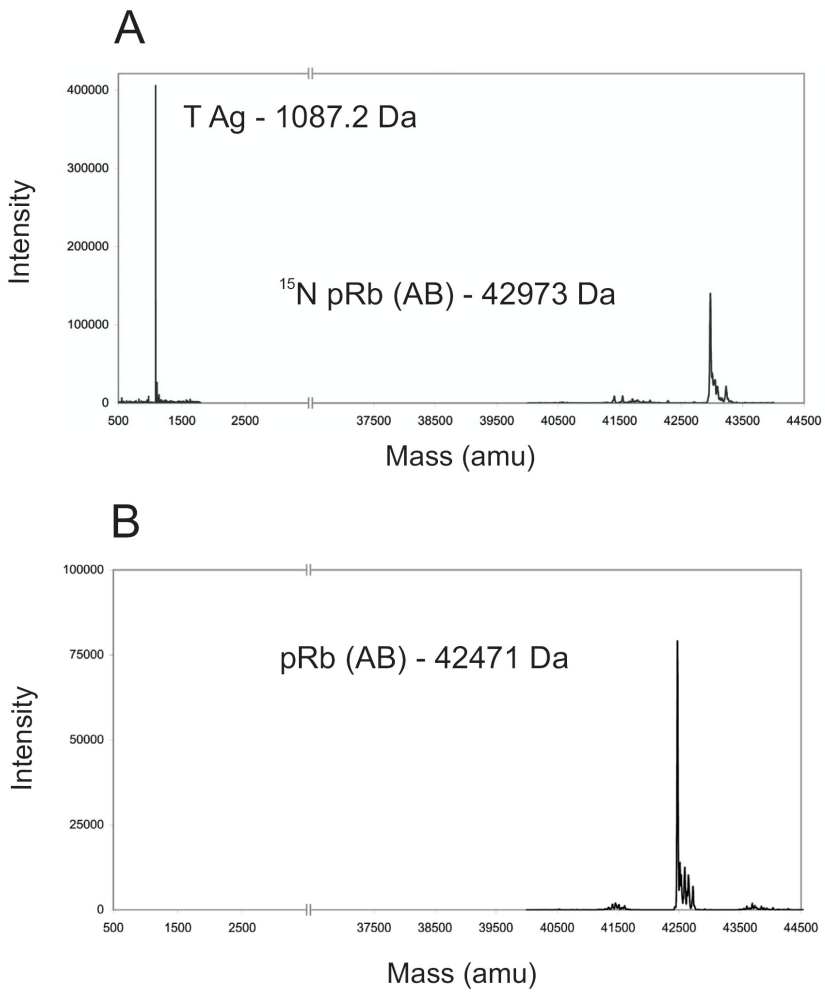


Figure 3.2.3. Mass spectrometry results of: A) the gel filtration chromatography purified sample of mixture 1:1 ^{15}N pRb (AB) with the large T antigen peptide - both the protein and the peptide are detectable, indicating physical interaction between the two. B) The sample of mixture 1:1 pRb (AB) with the MyoD purified by gel filtration chromatography - the detectable protein corresponds to pRb (AB).

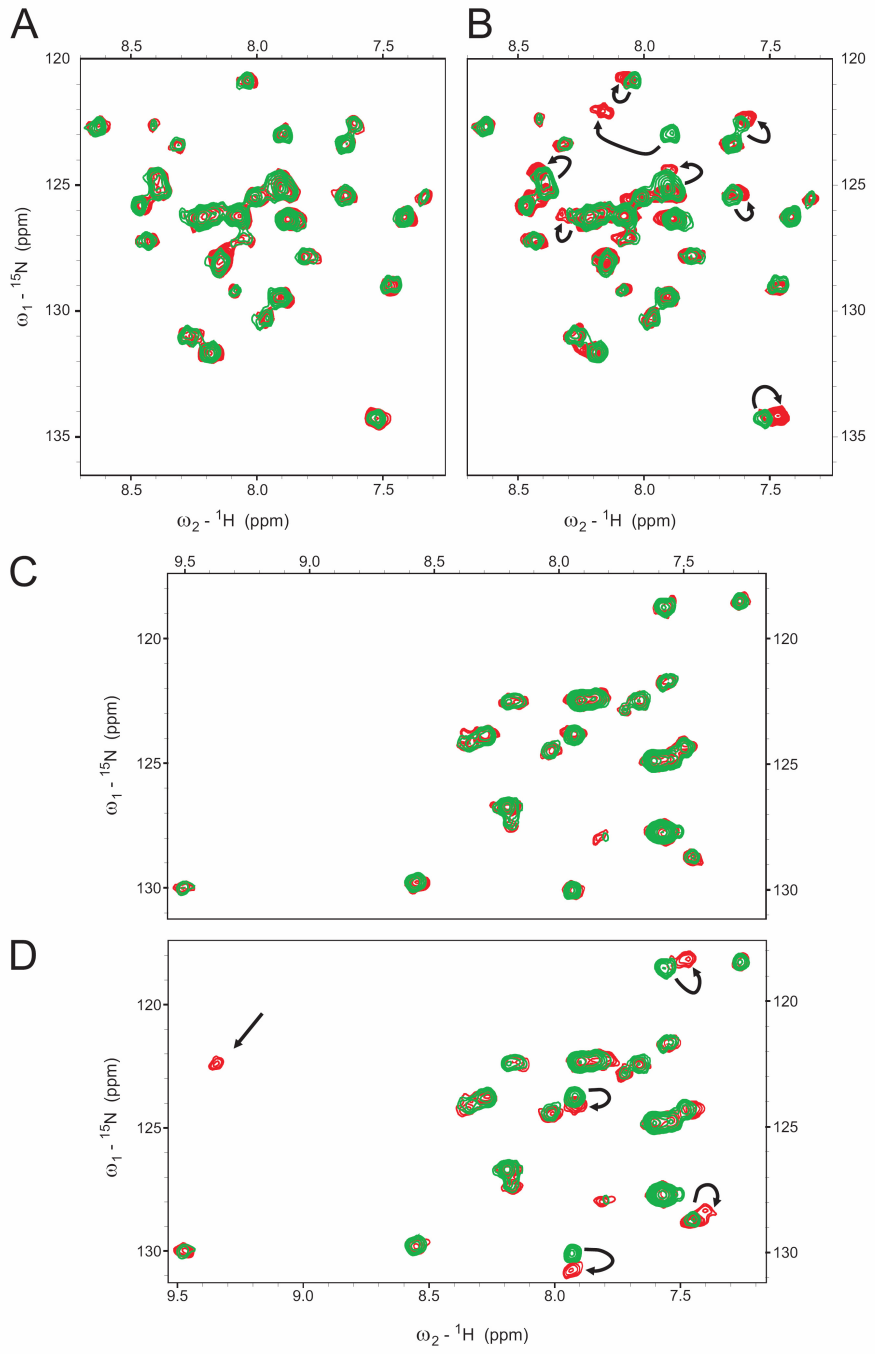


Figure 3.2.4. Titration of the ^{15}N-labeled pRb (AB) with Id-2 and MyoD. The ^{15}N-HSQC spectra of the A) selectively ^{15}N-Leu labeled pRb (construct AB) mixed in equimolar ratio with Id-2 (red), B) ^{15}N-Leu labeled pRb mixed in equimolar ratio with the E7-1 peptide (red), C) ^{15}N-Lys labeled pRb (construct AB) mixed in equimolar ratio with MyoD (red), and D) ^{15}N-Lys labeled pRb mixed in equimolar ratio with E7-0 peptide (red); pRb alone is shown in green in all panels. Arrows show cross-peaks that move. All spectra were recorded at pH 7.8.

It has been reported that MyoD, Id-2, and viral HPV16 E7 bind to the same region of pRb (Gu et al. 1993; Lasorella et al. 1996). Because bigger proteins are hardly accessible by NMR (as the relaxation rates are increasing and spectra become too complicated to interpret), we decided to check interactions only with some specific residues of pRb. The published crystal structure of the complex between the pRb small pocket and the HPV E7 peptide as well as with T large antigen, indicate that there are several lysine and leucine residues of pRb situated in the B domain of pRb that participate in the LXCXE-pRb complex formation (Lee et al. 1998; Dahiya et al. 2000; Kim et al. 2001; Dick and Dyson 2002). To test interactions with these particular amino acids we used selectively ^{15}N labeled pRb-AB with ^{15}N-Lys or ^{15}N-Leu. Titration of ^{15}N-Leu pRb-AB with the unlabeled Id-2 did not cause any changes in the positions of cross peaks of the ^{1}H-^{15}N HSQC spectrum, proving no interaction. The same result was observed in the reaction of the ^{15}N-Lys pRb-AB with the unlabeled MyoD. As a positive control we were able to see changes in the position and intensities of peaks after complex formation of the ^{15}N Lys and Leu labeled pRb with E7-0 and E7-1 peptides (Figure 3.2.4).

3.2.1.8 pRb and L/IXCXE peptides from HPV E7, SV40 large T antigen, HDAC1 and PAI-2

In this part of my work we tested the binding of pRb-AB wt and two its mutants with the LXCXE proteins and peptides. Our ITC experiments showed that dissociation constant (K_D) for binding with the wild type pRb was about 0.07 for E7 LXCXE peptide and about 0.44 ± 0.06 for the large T antigen peptide (Table 3.2.2; Figure 3.2.5). It has been reported that a pRb mutant C706F was defective in its ability to bind to the LXCXE sequence containing viral proteins (Kratzke et al. 1992). Another mutant, K713A, was also shown not to bind LXCXE sequences in one report (Dahiya et al. 2000), whereas other report suggested the opposite (Chan et al. 2001). C706 was proposed to be important in proper folding of the B pocket of pRb whereas K713 makes a direct contact with the backbone of the LXCXE protein (Lee et al. 1998). Our ITC experiments showed that both of these mutants bind tightly to the LXCXE peptide of HPV E7 with K_D 0.21 ± 0.04 μM for the K713A mutant and 0.044 ± 0.02 μM for C706F (Figure 3.2.6).

