Doctoral Thesis

Characterization of a novel ankyrin-like protein with transmembrane domains that is lost after oncogenic transformation

Author(s):
Jaquemar, Daniel

Publication Date:
1999

Permanent Link:
https://doi.org/10.3929/ethz-a-003823246

Rights / License:
In Copyright - Non-Commercial Use Permitted
Characterization of a Novel Ankyrin-like Protein with Transmembrane Domains that Is Lost after Oncogenic Transformation

A dissertation submitted to the
Swiss Federal Institute of Technology (ETH) Zürich
for the degree of
Doctor of Natural Sciences

presented by
Daniel Jaquemar
dipl. natw. ETH
born 6.5.1970
citizen of Austria

accepted on the recommendation of
Prof. Dr. K. H. Winterhalter, examiner
PD Dr. B. Trueb, co-examiner
Acknowledgements

I am grateful to Prof. Dr. K. H. Winterhalter for his continuous support and interest in my work during the preparation of this thesis.

I am especially indebted to my direct supervisor, PD Dr. Beat Trueb. He had always the time to discuss problems concerning my work. His kindness made it a pleasure working in his lab. Moreover, I will never forget the numerous "Fondue" and "Spargel" feasts held at his home.

I also enjoyed working with Dr. Jürg Zumbrunn, Thomas Schenker, Daniele Belluoccio, Martin Imhof, and Markus Wiedemann. They all contributed to the excellent atmosphere present in our lab. The many scientific and non-scientific moments and discussions will be unforgettable.

I would also like to thank my parents for always supporting me.

Last but not least, I want to thank Katia for her love and her patience during all these years.

This work was supported by grants from the Swiss National Science Foundation (31-50571.97), the Bernese Cancer Liga, and the M.E. Müller Foundation.
# Table of Contents

Summary .................................................................................................................. 6

Zusammenfassung .................................................................................................... 8

1. Introduction ........................................................................................................ 10
   1.1. Cancer ............................................................................................................ 10
       1.1.1. Oncogenes and tumor suppressors ......................................................... 10
       1.1.2. MMPs .................................................................................................... 14
       1.1.3. Cell adhesion ......................................................................................... 16
       1.1.4. Risk factors ......................................................................................... 18
   1.2. Subtractive library ....................................................................................... 20
   1.3. Zyxin ............................................................................................................ 21
       1.3.1. Other proteins related to zyxin ............................................................... 23
       1.3.2. Binding partners of zyxin .................................................................... 23
       1.3.3. ActA ..................................................................................................... 27
       1.3.4. Possible nuclear function of zyxin ......................................................... 29
   1.4. Collagen VI .................................................................................................. 29
   1.5. GFP ............................................................................................................. 32
   1.6. Aim of the study ......................................................................................... 35

2. Results ............................................................................................................... 36
   2.1. Cloning ......................................................................................................... 36
   2.2. Amino acid sequence .................................................................................. 39
   2.3. Abundance of p120 mRNA ......................................................................... 43
   2.4. Genomic PCR .............................................................................................. 45
   2.5. Localization of the p120 gene ..................................................................... 46
   2.6. Construction of a full length clone ............................................................. 48
   2.7. Characterization of the protein .................................................................. 49
   2.8. Expression of p120 in vitro ........................................................................ 50
   2.9. Expression in cell culture ........................................................................... 53
   2.10. Immunofluorescence ............................................................................... 54
   2.11. GFP studies ............................................................................................... 54

3. Discussion ......................................................................................................... 57
   3.1. Proteins containing ARDs .......................................................................... 57
   3.2. The TRP protein family ............................................................................. 60
   3.3. p120 ............................................................................................................ 63
4.4.5. Bacterial protein expression (GST system) .................................................. 96
4.4.6. Affinity column ....................................................................................... 97
4.4.7. Affinity purification of polyclonal antibodies ............................................. 97
4.4.8. ELISA ....................................................................................................... 98
4.4.9. In vitro transcription/translation ................................................................. 98
4.4.10. Immunofluorescence .............................................................................. 99
4.5. Cell culture .................................................................................................. 100
   4.5.1. Splitting of cells .................................................................................... 100
   4.5.2. Transfection of eukaryotic cells .............................................................. 101
   4.5.3. GFP expression ...................................................................................... 101

5. Literature ........................................................................................................ 102

Curriculum Vitae .................................................................................................. 115
Summary

The expression of a number of proteins is altered after the oncogenic transformation of a cell. Collagen VI expression for instance is down regulated by transformation and has been extensively studied in our lab. To characterize more such proteins, a subtractive cDNA library was constructed. It contained a large number of cDNA clones coding for proteins that were expressed in normal human lung fibroblasts (IMR90), but not in their SV40 transformed counterparts. Several of these novel cDNA clones coded for unknown proteins. One of them, termed p120, was chosen for further analysis.

The full-length cDNA sequence of p120 was isolated by screening a lambda phage cDNA library and by applying the RACE technique. The resulting cDNA encodes a protein of 1119 amino acids with a calculated molecular weight of 127.4 kD.

The comparison of the predicted amino acid sequence of p120 with all entries of the Brookhaven Protein Data Bank led to the conclusion that p120 is composed of two big domains. The N-terminal region contains fifteen well and three less well conserved ankyrin-like repeat domains (ARDs). Each such ARD is 33 amino acids in length and forms a helix-turn-helix motif. The C-terminal region of p120 contains seven hydrophobic elements of 20 amino acids each that probably form six transmembrane domains and one pore loop. Moreover, this region shows a strong homology with members of the trp family of calcium channel proteins and other transmembrane proteins.

Interestingly, a protein that was predicted from the genome project of the nematode C. elegans shows 28% identity (37% similarity) over the full length of our novel protein. This finding suggests that p120 is conserved throughout evolution and that it therefore might participate in basic and important processes.

As expected, p120 expression was down regulated in the vast majority of the tumor cell lines analyzed for mRNA. The abundance of p120 mRNA in human tissues was too low to be analyzed by Northern blotting experiments. The application of a very sensitive PCR approach established that about 500 mRNA molecules representing p120 are present in one IMR90 fibroblast. In contrast to this, the abundance of p120 mRNA in a human embryo is about 1000 times lower. This suggests that p120 is either expressed at a very low level in many cells of the body or that p120 expression is limited to a special cell type.