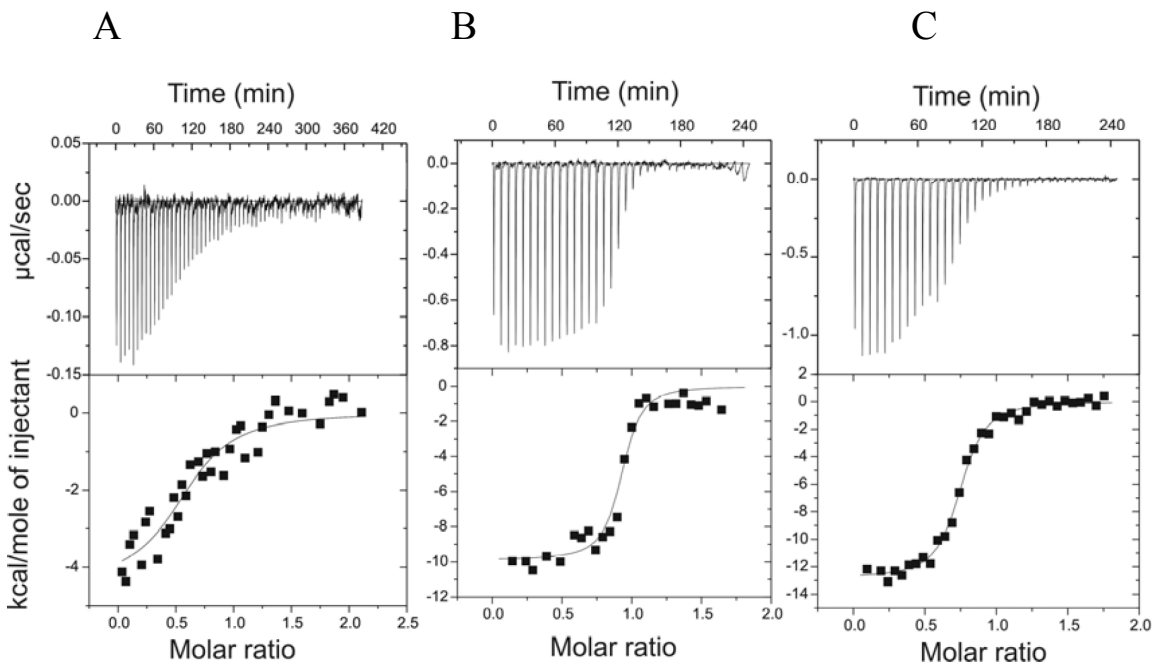


Figure 3.2.5. ITC titration curves for various LXCXE peptides with the pRb-AB. (Upper) The raw data of ITC experiment each performed at 20-22°C. (Lower) The integrated heat changes, corrected for the heat of dilution, and fitted curve based on a single site model. A) pRb-AB titrated with HDAC1 peptide (17mer). B) pRb-AB titrated with HPV E7 peptide. C) pRb-AB titrated with Large T antigen peptide.

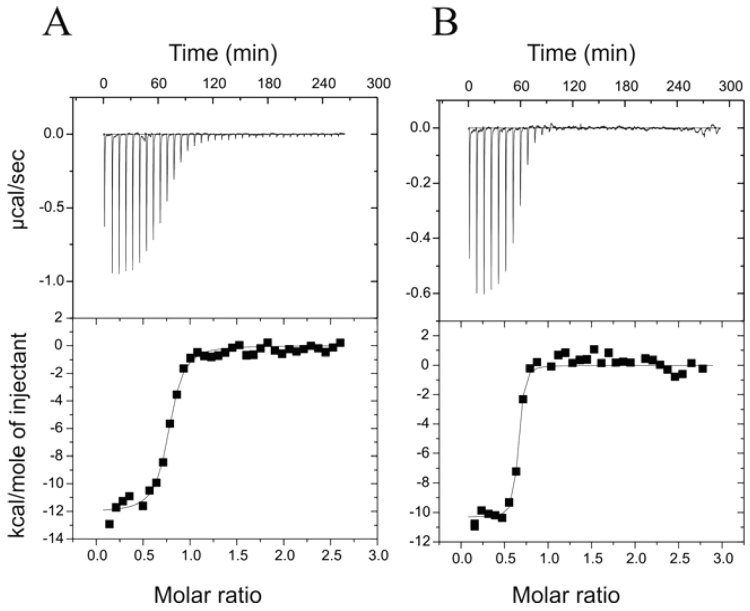


Figure 3.2.6. ITC curves for pRb-AB mutants and E7 peptide. A) pRb-AB (K713A) and E7 peptide. B) pRb-AB (C706F) and E7 peptide.

ITC experiments were additionally supported by gel filtration chromatography. Subsequently, we have used NMR to check the structural integrity of the mutants (Pellecchia et al. 2002; Rehm et al. 2002). The NMR showed that these mutations do not disrupt the structural integrity of the small pocket as no noticeable differences were observed compared to the wild type pRb-AB pocket in the NMR spectra.

A much weaker binding was measured for the interaction of the pocket domain of pRb and the IXCXE peptide of HDAC1 (17mer). In this case the equilibrium dissociation constant K_D was estimated to be about 10 μM (Figure 3.2.5) We tested also two additional HDAC1 peptides: a shorter - 9 amino acids version and a longer - 20mer (Table 3.2.2). ITC showed only weak binding between pRb-AB and all the three peptides, the estimated K_D roughly to be about 10 - 20 ± 3 μM. We also checked the strength of this binding using mass spectrometry combined with gel filtration. The pRb-AB and the three HDAC1 peptides were mixed in 1:3 molar ratios and passed through a S75 analytical Superdex column. A mass corresponding only to pRb-AB was detected in the eluent. In the case of HDAC1 there is isoleucine instead of conserved leucine at the beginning of LXCXE sequence. To check if this difference can have influence on interaction with pRb we performed ITC experiment with E7 mutated peptide - D**I**YCYEQLN. Binding of the peptide to the pRb-AB occurred with approximately the same affinity as wild type E7 peptide (K_D 0.32 μM) showing that this change does not disrupt the binding.

As the next approach to check the interactions, I used NMR spectroscopy. Titration of the ^{15}N-Lys pRb-AB with the HDAC1 peptide exhibited five lysines peak shifts after complex formation (Figure 3.2.7). The positive control experiment constituted the titration of the ^{15}N-Lys pRb-AB with the unlabeled E7 peptide. NMR spectra showed two separate sets of ^{1}H-^{15}N HSQC resonances observed in the intermediate stages of titration, one corresponding to the free pRb-AB and the other to the pRb-AB bound to the E7 peptide. During the titration experiment between the HDAC1 peptide and pRb-AB we noticed a continuous movement of five NMR peaks upon addition of increasing amounts of the peptide on HSQC spectra (Figure 3.2.7 and 3.2.8). This indicates that binding of the peptide to the pRb-AB is in fast exchange, which indicates weak binding that is the opposite of low exchange binding pRb/E7 peptide. We calculated the K_D for the pRb-HDAC1 complex formation from the NMR titrations to be 40 ± 10 μM at 27°C, which agrees with an approximate value estimated from ITC (10 ± 3 μM at 20°C). For the smaller version of the HDAC1 peptide, the changes in the chemical shift values were similar but smaller.

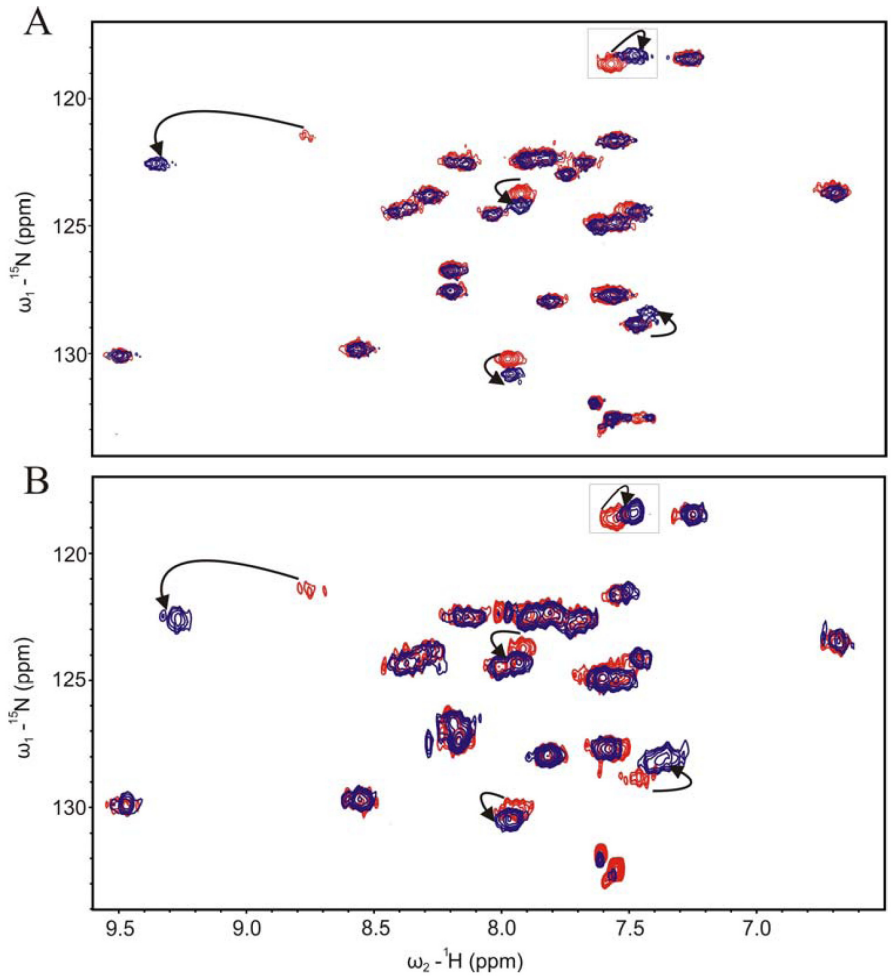


Figure 3.2.7. The ^{15}N lysine labeled pRb-AB was titrated with increasing concentrations of the E7 peptide (panel A) and ^{1}H-^{15}N HSQC spectra were recorded at each step of ligand addition. Only two steps of titrations are shown: the starting (red, untitrated) and the final step (blue, at an E7 peptide/pRb-AB molar ratio of 1.60). Upon addition of E7 peptide five resonances were perturbed (arrows), indicating that these lysines are involved in the binding to pRb-AB. In panel B a fresh sample of the ^{15}N lysine labeled pRb-AB was titrated with the IXCXE peptide from HDAC1. Only two steps of titrations are shown, red - the starting (untitrated) and blue - the final (at HDAC1 peptide/pRb-AB molar ratio of 3.76). Features of the cross peak (marked box, at 118.75 ppm, ω_1-^{15}N; 7.56 ppm, ω_2-^{15}H chemical shift values) in A) and B) are described in detail in Figure 3.2.8.

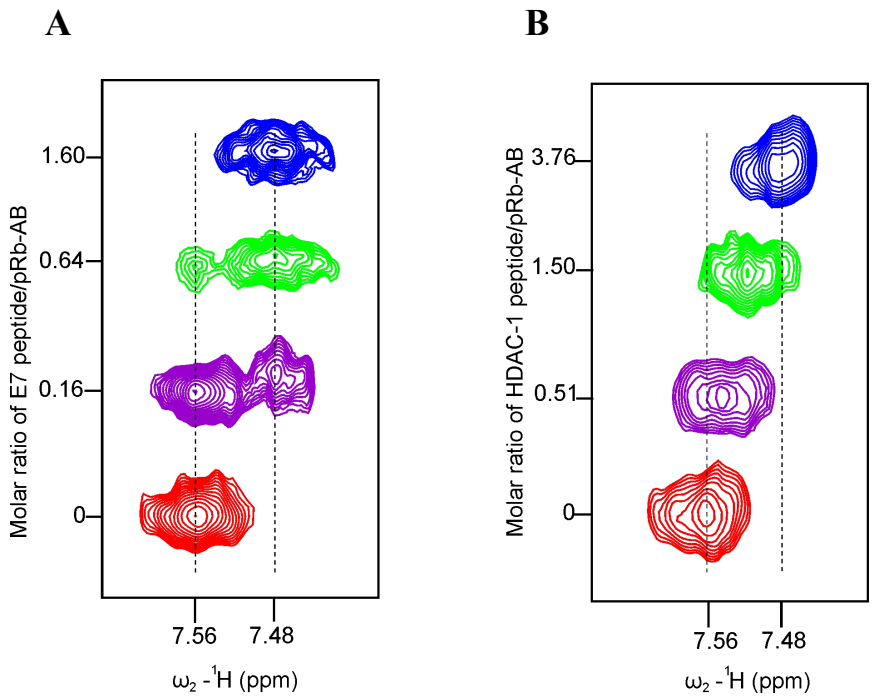


Figure 3.2.8. The cross peak (at 118.75 ppm ω1-15N; 7.56 ppm ω2-1H, Figure 3.2.7) was chosen to demonstrate the difference between tight A) (E7 peptide/pRb-AB) and weak B) (HDAC1 peptide/pRb- AB) binding. Four steps of titrations are shown here at the specified peptide/pRb-AB molar ratio.

It has been reported that the plasminogen activator inhibitor 2 (PAI-2) protects the pRb protein before degradation in the nucleus were both proteins colocalize (Darnell et al. 2003). Because PAI-2 has a LXCXE motif, I performed binding tests between several PAI-2 and pRb constructs. ITC and NMR experiments showed no binding between a 17 amino acids peptide of PAI-2 and the Rb-AB protein. I also proved the lack of the interaction of the PAI-2 wt and PAI-2 ΔCD with the pocket domain of pRb (Rb-AB) using gel filtration chromatography. Additionally to the LXCXE sequence PAI-2 wt possess a PENF motif which is localized in the CD interhelical loop of PAI-2 wt, lacking in the ΔCD mutant. Some data suggest that PAI-2 binds to the C-pocket of pRb through this motif (Darnell et al. 2003). To check this LXCXE independent interaction I performed additionally a gel filtration analysis of binding between PAI-2 wt and Rb-ABC as well as C-Rb. As a negative control, I checked the ΔCD mutant binding to C-Rb. In all cases I saw no binding between these proteins (Figure 3.2.9).

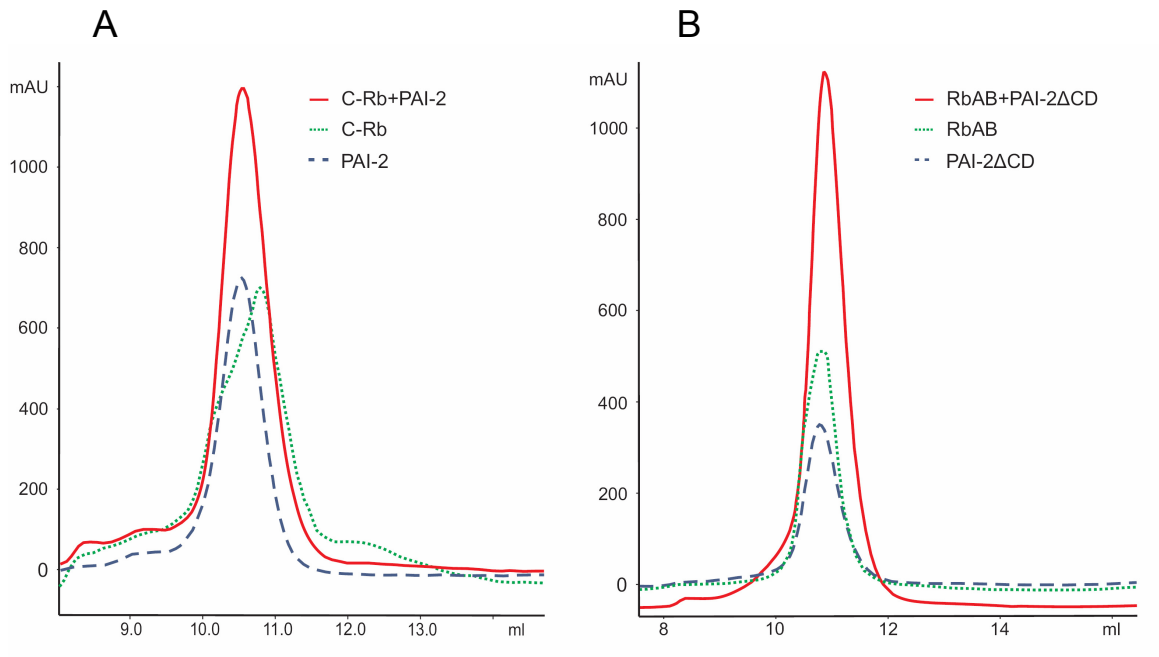


Figure 3.2.9. Monitoring of the protein–protein complexes formation using gel filtration chromatography. Both panels show no complex formation between C-Rb (dimer)-dotted line, and PAI-2 wt – dash line (panel A) and pRb-AB - dotted line and PAI-2 ΔCD – dash line (panel B). Straight line shows elution profile of 1:1 molar mixture of pRb and PAI-2 constructs.

3.2.2 Discussion

3.2.2.1 The HLH proteins MyoD and Id-2 does not interact directly with pRb

MyoD is a transcription factor belonging to the HLH protein family (Puri et al. 1997; Cenciarelli et al. 1999; Polesskaya et al. 2001; Magenta et al. 2003) that plays a key role in the differentiation of all skeletal muscle lineages (Weintraub et al. 1991). *In vitro* protein binding and immunoprecipitation studies indicated that both MyoD and myogenin bind to the pocket domain of underphosphorylated pRb directly through the bHLH domain (Gu et al. 1993). Several other *in vivo* studies however could find no evidence for such binary interactions suggesting that if there is any direct interaction it appears to be weak, making the significance of the *in vitro* MyoD-pRb interaction questionable (Puri et al. 1997; Cenciarelli et al. 1999; Zhang et al. 1999; Zhang et al. 1999; Guo et al. 2003). Using the two-hybrid system Zhang et al. (1999a) could not detect direct MyoD or myogenin-pRb interactions.

Id-1 to Id-4 are another category of mammalian HLH proteins (Pagliuca et al. 1995). These proteins promote cell proliferation by binding to transcription factors from

the bHLH subclass of proteins and inhibit their ability to bind DNA (Norton et al. 1998). Id-2 and MyoD have a high degree of sequence homology. Only Id-2, and not other Id proteins, was reported to disrupt the antiproliferative effects of pRb proteins via direct interactions (Lasorella et al. 1996). The HLH domain of Id-2 was shown to interact with pRb (the small pocket domain), p107, and p130 (Iavarone et al. 1994; Lasorella et al. 1996).

Our approach was to verify this data by carrying out direct protein-protein interactions *in vitro*. By using purified constructs, situations where proteins interact indirectly by multiprotein complexes can be easily eliminated making interpretation of the data clearer. In our research we used NMR spectroscopy, gel filtration chromatography combined with mass spectrometry and affinity chromatography pull-down assays (Diercks et al. 2001; Pellecchia et al. 2002; Meyer and Peters 2003). We did not use the yeast two hybrid system, immunoprecipitation and autoradiographic detection methods as it is known that these methods are prone to give false positive results (Gu et al. 1993; Lasorella et al. 1996; Gietz et al. 1997; Toby and Golemis 2001).

We were unable to find direct interactions between pRb and the HLH proteins MyoD and Id-2. However we could detect interactions between MyoD and Id-2, MyoD and DNA, as well as between pRb and viral oncoproteins proving that methods used in our experiments are effective in testing *in vitro* protein-protein interactions. It is possible that Id-2 and MyoD influences directly or indirectly the expression levels of the pRb gene without the physical protein–protein interaction with pRb. MyoD is a transcription factor that can interact with p300 and CBP proteins of which both are histone acetyltransferases (Halevy et al. 1995). These interactions facilitate binding to a promoter sequence CRE. MyoD together with CBP and CREB proteins trigger transcriptions of pRb (Magenta et al. 2003) and p21 (Halevy et al. 1995; Cenciarelli et al. 1999) - both promoters of cell differentiation. It was also shown that augmented levels of p57, p27, and probably p21 cause increased stability of MyoD (Zabludoff et al. 1998; Reynaud et al. 1999).

It has been shown that the bHLH protein E2A has the ability to block G1 to S phase transition at least through the transcriptional activation of the p21Cip1 gene. Id proteins are able to inhibit E2A mediated expression of the p21Cip1 gene (Prabhu et al. 1997). The p21Cip1 binds to and inactivates CDKs and enhances the activity of pRb, thereby blocking G1 to S phase transition. Therefore, it is possible that Id proteins mediate G1 to S phase transition through the downregulation of p21Cip1 gene expression by heterodimerization with E2A.

3.2.2.2 Role of the complex formation between pRb and LXCXE proteins and peptides

The retinoblastoma gene product has been shown to restrict cell proliferation, promote cell differentiation, and inhibit apoptosis. The small pocket of pRb (residues 379-772) binds to the LXCXE-like sequence containing proteins (Lee and Cho 2002) and includes also a primary binding site for E2Fs (Flemington et al. 1993; Mundle and Saberwal 2003; Xiao et al. 2003). The C pocket region, that together with A and B domains form large pocket of pRb (amino acids 379-928), is responsible for binding to the nuclear c-Abl tyrosine kinase. It has also an additional binding site for E2F, required for its physiological activity, and is necessary for binding to other cellular proteins like MDM2 (Xiao et al. 1995) and cdk4/D1 complex (Pan et al. 2001). There are several viral proteins, including E7 from the human papilloma virus (HPV), and the T large antigen from simian virus 40 (SV40), E1a from adenovirus, and the Tax oncoprotein from the human T-cell leukemia virus type-I (HTLV-I) connected with pRb inactivation. All of these oncoproteins bind to the small pocket of pRb through the LXCXE motif. Cellular proteins, such as HDAC1, HDAC2, cyclin D1, BRCA1 and PAI2, possess an LXCXE-like motif in their sequences but for some of them it is not essential for interaction with pRb. (Ewen et al. 1993; Dunaief et al. 1994; Brehm et al. 1998; Luo et al. 1998; Magnaghi-Jaulin et al. 1998; Adams et al. 1999; Fan et al. 2001; Darnell et al. 2003). Despite of the number of studies there are still contradictory evidences concerning interaction between pRb and LXCXE proteins. In some reports HDAC1 binds directly to the repressive, hypophosphorylated form of pRb (Luo et al. 1998) by the same motif as viral oncoproteins (Brehm et al. 1998; Magnaghi-Jaulin et al. 1998), whereas others suggest that HDAC1 and HDAC2 do not utilize their LXCXE-like sequences for interaction with the small pocket of pRb (Chen and Wang 2000; Kennedy et al. 2001; Dick and Dyson 2002); instead, recruitment of HDACs to pRb occurs via an intermediary, LXCXE containing protein, RBP-1 (Dunaief et al., 1994).