Genomic PCR proved that p120 is encoded in the human genome. The additional use of the FISH technique showed that the gene is located on human chromosome 8, at position 8q13.
In vitro experiments led to the conclusion that expression of p120 is tightly regulated. In keeping with this is the high content of unusual triplet codons in the ORF (open reading frame) of p120. Moreover, the cDNA sequence around the first ATG codon does not follow the Kozak consensus sequence, but shows characteristics of oncogene initiation regions.

Polyclonal antibodies raised against p120 proved that our novel protein occurs as an integral membrane protein in SW872 and IMR90 cells. Unfortunately, the exact subcellular localization of p120 remains elusive since neither immunofluorescence experiments nor transfection of GFP-p120 fusion proteins into cells yielded clear results.

All these findings could also signify that an as yet unidentified binding partner of p120 is needed for correct function and subcellular localization of p120.

Taking into account all observations, I propose the following model of p120 structure: The N-terminal, cytoplasmic domain is comprised of three subdomains; each consisting of six ARDs. This entire domain is probably responsible for binding several different proteins, thereby bringing them in close contact to each other. The six transmembrane helices and a pore loop present in the C-terminal region most likely form a channel like structure. For proper functioning, four such p120 monomers assemble into either a homotetramer or a heteromultimer together with an as yet unidentified partner protein. It remains to be demonstrated in which cellular process p120 is involved.
Zusammenfassung


Die volle Länge der cDNA wurde durch Absuchen einer Lambda Phagen cDNA Bibliothek und durch Anwendung der RACE Technik isoliert. Diese cDNA kodiert für ein Protein mit 1119 Aminosäuren und einem errechneten Molekulargewicht von 127.4 kD.


Interessanterweise zeigte ein Protein, welches aus dem Genom-Projekt von C. elegans hervorging, 28% Identität (37% Ähnlichkeit) über die volle Länge von p120. Dies deutet darauf hin, dass p120 während der Evolution konserviert wurde und in einem grundlegenden und damit wichtigen Prozess eine Rolle spielt.

Die Expression von p120 war in den meisten der untersuchte Tumor Zelllinien wie erwartet herunterreguliert. Das Vorkommen von p120 mRNA in menschlichen Geweben scheint jedoch zu niedrig zu sein, um mit Hilfe von Northern Blots nachgewiesen werden zu können. Die Anwendung einer sehr empfindlichen PCR Methode führte zum Schluss, dass ungefähr 500 p120 mRNA Moleküle in einem IMR90 Fibroblasten vorkommen. Im Gegensatz dazu ist die Häufigkeit von p120 mRNA in einem menschlichen Embryo ungefähr tausendmal tiefer. Dies bedeutet, dass p120 entweder in einer grossen Anzahl von Körperzellen nur sehr wenig exprimiert wird, oder dass p120 nur in ganz speziellen Zellen vorkommt.
1. Introduction

1.1. Cancer

The history of humankind is also a history of diseases and their prevention. Although advances in hygiene and medicine will never overcome death, they caused a dramatic increase in life expectancy especially in the last couple of centuries. This led to the increased occurrence of illnesses connected with aging. Therefore, cancer has become a major cause of death not only in the industrial world. It has become clear during the last years that cancer is more or less a genetic disease. This means that a mutation in a gene causes a malfunction of the resulting protein. The altered function of such a protein can then contribute to the transformation of a cell. However, there are two big differences between cancer and the classic genetic diseases: First, cancer is mostly caused by somatic mutations and second, cancer arises not from a single mutation ("hit") but rather from the accumulation of mutations over decades. Recent studies concluded that 3 to 6 mutations are required for initiation of neoplastic growth (Vogelstein and Kinzler, 1993). The fact that there exists an exponential relationship between age and tumor incidence supports this concept. Tumors grow by clonal evolution driven by mutation. This means that a cell gaining a growth advantage from a mutation overgrows its neighboring cells. This cell then divides and grows into a tumor. If a cell in this tumor then acquires a new mutation, the same selection process will take place. The sequential acquisition of mutations explains why even after heavy x-ray exposure it takes years or decades for cancer to develop. One type of cancer seems to contradict this concept: Tumors of the eye and kidney occurring in children. A possible explanation might be that fewer mutations are necessary because a hereditary preposition is present. Hereditary factors might include deficiencies in the immune system or mutations already existing in the genome.

1.1.1. Oncogenes and tumor suppressors

A DNA lesion stands at the beginning of every mutated protein (Ames et al., 1995). Lesions can result in point mutations, deletions, or translocations. These lesions are caused by oxidative by-products of normal cell metabolism or by chemicals directly rupturing the DNA backbone. Another source for this damage is cell division, and the more a cell divides, the more lesions accumulate. This theoretical consideration is confirmed by the finding that cancer develops rarely in nondividing cells. Causes for an increased cell division can be hormones, excess calories, chronic inflammations, or
chemicals. It is obvious that lesions occurring in stem cells are the most deleterious. The contribution of chronic inflammation to cancer development should be considered significant, since the immune system fights this inflammation by producing huge amounts of reactive oxygen species. Experiments in rats showed that the contribution of oxidative by-products from normal cell metabolism to cell aging and ultimately cancer is tremendous: A rat cell is exposed to about 100,000 DNA damaging events, co-called "hits", per day. Most of them are repaired, but some damage remains. This means that a 2-year-old rat has about one million DNA lesions in its genome per cell and this is about twice as much as a young rat. A human cell has to face about 10,000 "hits" per cell and day. Not only the lower hit rate but also better repair and control mechanisms are responsible for a lower accumulation rate of DNA lesions in human cells.

There is substantial evidence from in vivo observations supporting the "hit" theory (Fig. 1.1.): By definition, oncogenes promote cell growth (Tab. 1.1.).

### Growth factors and their receptors

- **PDGF** (platelet derived growth factor); involved in brain tumors (glioma).
- **erb-B** receptor of epidermal growth factor (EGF), involved in glioblastoma and breast cancer.
- **RET** contributes to tumors of the thyroid gland.