Our data suggest that HDAC1 binds to the small pocket of pRb by the LXCXE-like motif but interactions mediated by the same motif between pRb and viral proteins E7 and T large antigen are much stronger. The HDAC1 protein has isoleucine instead of a conserved leucine in the LXCXE motif. To check if this mutation has influence on the binding strength, we performed additional tests with E7 mutated peptide – D**I**YCYEQLN were leucine was replaced by isoleucine. The mutated peptide binds with almost the same strength to the pRb pocket suggesting importance of additional residues of LXCXE

proteins in binding to pRb. Moreover, we were unable to notice any binding between pRb and the conserved LXCXE-like sequence from PAI-2. To explain why proteins possessing the LXCXE sequence can bind the small pocket domain of pRb (pRb-AB) with different kinetics or do not bind at all we have to consider structural elements involved in the binding, as well as the functions of the proteins that interact with pRb.

A structural data of the interactions of the LXCXE peptides from E7 and T antigen with the pocket domain of pRb indicate that a hydrophobic groove in the pRb pocket domain B forms the binding site. Few conserved amino acids, like for example, Tyr 709, Lys 713, Tyr 756, and Asn 757 of pRb make direct contacts with the backbone of the LXCXE peptide, some others like Ala 562, Ser 567, Arg 661 and Cys 706 are responsible for stabilizing the folded state of the A-B domain (Kaye et al. 1990; Lee et al. 1998; Kim et al. 2001). Our results proved that even if some of these residues are mutated, binding with oncoproteins is still preserved. It is probable that the evolutionary created ability to bind pRb is much stronger in the case of viral proteins because they must compete with cellular components, such as HDAC1 or E2F, for the pRb interaction site. E2F binds to the interface of the A and B domains making interactions with conserved residues from both. It was reported that the binding of E2F peptides to the small pocket region of pRb is independent of the binding of the E7 peptide based on the LXCXE motif, but a full length E7 protein abolishes the E2F binding capability to pRb (Xiao et al. 2003).

The LXCXE motif of T large antigen is a part of the longer unstructured sequence. The E7 9 amino acids peptide used in structural studies also does not form any secondary structure by itself but is difficult to speculate how it behaves as a part of a bigger domain. Both of them bind tightly in an extended conformation to the B pocket of pRb making hydrophobic contacts and hydrogen bonds with amino acids of the groove. (Lee et al. 1998; Kim et al. 2001). Secondary structure predictions of the HDAC1 protein reveal that its LXCXE-like motif is also a part of long unstructured chain. This can suggest an analogous mechanism of the binding with pRb. The weaker affinity of HDAC1 peptides to the LXCXE binding site of pRb can be important for the involvement of both proteins in multiprotein pathways that require high dynamics of interactions and high degree of precision. Proteins that play critical roles in the cell cycle regulation must be quickly released from the complex whereas viral proteins with only one possible target can act with higher affinity.

The crystal structure of PAI-2 reveals that its LXCXE motif is located on the sharp kink between a β-sheet and α–helix. This structural motif can prevent a close contact

formation between the B domain of pRb and PAI-2. Leu, Cys and Glu residues of PAI-2 protein could not form a similar pattern of interactions with pRb residues as seen in corresponding amino acids of viral oncoproteins, which touch directly the B-box groove. Moreover the LXCXE sequence of PAI_2 is exposed to the solvent as a part of hydrophilic surface that can reduce capacity of the LXCXE motif to interact with the hydrophobic groove of the B pocket.

The plasminogen activator inhibitor-2 (PAI-2) was proposed to be involved in the inhibition of the retinoblastoma protein degradation and enhancing its E2F repression activity by direct binding to the pRb (Darnell et al. 2003). *In vitro* protein studies, coimmunoprecipitation, and autoradiography data indicate that the PAI-2 protein binds to the pocket domain of pRb through the LXCXE motif and to the C pocket of pRb through the C-D interhelical region carrying the PENF motif. The PENF and LXCXE homology motifs are shared by many cellular and viral proteins that bind pRb. Some of them, such as BRCA1, bind to pRb using a LXCXE independent mechanism (Fan et al. 2001). There are also a number of proteins containing the PENF homology motif but not LXCXE, such as HDAC3 (Dahiya et al. 2000) or c-Abl, that are able to bind to the C-pocket of pRb (Welch and Wang 1993). However in all these cases there are no strong evidences on the PENF mediated binding to pRb. In our studies we did not notice any binding between PAI-2 and pRb using variety of constructs and methods, which may indicate on an alternative pRb - PAI-2 interaction pathway.

It was reported that PAI-2 synthesis is rapidly induced in response to the inflammatory cytokine, tumor necrosis factor-alpha (TNFα) in monocytes, macrophages and many tumors (Pytel et al. 1990; Darnell et al. 2003). PAI-2 is able to protect HeLa tumor cells from the TNF induced apoptosis. The data demonstrate that the C-D interhelical domain, as well as the Arg380 in P1 position of reactive site loop (RSL), are required for this protection, because cells expressing PAI-2 R380A and PAI-2 ΔCD mutants were still sensitive to the TNF-induced cell death (Dickinson et al. 1995; Dickinson et al. 1998). These data suggest that only an active form of PAI-2, possessing the protease activity, can act as an apoptosis inhibitor. It was postulated that PAI-2 might be an important factor in regulating cell death through the inhibition of a proteinase involved in the TNF-induced apoptosis (Dickinson et al. 1995). PAI-2 R380A and PAI-2 ΔCD mutants failed to protect pRb from degradation, which also suggests the connection between the PAI-2 proteinase inhibition activity and the restoration of pRb. PAI-2 may inhibit proteinases that are responsible for the pRb degradation, without direct binding to

pRb, through a similar mechanism to that which is responsible for the apoptosis inhibition induced by TNFα. Moreover, TNF indirectly act as a pRb activator which promotes the cell-cycle arrest and apoptosis inhibition (Shen et al. 2004). This can explain the elevation of pRb protein levels in the presence of PAI-2 suggested by Darnell et al. (2003).

Other serpin that has been described as a tumor suppressor protein was maspin that also contains Arg in the P_1 position of RSL (Zou et al. 1994). Both proteins, PAI-2 and maspin, were postulated to be responsible for blocking the cancer cell motility which was connected with their protease inhibition activities (Sheng et al. 1996).

4 SUMMARY

The thesis presented here is a result of the studies carried out in the Department of Structural Research at the Max Planck Institute for Biochemistry in Martinsried. The work has been based on two related projects. The first project is focused on the X-ray structure determination of the dimerization domain of the fibroblast growth factor receptor 1 (FGFR1) oncogene partner (FOP). The second project concerned the biochemical studies of the retinoblastoma protein (pRb) and its interaction with the proteins involved in the regulation of the cell cycle and cell differentiation.

The FOP protein was originally discovered as a fusion partner for the FGF-receptor 1 in chromosomal translocations giving rise to the myeloproliferative disorder (MPD). Dimerization abilities of the N-terminal part of FOP were postulated to be responsible for the FGFR1 induced, uncontrolled cells proliferation resulting in MPD. A subsequent proteomics screen identified FOP as a component of the centrosome. Targeting of the dimerized FOP-FGFR1 fusion protein to the centrosome contributes to oncogenesis. The presented 1.6 Å resolution structure shows the N-terminal domain of the FOP protein, containing a LisH-like motif that is part of a dimerization domain involved in the regulation of microtubule dynamics (by binding to dynein or to microtubules directly). FOP is also necessary for proteins like CAP350 (centrosome associated protein 350) and EB1 (plus-end binding protein) to associate with microtubules. The data presented in this thesis suggest that the LisH domain is important but not sufficient for the centrosomal localization, dimerization and multi-protein complex formation. Therefore I propose the LisH-like motif to be a part of a larger structural and functional domain.

The pRb was the first tumour suppressor protein identified through human genetic studies. It plays a central role in the regulation of cell cycle progression, by regulation of the E2F family of transcription factors that control the expression of genes involved in cell proliferation and survival. The pRb is also a target for oncoproteins of DNA tumour viruses. A large group of cellular proteins has been reported to bind the pocket domains of pRb directly. Because some of the data about the interactions between these proteins have been inconsistent, I investigated the binding between pRb and its interaction partners (MyoD, Id-2, HDAC1, PAI-2, HPV E7, and the SV40 large T antigen) using *in vitro* methods like NMR spectroscopy, mass spectrometry, isothermal titration calorimetry (ITC), the affinity chromatography pull-down assays, and gel filtration chromatography.

We found that there are no direct protein-protein interactions *in vitro* between pRb and myogenic differentiation promoting (MyoD) and inhibiting (Id-2) factors. Therefore, an indirect mechanism of the pRb regulation was proposed. It was suggested that MyoD and Id-2 could influence the expression level of the pRb gene without physical interaction with pRb, by direct binding to its promoter or through additional binding partners. Several viral oncoproteins (*e.g.* HPV E7 and SV40 large T antigen) and about 20 cellular proteins are known to bind to the pRb small pocket domain via a LXCXE motif. In this thesis we found that kinetics of binding through this motif differs between proteins and is dependent on the secondary structure of the LXCXE motif and also on the flanking residues.

5 ZUSAMMENFASSUNG

Die hier vorgestellte Arbeit ist das Ergebnis von Forschungsstudien in der Abteilung Strukturforschung am Max Planck Institut für Biochemie, Martinsried. Die Arbeit unterteilt sich in zwei Blöcke. Das erste Projekt befasst sich mit der Lösung der Röntgenstruktur der Dimerisierungsdomäne des Fibroblast Wachstumsfaktor Rezeptor 1 (FGFR1) Oncogen Partner (FOP). Das Zweite behandelt biochemische Untersuchungen am Retinoblastoma Protein (pRb) und dessen Interaktion mit Proteinen der Zellzyklus Regulation und der Zelldifferenzierung.

Das FOP Protein wurde ursprünglich als ein Fusionspartner für den FGF-Rezeptor 1 in chromosonalen Translokationen gefunden, die für die *myeloproliferative disorder* (MPD) verantwortlich ist. Die Dimerisierungsfähigkeit der N-terminalen Domäne von FOP wurde als Verursacher der FGFR1 induzierten unkontrollierten Zellproliferation angenommen. Ein weiterführender Proteomic Screen identifizierte FOP als eine Komponente des Centrosoms. Die Lokalisierung des dimeren FOP-FGFR1 Fusionsproteins im Centrosom mitverursacht dessen oncogene Eigenschaften. Die hier vorgestellte 1,6 Å Auflösung Struktur zeigt den N-terminalen Teil des FOP Proteins, und enthält ein LisH ähnliches Motiv, welches eine Dimerisierungsdomäne enthält und in der Regulation der Dynamik der Micotubuli (durch Bindung an Dynein oder direkt an Microtubuli) beteiligt ist. FOP wird für die Proteine CAP350 (*centrosome associated protein 350*) und EB1 (*plus-end binding protein*) benötigt um mit den Microtubuli zu assoziieren. Die hier vorgestellten Daten implizieren, dass das LisH-ähnliche Motiv wichtig, aber nicht ausreichend für die centrosomale Lokalisation, Dimerisation und Multiprotein Komplex Formation ist. Daher nehme ich an, dass das LisH-ähnliche Motiv ein Teil einer größeren strukturellen und funktionalen Domäne ist.

Das Retinoblastoma Protein war das erste Tumorsupressor Protein, das in Humangenetik Studien gefunden wurde. Es hat eine zentrale Rollen in der Regulation des Voranscheitens des Zellzyklus durch die Kontrolle der Familie der E2F Transkriptionsfaktoren, die ihrerseits für die Expression von Proteinen für die DNA Replikation und Zellproliferation verantwortlich sind. Das pRb Protein ist Ziel vieler viraler Oncoproteine von DNA Tumorviren. Es wurde von einer großen Gruppe von zellularen Proteinen berichtet, dass sie an die Pocket Domäne von pRb direkt binden. Da einige dieser Daten über die Interaktion mit pRb widersprüchlich sind haben wir einige der

pRb Bindungspartner (MyoD, Id-2, HDAC1, PAI-2, HPV E7 und SV40 large T antigen) näher untersucht. Dazu wurden verschiedene *in vitro* Methoden wie NMR Spektroskopie, Massenspektrometrie, Isothermal Titrations Calorimetrie (ITC), Affinitäs Chromatographie pull-down Experimente und Gelfiltrations-Chromatographie verwendet.

Es konnte keine direkte Protein-Protein Interaktion in vitro zwischen pRb und MyoD (*myogenic differentiation promoting*) und Id-2 (*inhibiting factor*) gefunden werden. Daher wird ein indirekter Mechanismus der pRb Regulation angenommen und MyoD und Id-2 beeinflussen die Expression des pRb Gens ohne direkt mit pRb zu interagieren, indem sie an dessen Promoter binden oder weiter Bindungspartner rekrutieren. Von vielen virale Oncoproteine (z.B., HPV E7 und SV40 large T antigen) und ungefähr 20 zellulare Proteine ist bekannt, dass sie an die pRb Pocket Domäne über ein LXCXE Motiv binden. In dieser Arbeit haben wir herausgefunden, dass die Kinetik der Bindung über dieses Motiv sich sehr stark zwischen den Proteinen variiert und ebenfalls die Sekundärstruktur und die flankierenden Reste des Motivs zu dessen Bindungsstärke beitragen.

6 REFERENCES

Adams, P.D., Gopal, K., Grosse-Kunstleve, R.W., et al. (2004). "Recent developments in the PHENIX software for automated crystallographic structure determination." *J Synchrotron Radiat* **11**: 53-55.

Adams, P.D., Li, X., Sellers, W.R., et al. (1999). "Retinoblastoma protein contains a C-terminal motif that targets it for phosphorylation by cyclin-cdk complexes." *Mol Cell Biol* **19**: 1068-1080.

Aguiar, R.C., Macdonald, D., Mason, P.J., et al. (1995). "Myeloproliferative disorder associated with 8p11 translocations." *Blood* **86**: 834-835.

Andersen, J.S., Wilkinson, C.J., Mayor, T., et al. (2003). "Proteomic characterization of the human centrosome by protein correlation profiling." *Nature* **426**: 570-574.

Antalis, T.M., La Linn, M., Donnan, K., et al. (1998). "The serine proteinase inhibitor (serpin) plasminogen activation inhibitor type 2 protects against viral cytopathic effects by constitutive interferon alpha/beta priming." *J Exp Med* **187**: 1799-1811.

Arlot-Bonnemains, Y. and Prigent, C. (2002). "Cell cycle. A trigger for centrosome duplication." *Science* **295**: 455-456.

Arnold, H.H. and Braun, T. (1996). "Targeted inactivation of myogenic factor genes reveals their role during mouse myogenesis: a review." *Int J Dev Biol* **40**: 345-353.

Arnold, H.H. and Winter, B. (1998). "Muscle differentiation: more complexity to the network of myogenic regulators." *Curr Opin Genet Dev* **8**: 539-544.

Bassi, M.T., Ramesar, R.S., Caciotti, B., et al. (1999). "X-linked late-onset sensorineural deafness caused by a deletion involving OA1 and a novel gene containing WD-40 repeats." *Am J Hum Genet* **64**: 1604-1616.

Bengal, E., Flores, O., Rangarajan, P.N., et al. (1994). "Positive control mutations in the MyoD basic region fail to show cooperative DNA binding and transcriptional activation in vitro." *Proc Natl Acad Sci U S A* **91**: 6221-6225.

Bertrand, N., Castro, D.S. and Guillemot, F. (2002). "Proneural genes and the specification of neural cell types." *Nat Rev Neurosci* **3**: 517-530.

Bettencourt-Dias, M., Rodrigues-Martins, A., Carpenter, L., et al. (2005). "SAK/PLK4 is required for centriole duplication and flagella development." *Curr Biol* **15**: 2199-2207.

Blattner, C., Sparks, A. and Lane, D. (1999). "Transcription factor E2F-1 is upregulated in response to DNA damage in a manner analogous to that of p53." *Mol Cell Biol* **19**: 3704-3713.

Blow, D. (2002). *Outline of Crystallography for Biologist*. New York, Oxford University Press.

Borger, D.R. and DeCaprio, J.A. (2006). "Targeting of p300/CREB binding protein coactivators by simian virus 40 is mediated through p53." *J Virol* **80**: 4292-4303.

Bradford, M.M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Anal Biochem* **72**: 248-254.

Brehm, A., Miska, E.A., McCance, D.J., et al. (1998). "Retinoblastoma protein recruits histone deacetylase to repress transcription." *Nature* **391**: 597-601.

Bricogne, G., Vonrhein, C., Flensburg, C., et al. (2003). "Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0." *Acta Crystallogr D Biol Crystallogr* **59**: 2023-2030.

Browaeys-Poly, E., Fafeur, V., Vilain, J.P., et al. (2005). "ERK2 is required for FGF1-induced JNK1 phosphorylation in Xenopus oocyte expressing FGF receptor 1." *Biochim Biophys Acta* **1743**: 1-4.

Camilleri, C., Azimzadeh, J., Pastuglia, M., et al. (2002). "The Arabidopsis TONNEAU2 gene encodes a putative novel protein phosphatase 2A regulatory subunit essential for the control of the cortical cytoskeleton." *Plant Cell* **14**: 833-845.

Caracciolo, V., Reiss, K., Khalili, K., et al. (2006). "Role of the interaction between large T antigen and Rb family members in the oncogenicity of JC virus." *Oncogene* **25**: 5294-5301.

Cardone, M.H., Roy, N., Stennicke, H.R., et al. (1998). "Regulation of cell death protease caspase-9 by phosphorylation." *Science* **282**: 1318-1321.

CCP4 (1994). "The CCP4 suite: programs for protein crystallography." *Acta Crystallogr D Biol Crystallogr* **50**: 760-763.

Cenciarelli, C., De Santa, F., Puri, P.L., et al. (1999). "Critical role played by cyclin D3 in the MyoD-mediated arrest of cell cycle during myoblast differentiation." *Mol Cell Biol* **19**: 5203-5217.

Chan, H.M., Smith, L. and La Thangue, N.B. (2001). "Role of LXCXE motif-dependent interactions in the activity of the retinoblastoma protein." *Oncogene* **20**: 6152-6163.

Chang, F., Lee, J.T., Navolanic, P.M., et al. (2003). "Involvement of PI3K/Akt pathway in cell cycle progression, apoptosis, and neoplastic transformation: a target for cancer chemotherapy." *Leukemia* **17**: 590-603.

Chen, T.T. and Wang, J.Y. (2000). "Establishment of irreversible growth arrest in myogenic differentiation requires the RB LXCXE-binding function." *Mol Cell Biol* **20**: 5571-5580.

Cheung, W.L., Turner, F.B., Krishnamoorthy, T., et al. (2005). "Phosphorylation of histone H4 serine 1 during DNA damage requires casein kinase II in S. cerevisiae." *Curr Biol* **15**: 656-660.

Chien, C.T., Bartel, P.L., Sternglanz, R., et al. (1991). "The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest." *Proc Natl Acad Sci U S A* **88**: 9578-9582.

Cobrinik, D. (2005). "Pocket proteins and cell cycle control." *Oncogene* **24**: 2796-2809.

Cross, M.J., Lu, L., Magnusson, P., et al. (2002). "The Shb adaptor protein binds to tyrosine 766 in the FGFR-1 and regulates the Ras/MEK/MAPK pathway via FRS2 phosphorylation in endothelial cells." *Mol Biol Cell* **13**: 2881-2893.

Dahiya, A., Gavin, M.R., Luo, R.X., et al. (2000). "Role of the LXCXE binding site in Rb function." *Mol Cell Biol* **20**: 6799-6805.

Dammermann, A., Desai, A. and Oegema, K. (2003). "The minus end in sight." *Curr Biol* **13**: R614-624.

Darnell, G.A., Antalis, T.M., Johnstone, R.W., et al. (2003). "Inhibition of retinoblastoma protein degradation by interaction with the serpin plasminogen activator inhibitor 2 via a novel consensus motif." *Mol Cell Biol* **23**: 6520-6532.

Davis, R.L., Weintraub, H. and Lassar, A.B. (1987). "Expression of a single transfected cDNA converts fibroblasts to myoblasts." *Cell* **51**: 987-1000.

DeCaprio, J.A., Ludlow, J.W., Figge, J., et al. (1988). "SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene." *Cell* **54**: 275-283.

Delaval, B., Lelievre, H. and Birnbaum, D. (2005). "Myeloproliferative disorders: the centrosome connection." *Leukemia* **19**: 1739-1744 (a).

Delaval, B., Letard, S., Lelievre, H., et al. (2005). "Oncogenic tyrosine kinase of malignant hemopathy targets the centrosome." *Cancer Res* **65**: 7231-7240 (b).

Demiroglu, A., Steer, E.J., Heath, C., et al. (2001). "The t(8;22) in chronic myeloid leukemia fuses BCR to FGFR1: transforming activity and specific inhibition of FGFR1 fusion proteins." *Blood* **98**: 3778-3783.

Dick, F.A. and Dyson, N.J. (2002). "Three regions of the pRB pocket domain affect its inactivation by human papillomavirus E7 proteins." *J Virol* **76**: 6224-6234.

Dickinson, J.L., Bates, E.J., Ferrante, A., et al. (1995). "Plasminogen activator inhibitor type 2 inhibits tumor necrosis factor alpha-induced apoptosis. Evidence for an alternate biological function." *J Biol Chem* **270**: 27894-27904.

Dickinson, J.L., Norris, B.J., Jensen, P.H., et al. (1998). "The C-D interhelical domain of the serpin plasminogen activator inhibitor-type 2 is required for protection from TNF-alpha induced apoptosis." *Cell Death Differ* **5**: 163-171.

Diercks, T., Coles, M. and Kessler, H. (2001). "Applications of NMR in drug discovery." *Curr Opin Chem Biol* **5**: 285-291.

Dixon, J., Ellis, I., Bottani, A., et al. (2004). "Identification of mutations in TCOF1: use of molecular analysis in the pre- and postnatal diagnosis of Treacher Collins syndrome." *Am J Med Genet A* **127**: 244-248.

Donaldson, M.M., Tavares, A.A., Hagan, I.M., et al. (2001). "The mitotic roles of Polo-like kinase." *J Cell Sci* **114**: 2357-2358.