### Cytoplasmic proteins involved in signal transduction

- **ras** several ras proteins exist, they contribute to cancers of the lung, ovary, colon, pancreas, and leukemia.

### Transcription factors

- **myc** family of different proteins involved in leukemia, neuroblastoma, glioblastoma and cancers of the lung, stomach and breast.

### Other proteins

- **Bcl-1** codes for cyclin D1, stimulates the cell cycle and is involved in tumors of the head and neck.
- **Bcl-2** averts apoptosis and contributes to B-cell lymphomas.
- **MDM2** antagonist for tumor suppressor p53, involved in cancers of connective tissues (sarcomas).

**Tab. 1.1.: Oncogenes.** In their unmutated form, all these proteins are responsible for important cellular functions. But when they are mutated, they become oncogenes.

In their untransformed form, as proto-oncogene, all these proteins are essential for proper cell regulation. But when they are mutated or expressed at the wrong time or place, they become oncogenes. However, the majority of oncogenes are not able to change cell growth on their own when they are transfected into primary rodent
fibroblasts. Only after a second oncogene is cotransfected, cell foci are formed that cause tumors when transplanted into mice. Interestingly, not all combinations of oncogenes seem to work. This might be explained by the fact that not only one, but several cell cycle regulation pathways exist and that the disruption of only one cycle is not enough for initiating neoplastic growth. Since some oncogenes act on the same regulation mechanism there are enough intact checkpoints left to insure proper cell growth.

**Fig. 1.1:** Colon tumor development. Either the loss of tumorsuppressors or the activation of oncogenes contributes to the progression of normal epithelial cells into malignant carcinoma cells.

In order to maintain the integrity of an organism, the amount of cells produced have to match the number of cells dying at the same time. Programmed cell death, or apoptosis, is very important in maintaining this delicate equilibrium. During normal development and growth, cells become apoptotic when they have fulfilled their task. Cancer cells somehow can escape programmed cell death. A well-studied system are the oncogenes Bcl-2 and c-myc: Normally, overexpression of c-myc leads to neoplastic growth. However, if these c-myc overexpressing cells are grown under limiting conditions, such as medium starvation, they become apoptotic. Normal cells would rest in G0 under such conditions. When the c-myc overexpressing cells are transfected with a plasmid coding for Bcl-2, no apoptosis occurs (Bissonette et al., 1992; Fanidi et al., 1992). In vivo, tumor cells might often be in such starving situations, especially when they detach from their neighboring cells and start to migrate. It has been shown that c-myc overexpression and Bcl-2 mutation often occur together in lymphomas. Another example comes from experiments with transgenic mice: Animals that expressed either c-myc or ras under the control of the mouse mammary tumor virus promoter developed breast cancer. Double transgenic mice expressing c-myc and ras developed tumors much faster, meaning that the two oncogenes had a synergistic effect (Sinn et al., 1987). One puzzling fact remains to be explained however. The incorporation of c-myc in the mouse germ line under the control of a breast cell specific promoter should lead to the transformation of every breast cell that overexpresses this gene. However, this is not the case, only a few cells become neoplastic. The reason for this is not known, but an explanation might be that
the overexpression of two oncogenes is not sufficient for initiating tumor growth and that further mutations have to occur. It is likely that other combinations of oncogenes work together to alter the complex balance between growth and death that determines whether a transformed cell will expand into a tumor.

In humans, however, the focus of research lies not on oncogenes but on the investigation of tumor suppressor genes. Recent findings showed that mostly the mutation of tumor suppressors and not the mutation of oncogenes ultimately lead to cancer. About a dozen of different proteins that keep transformed cells under control are well characterized (Tab. 1.2.). A good example for the importance of tumor suppressors is provided by the human papilloma virus: Two proteins encoded in the viral genome act as oncogenes by binding to two well known human tumor suppressive proteins (p53, RB) and thereby initiate neoplastic growth. Close examination of protein expression pattern in tumors showed that one mutated protein seems to be very important in enabling transformed cells to avert apoptosis: The protein p53 was shown to be mutated in 50% of all human tumors. p53 is associated with the replication and repair protein RPA and is thus able to sense DNA damage. Moreover, p53 controls the transition from G1 to S and is therefore an essential element in limiting cell proliferation.

Cytoplasmic proteins

- **APC**: contributes to cancers of colon and stomach.
- **DPC4**: normally involved in growth-inhibitory signal transduction; occurs in pancreas cancer.
- **NF-1**: inhibitor of ras, contributes to neurofibromas, tumors of the peripheral nervous system and myeloid leukemia.
- **NF-2**: involved in tumors of the brain.

Nuclear proteins

- **MTS1**: component of the cell-cycle, normally inhibits growth, affected in a variety of tumors.
- **RB**: very important in cell cycle control, occurs in tumors of retina, bone, breast and bladder.
- **p53**: induces apoptosis, is mutated in 50% of all tumors.
- **WT1**: affected in kidney tumor (Wilms-tumor).

Other proteins

- **BRCA1**: contributes to breast and ovary cancer.
- **BRCA2**: involved in breast cancer.
- **VHL**: involved in kidney cancer.

**Tab. 1.2.: Tumor suppressors.** Many tumor suppressors play important roles in cell cycle regulation and growth control.
Another very well studied example is colorectal cancer where the individual steps leading to metastasis have been elucidated (Fig. 1.1.). A germ line mutation in the tumor suppressor APC leads to the growth of thousands of small tumors in the colon. This is an excellent example for the consequences that the loss of an essential player in cell regulation has for cell growth. Several other mutations in major regulators of cell cycle and growth then cause metastasis. Some further studies led to the conclusion that the chronological order of mutations is rather important for development of neoplastic growth. A mutation at the wrong time might rather constitute a disadvantage than an advantage for a cell.