Dormann, D. and Weijer, C.J. (2006). "Imaging of cell migration." *Embo J* **25**: 3480-3493.

Dowdy, S.F., Hinds, P.W., Louie, K., et al. (1993). "Physical interaction of the retinoblastoma protein with human D cyclins." *Cell* **73**: 499-511.

Duensing, S. and Munger, K. (2003). "Human papillomavirus type 16 E7 oncoprotein can induce abnormal centrosome duplication through a mechanism independent of inactivation of retinoblastoma protein family members." *J Virol* **77**: 12331-12335.

Dunaief, J.L., Strober, B.E., Guha, S., et al. (1994). "The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest." *Cell* **79**: 119-130.

Ekholm, S.V. and Reed, S.I. (2000). "Regulation of G(1) cyclin-dependent kinases in the mammalian cell cycle." *Curr Opin Cell Biol* **12**: 676-684.

Emes, R.D. and Ponting, C.P. (2001). "A new sequence motif linking lissencephaly, Treacher Collins and oral-facial-digital type 1 syndromes, microtubule dynamics and cell migration." *Hum Mol Genet* **10**: 2813-2820.

Engel, I. and Murre, C. (2001). "The function of E- and Id proteins in lymphocyte development." *Nat Rev Immunol* **1**: 193-199.

Enzenauer, C., Mengus, G., Lavigne, A., et al. (1998). "Interaction of human papillomavirus 8 regulatory proteins E2, E6 and E7 with components of the TFIID complex." *Intervirology* **41**: 80-90.

Ephrussi, A., Church, G.M., Tonegawa, S., et al. (1985). "B lineage--specific interactions of an immunoglobulin enhancer with cellular factors in vivo." *Science* **227**: 134-140.

Ewen, M.E., Sluss, H.K., Sherr, C.J., et al. (1993). "Functional interactions of the retinoblastoma protein with mammalian D-type cyclins." *Cell* **73**: 487-497.

Fan, G., Ma, X., Kren, B.T., et al. (1996). "The retinoblastoma gene product inhibits TGF-beta1 induced apoptosis in primary rat hepatocytes and human HuH-7 hepatoma cells." *Oncogene* **12**: 1909-1919.

Fan, S., Yuan, R., Ma, Y.X., et al. (2001). "Disruption of BRCA1 LXCXE motif alters BRCA1 functional activity and regulation of RB family but not RB protein binding." *Oncogene* **20**: 4827-4841.

Felsani, A., Mileo, A.M. and Paggi, M.G. (2006). "Retinoblastoma family proteins as key targets of the small DNA virus oncoproteins." *Oncogene* **25**: 5277-5285.

Feng, Y., Olson, E.C., Stukenberg, P.T., et al. (2000). "LIS1 regulates CNS lamination by interacting with mNudE, a central component of the centrosome." *Neuron* **28**: 665-679.

Ferrante, M.I., Giorgio, G., Feather, S.A., et al. (2001). "Identification of the gene for oral-facial-digital type I syndrome." *Am J Hum Genet* **68**: 569-576.

Ferreira, R., Magnaghi-Jaulin, L., Robin, P., et al. (1998). "The three members of the pocket proteins family share the ability to repress E2F activity through recruitment of a histone deacetylase." *Proc Natl Acad Sci U S A* **95**: 10493-10498.

Fields, S. and Sternglanz, R. (1994). "The two-hybrid system: an assay for protein-protein interactions." *Trends Genet* **10**: 286-292.

Fioretos, T., Panagopoulos, I., Lassen, C., et al. (2001). "Fusion of the BCR and the fibroblast growth factor receptor-1 (FGFR1) genes as a result of t(8;22)(p11;q11) in a myeloproliferative disorder: the first fusion gene involving BCR but not ABL." *Genes Chromosomes Cancer* **32**: 302-310.

Flemington, E.K., Speck, S.H. and Kaelin, W.G., Jr. (1993). "E2F-1-mediated transactivation is inhibited by complex formation with the retinoblastoma susceptibility gene product." *Proc Natl Acad Sci U S A* **90**: 6914-6918.

Foehr, E.D., Raffioni, S., Murray-Rust, J., et al. (2001). "The role of tyrosine residues in fibroblast growth factor receptor 1 signaling in PC12 cells. Systematic site-directed mutagenesis in the endodomain." *J Biol Chem* **276**: 37529-37536.

Giacinti, C. and Giordano, A. (2006). "RB and cell cycle progression." *Oncogene* **25**: 5220-5227.

Gibrat, J.F., Madej, T. and Bryant, S.H. (1996). "Surprising similarities in structure comparison." *Curr Opin Struct Biol* **6**: 377-385.

Gietz, R.D., Triggs-Raine, B., Robbins, A., et al. (1997). "Identification of proteins that interact with a protein of interest: applications of the yeast two-hybrid system." *Mol Cell Biochem* **172**: 67-79.

Gray, S.G. and Ekstrom, T.J. (2001). "The human histone deacetylase family." *Exp Cell Res* **262**: 75-83.

Greenbaum, S. and Zhuang, Y. (2002). "Regulation of early lymphocyte development by E2A family proteins." *Semin Immunol* **14**: 405-414.

Gu, W., Schneider, J.W., Condorelli, G., et al. (1993). "Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation." *Cell* **72**: 309-324.

Guasch, G., Mack, G.J., Popovici, C., et al. (2000). "FGFR1 is fused to the centrosome-associated protein CEP110 in the 8p12 stem cell myeloproliferative disorder with t(8;9)(p12;q33)." *Blood* **95**: 1788-1796.

Guasch, G., Ollendorff, V., Borg, J.P., et al. (2001). "8p12 stem cell myeloproliferative disorder: the FOP-fibroblast growth factor receptor 1 fusion protein of the t(6;8) translocation induces cell survival mediated by mitogen-activated protein kinase and phosphatidylinositol 3-kinase/Akt/mTOR pathways." *Mol Cell Biol* **21**: 8129-8142.

Guo, C.S., Degnin, C., Fiddler, T.A., et al. (2003). "Regulation of MyoD activity and muscle cell differentiation by MDM2, pRb, and Sp1." *J Biol Chem* **278**: 22615-22622.

Habedanck, R., Stierhof, Y.D., Wilkinson, C.J., et al. (2005). "The Polo kinase Plk4 functions in centriole duplication." *Nat Cell Biol* **7**: 1140-1146.

Halevy, O., Novitch, B.G., Spicer, D.B., et al. (1995). "Correlation of terminal cell cycle arrest of skeletal muscle with induction of p21 by MyoD." *Science* **267**: 1018-1021.

Harbour, J.W. and Dean, D.C. (2000). "Rb function in cell-cycle regulation and apoptosis." *Nat Cell Biol* **2**: E65-67.

Harbour, J.W., Luo, R.X., Dei Santi, A., et al. (1999). "Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1." *Cell* **98**: 859-869.

Harper, J.W. and Elledge, S.J. (1996). "Cdk inhibitors in development and cancer." *Curr Opin Genet Dev* **6**: 56-64.

Henthorn, P., Kiledjian, M. and Kadesch, T. (1990). "Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif." *Science* **247**: 467-470.

Hiebert, S.W., Packham, G., Strom, D.K., et al. (1995). "E2F-1:DP-1 induces p53 and overrides survival factors to trigger apoptosis." *Mol Cell Biol* **15**: 6864-6874.

Holm, L. and Sander, C. (1993). "Protein structure comparison by alignment of distance matrices." *J Mol Biol* **233**: 123-138.

Huntington, J.A., Read, R.J. and Carrell, R.W. (2000). "Structure of a serpin-protease complex shows inhibition by deformation." *Nature* **407**: 923-926.

Iavarone, A., Garg, P., Lasorella, A., et al. (1994). "The helix-loop-helix protein Id-2 enhances cell proliferation and binds to the retinoblastoma protein." *Genes Dev* **8**: 1270-1284.

Ito, A., Kawaguchi, Y., Lai, C.H., et al. (2002). "MDM2-HDAC1-mediated deacetylation of p53 is required for its degradation." *Embo J* **21**: 6236-6245.

Jankova, L., Harrop, S.J., Saunders, D.N., et al. (2001). "Crystal structure of the complex of plasminogen activator inhibitor 2 with a peptide mimicking the reactive center loop." *J Biol Chem* **276**: 43374-43382.

Jensen, L.J., Lagarde, J., von Mering, C., et al. (2004). "ArrayProspector: a web resource of functional associations inferred from microarray expression data." *Nucleic Acids Res* **32**: W445-448.

Jensen, P.H., Jensen, T.G., Laug, W.E., et al. (1996). "The exon 3 encoded sequence of the intracellular serine proteinase inhibitor plasminogen activator inhibitor 2 is a protein binding domain." *J Biol Chem* **271**: 26892-26899.

Ji, P., Jiang, H., Rekhtman, K., et al. (2004). "An Rb-Skp2-p27 pathway mediates acute cell cycle inhibition by Rb and is retained in a partial-penetrance Rb mutant." *Mol Cell* **16**: 47-58.

Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." *J Appl Crystallogr* **26**: 795-800.

Kajava, A.V. (1998). "Structural diversity of leucine-rich repeat proteins." *J Mol Biol* **277**: 519-527.

Kaye, F.J., Kratzke, R.A., Gerster, J.L., et al. (1990). "A single amino acid substitution results in a retinoblastoma protein defective in phosphorylation and oncoprotein binding." *Proc Natl Acad Sci U S A* **87**: 6922-6926.

Kennedy, B.K., Liu, O.W., Dick, F.A., et al. (2001). "Histone deacetylase-dependent transcriptional repression by pRB in yeast occurs independently of interaction through the LXCXE binding cleft." *Proc Natl Acad Sci U S A* **98**: 8720-8725.

Khochbin, S., Verdel, A., Lemercier, C., et al. (2001). "Functional significance of histone deacetylase diversity." *Curr Opin Genet Dev* **11**: 162-166.

Kim, H.Y., Ahn, B.Y. and Cho, Y. (2001). "Structural basis for the inactivation of retinoblastoma tumor suppressor by SV40 large T antigen." *Embo J* **20**: 295-304.

Kim, M.H., Cooper, D.R., Oleksy, A., et al. (2004). "The structure of the N-terminal domain of the product of the lissencephaly gene Lis1 and its functional implications." *Structure* **12**: 987-998.

King, K.L. and Cidlowski, J.A. (1998). "Cell cycle regulation and apoptosis." *Annu Rev Physiol* **60**: 601-617.

Kobe, B. and Deisenhofer, J. (1994). "The leucine-rich repeat: a versatile binding motif." *Trends Biochem Sci* **19**: 415-421.

Kratzke, R.A., Otterson, G.A., Lin, A.Y., et al. (1992). "Functional analysis at the Cys706 residue of the retinoblastoma protein." *J Biol Chem* **267**: 25998-26003.

Kruithof, E.K., Baker, M.S. and Bunn, C.L. (1995). "Biological and clinical aspects of plasminogen activator inhibitor type 2." *Blood* **86**: 4007-4024.

Kruse, K. and Julicher, F. (2006). "Dynamics and mechanics of motor-filament systems." *Eur Phys J E Soft Matter* **20**: 459-465.

Ku, N.O., Zhou, X., Toivola, D.M., et al. (1999). "The cytoskeleton of digestive epithelia in health and disease." *Am J Physiol* **277**: G1108-1137.

Lai, A., Lee, J.M., Yang, W.M., et al. (1999). "RBP1 recruits both histone deacetylase-dependent and -independent repression activities to retinoblastoma family proteins." *Mol Cell Biol* **19**: 6632-6641.

Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988). "The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins." *Science* **240**: 1759-1764.

Laskowski, R.A., Moss, D.S. and Thornton, J.M. (1993). "Main-chain bond lengths and bond angles in protein structures." *J Mol Biol* **231**: 1049-1067.

Lasorella, A., Iavarone, A. and Israel, M.A. (1996). "Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins." *Mol Cell Biol* **16**: 2570-2578.

Lassar, A.B., Buskin, J.N., Lockshon, D., et al. (1989). "MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer." *Cell* **58**: 823-831.

Lassar, A.B., Davis, R.L., Wright, W.E., et al. (1991). "Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo." *Cell* **66**: 305-315.

Lawrence, D.A., Ginsburg, D., Day, D.E., et al. (1995). "Serpin-protease complexes are trapped as stable acyl-enzyme intermediates." *J Biol Chem* **270**: 25309-25312.

Lee, C. and Cho, Y. (2002). "Interactions of SV40 large T antigen and other viral proteins with retinoblastoma tumour suppressor." *Rev Med Virol* **12**: 81-92.

Lee, J.O., Russo, A.A. and Pavletich, N.P. (1998). "Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7." *Nature* **391**: 859-865.

Lee, K.Y., Ladha, M.H., McMahon, C., et al. (1999). "The retinoblastoma protein is linked to the activation of Ras." *Mol Cell Biol* **19**: 7724-7732.

Leung, T.H., Hoffmann, A. and Baltimore, D. (2004). "One nucleotide in a kappaB site can determine cofactor specificity for NF-kappaB dimers." *Cell* **118**: 453-464.

Li, X., Zhang, Y.P., Kim, H.S., et al. (2005). "Gene therapy for prostate cancer by controlling adenovirus E1a and E4 gene expression with PSES enhancer." *Cancer Res* **65**: 1941-1951.

Loffler, H., Lukas, J., Bartek, J., et al. (2006). "Structure meets function--centrosomes, genome maintenance and the DNA damage response." *Exp Cell Res* **312**: 2633-2640.

Luo, R.X., Postigo, A.A. and Dean, D.C. (1998). "Rb interacts with histone deacetylase to repress transcription." *Cell* **92**: 463-473.

Ma, P.C., Rould, M.A., Weintraub, H., et al. (1994). "Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation." *Cell* **77**: 451-459.

Macdonald, D., Aguiar, R.C., Mason, P.J., et al. (1995). "A new myeloproliferative disorder associated with chromosomal translocations involving 8p11: a review." *Leukemia* **9**: 1628-1630.

Magenta, A., Cenciarelli, C., De Santa, F., et al. (2003). "MyoD stimulates RB promoter activity via the CREB/p300 nuclear transduction pathway." *Mol Cell Biol* **23**: 2893-2906.

Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., et al. (1998). "Retinoblastoma protein represses transcription by recruiting a histone deacetylase." *Nature* **391**: 601-605.

Massari, M.E. and Murre, C. (2000). "Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms." *Mol Cell Biol* **20**: 429-440.

Mateja, A., Cierpicki, T., Paduch, M., et al. (2006). "The dimerization mechanism of LIS1 and its implication for proteins containing the LisH motif." *J Mol Biol* **357**: 621-631.

Matsushime, H., Roussel, M.F., Ashmun, R.A., et al. (1991). "Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle." *Cell* **65**: 701-713.

McPherson, A. (1999). *Crystallization of Biological Macromolecules*. New York, CSHL Press.

McRee, D.E. (1999). "XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density." *J Struct Biol* **125**: 156-165.

Meraldi, P., Lukas, J., Fry, A.M., et al. (1999). "Centrosome duplication in mammalian somatic cells requires E2F and Cdk2-cyclin A." *Nat Cell Biol* **1**: 88-93.

Meraldi, P. and Nigg, E.A. (2002). "The centrosome cycle." *FEBS Lett* **521**: 9-13.

Metge, B., Ofori-Acquah, S., Stevens, T., et al. (2004). "Stat3 activity is required for centrosome duplication in chinese hamster ovary cells." *J Biol Chem* **279**: 41801-41806.

Meyer, B. and Peters, T. (2003). "NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors." *Angew Chem Int Ed Engl* **42**: 864-890.

Miura, T.A., Morris, K., Ryan, S., et al. (2003). "Adenovirus E1A, not human papillomavirus E7, sensitizes tumor cells to lysis by macrophages through nitric oxide- and TNF-alpha-dependent mechanisms despite up-regulation of 70-kDa heat shock protein." *J Immunol* **170**: 4119-4126.

Molkentin, J.D., Black, B.L., Martin, J.F., et al. (1995). "Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins." *Cell* **83**: 1125-1136.

Mori, S., Nishikawa, S.I. and Yokota, Y. (2000). "Lactation defect in mice lacking the helix-loop-helix inhibitor Id2." *Embo J* **19**: 5772-5781.

Moroni, M.C., Hickman, E.S., Lazzerini Denchi, E., et al. (2001). "Apaf-1 is a transcriptional target for E2F and p53." *Nat Cell Biol* **3**: 552-558.

Morozov, A., Shiyanov, P., Barr, E., et al. (1997). "Accumulation of human papillomavirus type 16 E7 protein bypasses G1 arrest induced by serum deprivation and by the cell cycle inhibitor p21." *J Virol* **71**: 3451-3457.

Morris, E.J. and Dyson, N.J. (2001). "Retinoblastoma protein partners." *Adv Cancer Res* **82**: 1-54.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., et al. (1989). "Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance." *Methods Enzymol* **177**: 44-73.

Mundle, S.D. and Saberwal, G. (2003). "Evolving intricacies and implications of E2F1 regulation." *Faseb J* **17**: 569-574.

Murre, C., McCaw, P.S., Vaessin, H., et al. (1989). "Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence." *Cell* **58**: 537-544.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." *Acta Crystallogr D Biol Crystallogr* **53**: 240-255.

Nakamura, M., Masuda, H., Horii, J., et al. (1998). "When overexpressed, a novel centrosomal protein, RanBPM, causes ectopic microtubule nucleation similar to gamma-tubulin." *J Cell Biol* **143**: 1041-1052.

Nicolas, E., Morales, V., Magnaghi-Jaulin, L., et al. (2000). "RbAp48 belongs to the histone deacetylase complex that associates with the retinoblastoma protein." *J Biol Chem* **275**: 9797-9804.

Nigg, E.A. (2006). "Origins and consequences of centrosome aberrations in human cancers." *Int J Cancer*.

Norton, J.D., Deed, R.W., Craggs, G., et al. (1998). "Id helix-loop-helix proteins in cell growth and differentiation." *Trends Cell Biol* **8**: 58-65.

Novitch, B.G., Mulligan, G.J., Jacks, T., et al. (1996). "Skeletal muscle cells lacking the retinoblastoma protein display defects in muscle gene expression and accumulate in S and G2 phases of the cell cycle." *J Cell Biol* **135**: 441-456.

Ollendorff, V., Guasch, G., Isnardon, D., et al. (1999). "Characterization of FIM-FGFR1, the fusion product of the myeloproliferative disorder-associated t(8;13) translocation." *J Biol Chem* **274**: 26922-26930.

Pagliuca, A., Bartoli, P.C., Saccone, S., et al. (1995). "Molecular cloning of ID4, a novel dominant negative helix-loop-helix human gene on chromosome 6p21.3-p22." *Genomics* **27**: 200-203.

Pan, W., Cox, S., Hoess, R.H., et al. (2001). "A cyclin D1/cyclin-dependent kinase 4 binding site within the C domain of the retinoblastoma protein." *Cancer Res* **61**: 2885-2891.

Pancoast, M., Dobyns, W. and Golden, J.A. (2005). "Interneuron deficits in patients with the Miller-Dieker syndrome." *Acta Neuropathol (Berl)* **109**: 400-404.

Pellecchia, M., Sem, D.S. and Wuthrich, K. (2002). "NMR in drug discovery." *Nat Rev Drug Discov* **1**: 211-219.

Perrakis, A., Morris, R. and Lamzin, V.S. (1999). "Automated protein model building combined with iterative structure refinement." *Nat Struct Biol* **6**: 458-463.

Polesskaya, A., Naguibneva, I., Duquet, A., et al. (2001). "Interaction between acetylated MyoD and the bromodomain of CBP and/or p300." *Mol Cell Biol* **21**: 5312-5320.

Popovici, C., Adelaide, J., Ollendorff, V., et al. (1998). "Fibroblast growth factor receptor 1 is fused to FIM in stem-cell myeloproliferative disorder with t(8;13)." *Proc Natl Acad Sci U S A* **95**: 5712-5717.

Popovici, C., Zhang, B., Gregoire, M.J., et al. (1999). "The t(6;8)(q27;p11) translocation in a stem cell myeloproliferative disorder fuses a novel gene, FOP, to fibroblast growth factor receptor 1." *Blood* **93**: 1381-1389.

Powell, H.R. (1999). "The Rossmann Fourier autoindexing algorithm in *MOSFLM*." *Acta Cryst* **D55**: 1690-1695.

Prabhu, S., Ignatova, A., Park, S.T., et al. (1997). "Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins." *Mol Cell Biol* **17**: 5888-5896.

Pucci, B., Kasten, M. and Giordano, A. (2000). "Cell cycle and apoptosis." *Neoplasia* **2**: 291-299.

Puri, P.L., Avantaggiati, M.L., Balsano, C., et al. (1997). "p300 is required for MyoD-dependent cell cycle arrest and muscle-specific gene transcription." *Embo J* **16**: 369-383.

Puri, P.L., Iezzi, S., Stiegler, P., et al. (2001). "Class I histone deacetylases sequentially interact with MyoD and pRb during skeletal myogenesis." *Mol Cell* **8**: 885-897.

Pytel, B.A., Peppel, K. and Baglioni, C. (1990). "Plasminogen activator inhibitor type-2 is a major protein induced in human fibroblasts and SK-MEL-109 melanoma cells by tumor necrosis factor." *J Cell Physiol* **144**: 416-422.

Rehm, T., Huber, R. and Holak, T.A. (2002). "Application of NMR in structural proteomics: screening for proteins amenable to structural analysis." *Structure (Camb)* **10**: 1613-1618.

Reiter, A., Sohal, J., Kulkarni, S., et al. (1998). "Consistent fusion of ZNF198 to the fibroblast growth factor receptor-1 in the t(8;13)(p11;q12) myeloproliferative syndrome." *Blood* **92**: 1735-1742.

Remold-O'Donnell, E. (1993). "The ovalbumin family of serpin proteins." *FEBS Lett* **315**: 105-108.