In most cases, not the original tumor, but daughter colonies spread throughout the body are responsible for death. This destruction of other organs and parts of the body, called metastasis, is therefore examined very thoroughly. It has been shown that less than every thousandth cancer cell that detaches from the primary tumor is able to establish new daughter colonies. These solitary cells are distributed either through the blood or the lymph system. They circulate for about 8 to 24 hours in the blood before they attach to the endothelium of the blood vessels and invade adjacent tissue. Similar to other cells, cancer cells need nutrients and therefore have to be well supplied with blood. One of the most important characteristics of tumors is their ability to induce neovascularization. This renders them independent of endogenous blood vessels and makes every place in the body susceptible for tumor growth. Therefore, many attempts have been made to develop drugs that inhibit angiogenesis. A very recent report now suggests that epigallocatechin-3-gallate, an ingredient of green tea, is a potent inhibitor of angiogenesis (Cao and Cao, 1999) and that drinking green tea might reduce the cancer risk. These exciting new findings will have to be confirmed by other, more sophisticated experiments, of course.

1.1.2. MMPs

During their lifespan, tumor cells need to cross the basement membrane of a number of tissues and organs. These basement membranes, composed of a dense meshwork of collagens (especially collagen IV), laminins and heparan sulfate proteoglycans, do not contain holes large enough to allow passive migration of cells. Therefore, the dissemination of cancer cells has to be an active process and it has been shown that cancer cells can recruit normal cells to assist them in this process. For these reasons, the excessive proteolysis of proteins in the ECM is a sure sign for malignancy. Three steps are required for successful migration through the ECM (Stetler-Stevenson et al., 1993): First, the transformed cells need to attach to the lamina with special receptors. Second, normal cells and the tumor cell itself produce ECM destroying proteins. Finally,
pseudopods are developed that enable the cancer cells to migrate through the proteolytic hole of the ECM. This whole process is not a special feature of cancer cells but a normal characteristic of leukocytes en route to places of inflammation or cells involved in tissue remodeling and wound healing. However, it is not known why cancer cells somehow acquire the ability to migrate and use other cells for help. Events taking place in tissue remodeling and tumor invasion show a striking similarity (Johnsen et al., 1998). The same proteinases involved in metastasis are also involved in wound healing and post-lactational mammary gland involution. In all these processes, normal cells surrounding the affected place are recruited from cells directly involved. Although some features are common to all cancer cells, every cancer has its own protein expression pattern.

Quite some work was undertaken to further study the nature of the proteins involved in the degradation of ECM proteins. About 15 so-called matrix metalloproteinases (MMPs) have been isolated and characterized up to date (Shapiro, 1998). Normally, they are involved in embryo development and growth but also in tissue remodeling and repair. Their inappropriate or excessive expression can lead to arthritis, multiple sclerosis, tooth decay, cardiovascular disease and tumor progression. In an attempt to develop drugs that neutralize these undesirable effects, antibodies against MMPs were raised. These antibodies proved effective in stopping neoplastic growth, at least in vitro. It then turned out that such protease inhibitors, called tissue inhibitors of metalloproteinases (TIMPs), exist also in vivo. A rather strange finding was that cancer cells express MMPs and TIMPs at the same time. After closer examination, however, it became clear that the amount of expressed TIMP is far less than that of MMP. Up to now, several transgenic mice lacking different MMPs are available. Surprisingly, none of them shows serious defects during embryogenesis or after birth. The explanation for this probably lies in the redundancy of the MMP system. One has to wait for multiple knockout mice in order to really establish the exact function of every single member of this protein family. The MMP deficient mice were even able to produce viable offspring, although MMPs are known to be involved in menstruation, mammary involution, or trophoblast invasion. Nevertheless, results obtained from these mice helped to deepen the understanding of MMP function. For example, mice lacking MMP-9 (gelatinase B) appeared to have impaired angiogenesis in bone growth plate, leading to shorter bones in adult animals (Vu et al., 1998). A striking observation in most of these MMP deficient mice was their increased resistance against tumorigenesis. For example, in wildtype mice or humans, long time exposure to cigarette smoke leads to recruitment of inflammatory cells to the lung and consequently to emphysema. In MMP-12 deficient mice, no macrophages are recruited to the lung under such conditions and therefore no destruction of the tissue occurs. It is known for some time, that a ruptured ECM can initiate cell migration, but the signals causing this response remained
undiscovered. Recently, some light has been shed onto the origin of these signals. Fragments of ECM molecules cut off by MMPs have signaling effects on cells. MMP-2 (gelatinase A) cleavage of the laminin-5 γ2 chain exposes a cryptic site that induces the migration of malignant breast cells (Giannelli et al., 1997).

1.1.3. Cell adhesion

In normal cells, cell-cell contact inhibits growth of cells expressing a single oncogene and arrests them in an untransformed state. If this contact inhibition is lost due to killing of neighboring cells, neoplastic grows occurs. This leads to the conclusion that the extracellular matrix (ECM) is very important in keeping up the untransformed state, as seen in many tumors that express a reduced number and amount of proteins responsible for cell adhesion. Most if not all tumors showed that they have abnormal interactions with their environment. In some breast cancers, epithelial cells are incapable of producing a functional and well-organized basement membrane that would normally initiate growth arrest (Boudreau and Bissell, 1998). This malfunctioning is probably due to an impaired communication of the concerned cells with their environment. It has been shown that a number of cell adhesion molecules act as tumor suppressors by maintaining the genomic stability of the cell (Tlsty, 1998). Recently, it could be demonstrated that the loss of adhesion is the cause but not the consequence of malignancy (Perl et al., 1998). This is in accordance with the fact that detached cells have relaxed cell cycle checkpoints and are therefore more susceptible to DNA damage than cells attached to the ECM. Since p53 is mutated in about 50% of all tumors and is also an important element in controlling cell cycle progression, this observation makes a lot of sense.