Reynaud, E.G., Pelpel, K., Guillier, M., et al. (1999). "p57(Kip2) stabilizes the MyoD protein by inhibiting cyclin E-Cdk2 kinase activity in growing myoblasts." *Mol Cell Biol* **19**: 7621-7629.

Rhodes, G. (2000). *Crystallography Made Crystal Clear*, Academic press.

Romio, L., Fry, A.M., Winyard, P.J., et al. (2004). "OFD1 is a centrosomal/basal body protein expressed during mesenchymal-epithelial transition in human nephrogenesis." *J Am Soc Nephrol* **15**: 2556-2568.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sandal, T. (2002). "Molecular aspects of the mammalian cell cycle and cancer." *Oncologist* **7**: 73-81.

Sapir, T., Cahana, A., Seger, R., et al. (1999). "LIS1 is a microtubule-associated phosphoprotein." *Eur J Biochem* **265**: 181-188.

Sapir, T., Elbaum, M. and Reiner, O. (1997). "Reduction of microtubule catastrophe events by LIS1, platelet-activating factor acetylhydrolase subunit." *Embo J* **16**: 6977-6984.

Schagger, H. and von Jagow, G. (1987). "Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa." *Anal Biochem* **166**: 368-379.

Schneider, J.W., Gu, W., Zhu, L., et al. (1994). "Reversal of terminal differentiation mediated by p107 in Rb-/- muscle cells." *Science* **264**: 1467-1471.

Shen, C.P. and Kadesch, T. (1995). "B-cell-specific DNA binding by an E47 homodimer." *Mol Cell Biol* **15**: 4518-4524.

Shen, W.H., Yin, Y., Broussard, S.R., et al. (2004). "Tumor necrosis factor alpha inhibits cyclin A expression and retinoblastoma hyperphosphorylation triggered by insulin-like growth factor-I induction of new E2F-1 synthesis." *J Biol Chem* **279**: 7438-7446.

Sheng, S., Carey, J., Seftor, E.A., et al. (1996). "Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells." *Proc Natl Acad Sci U S A* **93**: 11669-11674.

Takaki, T., Fukasawa, K., Suzuki-Takahashi, I., et al. (2004). "Cdk-mediated phosphorylation of pRB regulates HDAC binding in vitro." *Biochem Biophys Res Commun* **316**: 252-255.

Tapscott, S.J. (2005). "The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription." *Development* **132**: 2685-2695.

Tietze, K., Oellers, N. and Knust, E. (1992). "Enhancer of splitD, a dominant mutation of Drosophila, and its use in the study of functional domains of a helix-loop-helix protein." *Proc Natl Acad Sci U S A* **89**: 6152-6156.

Toby, G.G. and Golemis, E.A. (2001). "Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions." *Methods* **24**: 201-217.

Umeda, M., Nishitani, H. and Nishimoto, T. (2003). "A novel nuclear protein, Twa1, and Muskelin comprise a complex with RanBPM." *Gene* **303**: 47-54.

Vassalli, J.D., Sappino, A.P. and Belin, D. (1991). "The plasminogen activator/plasmin system." *J Clin Invest* **88**: 1067-1072.

Vidwans, S.J., Wong, M.L. and O'Farrell, P.H. (1999). "Mitotic regulators govern progress through steps in the centrosome duplication cycle." *J Cell Biol* **147**: 1371-1378.

Vizmanos, J.L., Hernandez, R., Vidal, M.J., et al. (2004). "Clinical variability of patients with the t(6;8)(q27;p12) and FGFR1OP-FGFR1 fusion: two further cases." *Hematol J* **5**: 534-537.

von Mering, C., Jensen, L.J., Snel, B., et al. (2005). "STRING: known and predicted protein-protein associations, integrated and transferred across organisms." *Nucleic Acids Res* **33**: D433-437.

Wang, J. and Wilkinson, M.F. (2001). "Deletion mutagenesis of large (12-kb) plasmids by a one-step PCR protocol." *Biotechniques* **31**: 722-724.

Wang, W. and Malcolm, B.A. (1999). "Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis." *Biotechniques* **26**: 680-682.

Weintraub, H., Davis, R., Tapscott, S., et al. (1991). "The myoD gene family: nodal point during specification of the muscle cell lineage." *Science* **251**: 761-766.

Weintraub, H., Tapscott, S.J., Davis, R.L., et al. (1989). "Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD." *Proc Natl Acad Sci U S A* **86**: 5434-5438.

Welch, P.J. and Wang, J.Y. (1993). "A C-terminal protein-binding domain in the retinoblastoma protein regulates nuclear c-Abl tyrosine kinase in the cell cycle." *Cell* **75**: 779-790.

Whyte, P., Buchkovich, K.J., Horowitz, J.M., et al. (1988). "Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product." *Nature* **334**: 124-129.

Wiese, C. and Zheng, Y. (2000). "A new function for the gamma-tubulin ring complex as a microtubule minus-end cap." *Nat Cell Biol* **2**: 358-364.

Xiao, B., Spencer, J., Clements, A., et al. (2003). "Crystal structure of the retinoblastoma tumor suppressor protein bound to E2F and the molecular basis of its regulation." *Proc Natl Acad Sci U S A* **100**: 2363-2368.

Xiao, S., McCarthy, J.G., Aster, J.C., et al. (2000). "ZNF198-FGFR1 transforming activity depends on a novel proline-rich ZNF198 oligomerization domain." *Blood* **96**: 699-704.

Xiao, S., Nalabolu, S.R., Aster, J.C., et al. (1998). "FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome." *Nat Genet* **18**: 84-87.

Xiao, Z.X., Chen, J., Levine, A.J., et al. (1995). "Interaction between the retinoblastoma protein and the oncoprotein MDM2." *Nature* **375**: 694-698.

Yan, X., Habedanck, R. and Nigg, E.A. (2005). "A Complex of Two Centrosomal Proteins, CAP350 and FOP, Cooperates with EB1 in MT Anchoring." *Mol Biol Cell* **17**: 634-644.

Yu, H. (1999). "Extending the size limit of protein nuclear magnetic resonance." *Proc Natl Acad Sci U S A* **96**: 332-334.

Zabludoff, S.D., Csete, M., Wagner, R., et al. (1998). "p27Kip1 is expressed transiently in developing myotomes and enhances myogenesis." *Cell Growth Differ* **9**: 1-11.

Zhang, J.M., Wei, Q., Zhao, X., et al. (1999). "Coupling of the cell cycle and myogenesis through the cyclin D1-dependent interaction of MyoD with cdk4." *Embo J* **18**: 926-933.

Zhang, J.M., Zhao, X., Wei, Q., et al. (1999). "Direct inhibition of G(1) cdk kinase activity by MyoD promotes myoblast cell cycle withdrawal and terminal differentiation." *Embo J* **18**: 6983-6993.

Zhang, O. and Forman-Kay, J.D. (1995). "Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer." *Biochemistry,* **34**: 6784-6794.

Zou, X., Tsutsui, T., Ray, D., et al. (2001). "The cell cycle-regulatory CDC25A phosphatase inhibits apoptosis signal-regulating kinase 1." *Mol Cell Biol* **21**: 4818-4828.

Zou, Z., Anisowicz, A., Hendrix, M.J., et al. (1994). "Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells." *Science* **263**: 526-529.

7 SUPPLEMENTARY MATERIALS

7.1 List of programs and web-pages used to analyse sequences and, structures:

7.1.1 Nucleic acid

Single sequence

- Similarity search BLAST, BLAST 2
- Feature search
  - Restriction sites BCM, Webcutter, Molecular Toolkit
  - ORF, start, stop, poly(A)-signalling sites GeneScan, ORF finder, WebGene
  - Splice sites SSP by NN
  - Transcription factor binding sites TF Search
- PCR primer design BCM, Primer3, NetPrimer, SDM primer generator
- Search for orthologs and homologs BLAST, EST at NCBI
- Translation BCM
- Codon usage Alces, Codon Usage Database, E.coli codon usage, E.coli codon usage analyzer

Multiple sequences

- Aligment BCM
- Generation of consensus sequence BCM, BLAST
- Phylogenetic analisys Philip, Phylodendron

7.1.2 Proteins

Single sequence

- Similarity search BLAST, PropSearch
- Features
  - MW, pI, AA composition ProtParam at ExPaSy
  - Hydrophobicity BCM
  - Transmembrane region TMPred
  - Post-translational modification sites
    - glycosylation NetOGlyc

phosphorylation NetPhos

Signal sequences Psort

Secondary structure predictions PredictProtein, nnPredict

Motifs MAST, PRINTS BLAST

Molecular modelling SWISS-MODEL

Reverse translation Molecular Toolkit

Codon usage Alces, Codon Usage Database, E.coli codon usage, E.coli codon usage analyzer

Multiple sequences

Aligment BCM

Phylogenetic analisys Philip, Phylodendron

7.1.3 Expression system

Codon usage Alces, Codon Usage Database, E.coli codon usage, E.coli codon usage analyzer

Vectors databases VAST at NCBI

Strains databases ATCC, E.coli INDEX

7.1.4 Protocols

Protocol-Online, NWFSC Molecular Biology Protocols, Gerard R. Lazo

7.1.5 NMR data analysis

Assigment Sparky (Goddard 2001)

In-house software CCNMR (Cieslar et al. 1990)

7.1.6 X-ray data analysis

ARP/wARP (Perrakis et al. 1999)

MolRep (CCP4 1994)

Mosflm (Powell 1999)

PHENIX (Adams et al. 2004)

PROCHECK (Laskowski et al. 1993)

Refmac 5 (Murshudov et al. 1997)

SCALA (CCP4 1994)

SHARP (Bricogne et al. 2003)

Swiss-Pdb Viewer (www.expasy.org/spdbv)

WebLab Viewer (http://www.liv.ac.uk/ctichem/swalph.html)

XDS (Kabsch 1993)

XSCALE (Kabsch 1993)

XtalView/Xfit (McRee 1999)

ACKNOWLEDGEMENT

The thesis presented here is a result of the studies carried out in the Department of Structural Research at the Max Planck Institute for Biochemistry. I would like to thank everyone who contributed to this work.

In particular to Professor Robert Huber for the opportunity of working in his department.

To PD Doctor Nediljko Budisa for being my Doktorvater.

To my supervisor Doctor Tad Holak for his many scientific contributions, guidance, patience, and interest in my work.

To the NMR and MPI team (Loyola D'Silva, Madhumita Ghosh, Marcin Krajewski, Joma Kanikadu Joy, Sudipta Majumdar, Przemyslaw Ozdowy, Grzegorz Popwicz, Ulli Rothweiler, Mahavir Singh, Igor Siwanowicz, Dorota Ksiazek, Kinga Brongel, Anna Czarny, Aleksandra Szwagierczak, Tomasz Sitar, Anna Ducka, Anna Tochowicz, Xiumin Yan, Yuri Cheburkin) for their help, advice and good working atmosphere.

I would specially like to thank Magda for her friendship, help in all matters, and coffee time.

Last but not least, I would like to acknowledge my parents and Pawel for all the love, care and constant encouragement.