Two important members of these adhesion molecules are the cadherins and the integrins (Hynes, 1992). Cadherins are important in establishing cell polarity and proper cell differentiation. They are linked to the actin cytoskeleton via their cytoplasmic domain that binds to catenin. Upon calcium binding, two monomers assemble and form an active receptor. E-cadherin is directly involved in the progression of colon adenoma to carcinoma (Perl et al., 1998). Integrins are heterodimers composed of an α and a β chain and constitute highly versatile receptors that mediate cell adhesion, migration, and bidirectional signal transduction across the cell membrane (Dedhar and Hannigan, 1996). Up to now, 16 α and 8 β subunits have been identified. Every α chain can combine with several different β chains and vice versa. So far, 21 different αβ pairings are known (Chotia and Jones, 1997). Extensive studies were done on integrin function and a number of mice lacking a specific integrin subunit are available (Fässler et al., 1996). Many of these knockout mice exhibit
a lethal phenotype, and die either during embryogenesis or early in life. However, it came as a big surprise that few overlapping but many distinct phenotypes were obtained from these knockout mice. This might indicate that integrins participate in very different processes (Fig. 1.2.). To further examine integrin function, mice with subtle mutations in their integrin domains or conditional knockouts of integrin genes should be generated.

It seems that two different requirements have to be met for proper integrin signaling (Miyamoto et al., 1998; Howe et al., 1998): First, the ligand binding site has to be occupied and second, the integrin receptors need to assemble into a cluster.

![Integrin Signaling Diagram]

**Fig. 1.2.: Integrin signaling.** Integrins pass on the signal of extracellular ligands to intracellular signaling pathways (outside-in). On the other hand, intracellular proteins are able to modify the ligand binding properties of integrins (inside-out), thereby altering the outside-in signaling pathway (Yamada, 1997).

A variety of cell types require integrin-mediated adhesion to ECM proteins for survival (Meredith et al., 1993; Ruoslahti and Reed, 1994). Fibroblasts undergo reversible growth arrest when they are removed from the ECM. Even more dramatic is the response of endothelial and epithelial cells to the contact loss with the ECM. They undergo anoikis, as this integrin-dependent apoptosis is called (Frisch and Ruoslahti, 1997). Anoikis seems to be very important in vivo, because it prevents detached cells from attaching at new sites and developing tumors. It is therefore tempting to postulate, that apoptosis is the default pathway entered by cells if an appropriate signal is missing. For example, MDCK cells undergo apoptosis in suspension. However, the overexpression of Bcl-2 or
a constitutively active mutant of FAK renders these cells tumorigenic, despite the fact that these two proteins are not mitogenic in attached MDCK cells (Frisch et al., 1996). This is a further strong indication that resistance to apoptosis plays a significant role in tumorigenesis.

It was also demonstrated that a malignant breast epithelial cell could be restored to a normal state by downregulating the signaling from β1 integrins, thereby reestablishing the normal level of integrin signaling (Weaver et al., 1997). Another very interesting experiment showed how cell shape influences the cellular response to a signal (Kheradmand et al., 1998). Synovial fibroblasts were spread onto a substrate covered with an anti-integrin-α2β1-antibody. These flat cells expressed collagenase-1 independent of the activation of the GTPase Rac. In contrast, cells that were treated with a soluble antibody became roundish and required the GTPase Rac for correct expression of collagenase-1. MMP-2 processing of collagen I exposes a new binding site for α2β1. This allows melanoma cells to receive survival signals and to avert apoptosis. Cryptic binding sites on integrins could also be exposed by mechanical stress, as shown with fibronectin (Zhong et al., 1998). Furthermore, it was demonstrated that the binding of integrins to the ECM could induce the movement of mRNA and ribosomes to sites of focal adhesion (Chicurel et al., 1998). Integrins are also postulated to function as mechanotransducers in a variety of cell types (Shyy and Chien, 1997). Support for this postulated mechanosensing ability comes from experiments proving that the more rigid the ECM is, the more integrins are present at sites of cellular adhesion (Choquet et al., 1997).

Integrin mediated cell adhesion can be regulated either by changing the repertoire of integrin receptors present on the cell surface or by modulating the affinity of the integrins for their ligands (Dedhar and Hannigan, 1996). This modulation is a very important aspect of integrin function. A number of integrins become active (i.e. they are able to bind ligands) or alter their binding preferences in response to intracellular events. The exact mechanism by which integrins transduce their signal through the cell membrane is still not exactly known, but substantial evidence exists that conformational changes in the integrin molecule are responsible. Marcantonio and David (1997) postulate that ligand binding leads to a disruption of the binding between the α and the β chain. This disruption causes the prolongation of the transmembrane α-helix of the β-chain through the proximal cytoplasmic domain. The formation of a new α-helical domain in the cytoplasm then leads to signal transduction.

1.1.4. Risk factors
Microorganisms and viruses are also responsible for cancer. The virus causing hepatitis B for example is the major reason for the high incidence of liver cancer in developing countries. In contrast to the western world, almost nobody is vaccinated. Therefore, about 500 million people in Asia and Africa are infected annually with this virus that at first causes chronic liver inflammation. Another virus, the human papilloma virus, is transmitted sexually and causes cervix cancer. A third of all humans is infected with the bacterium heliobacter pylori, which can cause ulcer, gastritis, and ultimately stomach cancer. Extensive studies showed that asbestos promotes lung cancer by causing a chronic lung inflammation. Hormones are also suspected to contribute to a third of all cancers, mostly by promoting cell division. Because of the diminishing ozone layer and the ensuing increase in UV radiation, black skin cancer (melanoma) is becoming an ever-greater threat.

Hereditiy rarely leads directly to cancer, but it is surely involved in tumors occurring in children or in breast cancer. It has become clear in the last few years that a germ line mutation in the tumor suppressor gene erb-2 makes breast cells more susceptible for transformation.

One major risk for cancer development is smoking. Over a third of all deaths caused by cancer can be attributed to lung cancer alone. This number is all the more frightening since in contrast to many other cancers, the direct cause for lung cancer is well known. Furthermore, smoking is also responsible for a quarter of all deaths caused by heart diseases.

The contribution of chemical compounds present in the environment responsible for cancer development has probably been overestimated for the following reasons: Roasted coffee contains about 1000 different chemicals. In rodents, 19 of 26 substances tested proved to be carcinogenic. The amount of these 19 substances in a single cup of coffee is higher than the amount of synthetic substances in the diet in a whole year. Up to now, no study could show a significant difference in cancer incidence between coffee drinking people and people that do not drink coffee. Two things have to be considered when data obtained from animals are applied to humans: The amount of a substance given to the animals is much higher than the physiological concentration and humans have a much better safety system in their cells, enabling us to live much longer than rodents. Nevertheless, the contribution of the diet to cancer development is significant (Ames et al., 1995). In animal experiments, it could be shown that rats having access to an unlimited food supply develop more cancer than rats fed with a limited amount of calories. Members of the church of the latter day saints, who rarely drink alcohol, smoke or eat red meat but eat a lot of vegetables and fruit, have less cancer than the average American population. This is explicable, as antioxidants present in this diet lower the cancer risk. This is in accordance with the hypothesis that the presence of protective
substances is more important in cancer prevention than avoiding unhealthy factors, since
the endogenous "hit" rate is already very high. Exogenous factors become important
when the amount of antioxidants in the cell is already low. For example, in order to
show the same blood level of ascorbate as non-smokers, smokers should double their
dietary intake of ascorbate, but they rarely do. Some other chemicals present in food
might also contribute to tumor progression by depleting antioxidants: amino acids give
rise to nitrosamines and heterocyclic amines. Polycyclic hydrocarbons occur in charred
meat. Furanus or epoxids can arise from cooking sugars or fat, respectively. Heavy
alcohol consumption causes chronic inflammation and cirrhosis of the liver that can then
lead to cancer.

1.2. Subtractive library

As mentioned above, some proteins are either up- or down-regulated when normal cells
are transformed into tumorogenic cells. Since one example of a down-regulated protein
(collagen VI) had already been extensively studied in our group, we set out to identify
additional proteins that were absent from tumor cells. In order to accomplish this, we
constructed a subtractive library. As starting material for such a library, two different
cell types need to be available: a transformed and an untransformed cell. Normally, the
untransformed counterpart of spontaneous tumor cells is not available. Moreover, there
exists the risk to identify false positive proteins, if one uses tumor cells and an unrelated
untransformed cell line for construction of the subtractive library. Therefore, care must
be taken in choosing the two cell types for the subtractive library. We chose normal
human lung fibroblasts (WI38) and their SV40 transformed counterpart (VA13). The
resulting subtractive library contained more than 400 cDNA clones (Schenker et al.,
1994). 192 of them were used as probes on Northern blots loaded with total RNA
derived from normal and from SV40 transformed fibroblasts. 51 clones (26%) did not
yield any signal at all and were therefore further analyzed by sequencing. It turned out
that they represent contaminations of genomic DNA, since they were related to intronic
sequences and Alu repeats. 99 clones (52%) showed equally strong signals on both
RNA populations and were therefore not further investigated. It has to be noted,
however, that differences up to twofold were not considered significant since Northern
blotting experiments always have some built in experimental error.
42 cDNA clones (22%) fulfilled the prerequisite of down-regulation in transformed
fibroblasts. Further analysis by sequencing demonstrated that they coded for 24
different proteins, meaning that some proteins are represented by more than one clone.
These transformation-sensitive proteins can be grouped in four different classes (Tab.
1.3.): Components of the ECM, enzymes and inhibitors, proteins of the cytoskeleton and regulatory proteins.

ECM components
- Fibronectin (7)
- βig-h3 (5)
- Collagen VI(α2) (2)
- Collagen VI(α3) (1)
- Collagen XVIII (1)
- Galectin-1 (1)
- ECM-1 (1)

Enzymes and inhibitors
- TIMP-2 (4)
- Urokinase (3)
- Collagenase (1)
- Cathepsin B (1)
- Transglutaminase (1)
- L56 Protease (1)

Cytoskeletal proteins
- SM22 (1)
- MAP A1 (1)
- Vinculin (1)
- Zyxin-like (1)
- Ankyrin-like (1)

Regulatory proteins
- IAP (2)
- Betaglycan (2)
- IGFBP-5 (1)
- GTP-binding (1)
- Myosin Kinase (1)
- p406 (1)

Tab. 1.3.: Proteins encoded by the subtractive library. 42 cDNA clones coding for 24 different proteins were obtained. Numbers in parentheses stand for the number of independent cDNA clones present in the library coding for this protein.

Among these 24 proteins were six that were not yet described in the literature at that moment: L56 protease (Zumbrunn and Trueb, 1996a), Zyxin-like (Zumbrunn and Trueb, 1996b), GTP-binding (Schenker et al., 1994), Ankyrin-like (Jaquemar et al., 1999), ECM1 (Bhalerao et al., 1995), and p406. Several clones present in the library coded for proteins already known for their absence in tumor cells: Fibronectin (Hynes, 1990), collagen VI (Timpl and Chu, 1994) and βig-h3 (Skonier et al., 1994). In the following, I will discuss two proteins lost in SV40 transformed fibroblasts: Collagen VI and zyxin.

1.3. Zyxin

Zyxin was first discovered as a protein with an apparent size of 82 kD that localizes to focal adhesion plaques (Beckerle, 1986; Beckerle, 1997). Later it was isolated and
characterized extensively from avian smooth muscle (Crawford and Beckerle, 1991). After the complete cDNA sequence for the avian protein was determined, one discovered that it contains an unusual proline rich N-terminal region and three tandem repeats of the so-called LIM-domain (Sadler et al., 1992) (Fig. 1.3A.).

As with collagens, the proline rich region is responsible for the fact that zyxin migrates with an apparent molecular mass of 82 kD on a SDS polyacrylamide gel, although it has a calculated mass of 58.5 kD (542 amino acids). LIM domains are supposed to bind metal ions, especially zinc, and look very similar to the classical "zinc fingers" first identified in proteins that play important roles in transcriptional regulation and cell differentiation. The classical "zinc fingers" found in transcription factors (TFIIIA) or steroid receptors (estrogen receptor) are shorter in length and responsible for the DNA binding properties of these proteins. However, the zinc-fingers of LIM-domains function as modules for protein-protein interactions and not for binding DNA. Sadler et al. (1992) then showed that zyxin can bind zinc and therefore can build zinc-fingers (Fig. 1.3B.).

Three different groups then isolated and characterized the human equivalent of zyxin (Reinhard et al., 1995; Macalma et al., 1996; Zumbrun and Trueb, 1996b) and showed that the chicken and the human protein are highly conserved. Over the whole sequence, 58% identity can be observed. When only the C-terminal LIM-domains are considered,
this number rises to 74% (or 83% if conservative amino acids substitutions are included). Because of an insertion of a glutamine rich stretch in the middle of the proline rich N-terminal domain, the human zyxin is 572 amino acids in length compared to the chicken zyxin with 542. The murine zyxin was also cloned (Macalma et al., 1996) and it showed 88% identity with the human and 57% identity with the chicken zyxin at the amino acid level. Recently, zyxin from swine and from drosophila were also isolated.

The gene for the human zyxin has been mapped to chromosome 7 at position 7q34-q35 (Zumbrunn and Trueb, 1998). It might be of interest to mention some other genes present at this site. Filamin-2, the B1 subunit of laminin, and the α2 subunit of the capping protein of the muscle Z-line all play a role in the organization of actin filaments whereas 2,3-bisphosphoglycerate mutase and the E3 component of the pyruvate dehydrogenase complex are housekeeping genes. Even some proto-oncogenes (homologues of ras and V-raf, met and tim) and a tumor suppressor gene (tsg 7) are located at this site.

1.3.1. Other proteins related to zyxin

From a research project identifying genes that are involved in chromosomal translocations associated with lipomas, a benign form of mesenchymal tumors in humans, emerged a protein that is closely related to zyxin (Petit et al., 1996). The relationship is striking, since LPP (lipoma preferred partner) also comprises three C-terminal LIM-domains and an N-terminal proline rich region. The lipomas are caused by a rather strange rearrangement of the HMGIC locus, a high mobility group of non-histone chromosomal proteins. The lipoma cells often express chimeric proteins, consisting of the DNA binding domain of HMGIC and the LIM domains of LPP. The exact mechanism by which these chimeric proteins can cause lipomas is unclear. Another protein sharing sequence homologies with zyxin was isolated by a two-hybrid system setup to identify proteins that interact with the thyroid receptor in a ligand dependent fashion: Trip6 (Yi and Beckerle, 1998). It remains to be proven, however, whether these two proteins not only share sequence but also functional similarities with zyxin.

1.3.2. Binding partners of zyxin

Four proteins that bind zyxin emerged from extensive studies: CRP (Sadler et al., 1992), α-actinin (Crawford et al., 1992), VASP (Reinhard et al., 1995) and Vav (Hobert et al., 1996) (Fig. 1.4). The binding site for CRP resides in the first of the three LIM-domains (Schmeichel and Beckerle, 1994) while the binding sites for the other three proteins are located in the N-terminal proline rich region.
α-Actinin

| Actin-binding Domain | 1 | 2 | 3 | 4 | 5 |

CRP

| EFH | LIM |

Vav

| GTP/GDP Exchange | PIK | SH3 | SH2 | SH3 |

VASP

| EVH1 | Pro-rich |

Fig. 1.4.: Zyxin binding proteins. The 100 kD protein α-actinin contains an N-terminal actin binding domain and 4 repeats involved in homodimerization. At the C-terminus two calcium binding EF-hand motifs are located. CRP contains two LIM-domains separated by a glycine rich region. Vav has a rather large array of different protein motifs: One region is involved in guanine nucleotide exchange with small GTP-binding proteins of the Rho-family. A Pleckstrin homology domain (PH) is followed by a zinc-finger motif related to the one found in protein kinase C which can bind diacylglycerol (DAG). At the C-terminus, two SH3 and one SH2 domain can be found. VASP exhibits an N-terminal Ena/VASP homology domain (EVH1) involved in zyxin binding. Additionally, the proline rich stretch in the middle of the protein is responsible for profilin binding.

CRP

crpe (cystenic-rich-protein) is an evolutionarily conserved 23 kD protein expressed in a tissue-specific and transitory way during embryogenesis (Crawford et al., 1994). It is a member of a small family of closely related proteins that show similar molecular organization. Up to now, three proteins of this family are well characterized (CRP1, CRP2, CRP3) (Louis et al., 1997). Each consists of two LIM-domains separated by a glycine-rich region and a putative nuclear localization signal. Evidence exists, that the CRPs play an important role in myogenesis: They are able to bind zyxin and α-actinin and are therefore associated with the actin cytoskeleton. Moreover, they are expressed preferentially in different muscle types (CRP1: smooth muscle, CRP2: arteries, CRP3: striated muscle). Recently, it has been shown (Schmeichel and Beckerle, 1998), that residues residing in both LIM-domains of CRP1 are essential for zyxin binding.

VASP

vaso (vasodilator-stimulated phosphoprotein) is present at integrin-rich adhesion sites and might play a role in dynamic membrane activity (Holt et al., 1998). It was first identified as a substrate for cAMP- and cGMP-dependent kinases involved in the Abl pathway. Since it also binds to vinculin (Reinhard et al., 1996), another member of the focal adhesion plaque, it is thought to play an important role in the assembly of actin.
filaments and cell movement. The human VASP is composed of 380 amino acids and has a calculated mass of 39.8 kD. On a SDS polyacrylamide gel, it migrates with an apparent molecular mass of 46 kD. This shift is caused by a phosphorylated serine that changes the conformation of the protein dramatically and by a proline rich stretch in the middle of the protein. In vivo, VASP appears to assemble into a homotetramer. The N-terminal EVIII domain is responsible for binding zyxin and vinculin, while the central proline rich domain, which is quite distinct from the proline rich region of zyxin, binds profilin, a small g-actin binding protein implicated in actin filament organization and signal transduction.

Vav

The product of the proto-oncogene Vav, p95Vav, is expressed almost exclusively in cells of hematopoietic origin and appears to play a role in signal transduction, especially for antigen-receptor-dependent T- and B-cell activation. It seems to be a nucleotide exchange factor for the Rho family of small GTP-binding proteins. Deletion analysis showed that the C-terminal SH3 domain of Vav is both necessary and sufficient for zyxin binding (Hobert et al., 1996). It has been known for some time that Src homology 2 and 3 (SH2 and SH3) domains are involved in important protein-protein interactions occurring in signal transduction. Vav is expressed only in a limited number of cell types compared to zyxin. Therefore, it seems reasonable to assume that additional proteins responsible for Vav-like functions exist in other cell types. Possible candidates are members of the Src family, which were found to colocalize with zyxin.

α–Actinin

The first protein being identified as a binding partner of zyxin was α–actinin (Crawford et al., 1992). α–Actinin is a 100 kD protein that forms antiparallel dimers and then crosslinks actin filaments. Interestingly, the region in α–actinin responsible for zyxin binding could be identified as the N-terminal globular domain, which is also responsible for actin binding. Whether this means that zyxin and actin compete for α–actinin binding remains to be determined. α–Actinin is codistributed with actin filaments and is present in adhesion plaques where it is postulated to interact with members of the β1 integrin family. This also brings zyxin close to integrins, which is a prerequisite for the postulated signaling features of zyxin (see below).

In the following, I will discuss the α–actinin binding site of zyxin in greater detail, because this knowledge is required for the proper understanding of results I obtained
with GFP-zyxin constructs during the preparation of this thesis. Since these results fit
this chapter perfectly, I will also present and discuss them here.

J. Zumbrunn in our group set out to further narrow down the α-actinin binding site of
zyxin. Using blot overlay experiments, he could observe that radiolabeled α-actinin can
bind to GST fusion proteins containing the first 41 or 51 amino acids of zyxin. In
contrast to this, constructs containing only the first 18 or 27 residues could not bind to
α-actinin anymore. The minimal peptide still capable of binding α-actinin should
therefore correspond to positions 27 to 41. This region (GPVVAPKPKVNPFPRP) contains
either hydrophobic or positively charged amino acids. Interestingly, a quite
similar sequence is also found at position 264 - 280. This second region is conserved in
murine zyxin, but not in chicken zyxin or in LPP. In collaboration with M. Reinhard, we
now wanted to examine whether these two sites are functional in vivo (Reinhard et al.,
1999).

Three different deletion mutants of human zyxin were constructed: One lacking amino
acids 19 - 41 (zyxinΔ1), a second lacking positions 261-283 (zyxin-Δ2) and a third
lacking both sites (zyxinΔ12). These constructs were then expressed in a bacterial
system as GST fusion proteins by J. Zumbrunn and used in blot overlay experiments
with radiolabeled α-actinin. The only construct still capable of binding was GST-
zyxin/3. This leads to the conclusion, that only the first site at position 19-41 is
responsible for α-actinin binding and that the second is negligible.

In addition, by exploiting the GFP system, I could prove that the first site is essential
for correct subcellular localization of zyxin (Fig. 1.5.). Constructs with wildtype zyxin
(GFP-zyxin, Fig. 1.5a.) or GFP-zyxinΔ2 (Fig. 1.5c.) yielded the same results as
previously obtained by immunofluorescence experiments with antibodies against zyxin:
Staining of focal adhesion sites and stress fibers. GFP-zyxin fusion proteins lacking the
first α-actinin binding site (GFP-zyxinΔ1, Fig. 1.5b. and GFP-zyxinΔ12, Fig. 1.5d.)
stained fewer focal adhesion sites and no staining of the actin filaments was observed. In
contrast, even a bright fluorescence of the cytoplasm could be observed with these
constructs. Binding of zyxin to others proteins of the focal adhesion plaque (VASP, Vav
and CRP) is probably responsible for the still detectable staining at these sites. Actin
filaments originating at mitochondria could be produced by using a special expression
vector that directed zyxin to mitochondria. We could show that α-actinin is also
recruited to this site and that the first α-actinin binding site of zyxin is responsible for
this, whereas the contribution of the second site to these functions of zyxin seems to be
insignificant. This means that zyxin is both necessary and sufficient for the recruitment
of α-actinin to sites of actin filament assembly and the following elongation of actin
filaments. The mitochondria were chosen as target site since neither zyxin nor α-actinin
or any of the other zyxin binding partners are present at this location.
A group of researchers studying the ActA protein from Listeria monocytogenes helped to shed more light onto the exact role that zyxin plays at the site of actin filament assembly and organization. The intracellular bacterial parasite Listeria monocytogenes is a food-born pathogen that in mammals can cause life-threatening diseases such as encephalitis. It invades cells by escaping membrane bound endosomes. Once in the cytoplasm, it replicates and recruits the actin machinery of the host cell for its own purposes. Actin filled comets are assembled that enable the bacterium to move inside the cell and more importantly to invade adjacent cells, thereby causing a rapid spread of infection. It has been shown that a single bacterial protein, ActA, is sufficient for the assembly of these actin-filled comets. These comets show a striking similarity to the
organization of the actin filaments at the sites of cellular adhesion in mammalian cells. By probing mammalian cells with antibodies originally directed against ActA, Golsteyn et al. (1997) wanted to identify endogenous mammalian proteins exhibiting a similar function as ActA. A single protein with a size of 82 kD was isolated that turned out to be zyxin. To further examine zyxin function, a tag was added to zyxin that anchored it to the inner site of the plasma membrane, very much like ActA. This construct was used in transient transfection experiments with either cells that have a poorly organized actin cytoskeleton (HeLa cells) or cells with a very well organized actin cytoskeleton (simian fibroblasts such as CV-1 cells). The actin cytoskeleton of both cell types was severely disrupted, namely in a similar way ActA destroys this network. Interestingly, a chimeric protein composed of the N-terminal domain of ActA fused to the N-terminal and central domains of zyxin caused the same cellular response as ActA alone. This seems reasonable when the two proteins are compared: Like zyxin, ActA encompasses several proline rich regions in its N-terminal region, which makes it an ideal binding partner for the same proteins as zyxin. One major difference though is the membrane anchor of ActA. This anchor enables the bacterial protein to circumvent the α-actinin-integrin connection required for correct zyxin function. All these points discussed above lead to the following model of actin filament organization (Beckenle, 1998) (Fig. 1.6.).

Fig. 1.6.: Proposed function of zyxin at sites of actin filament organization. In a typical mammalian cell, zyxin is connected to integrin via α-actinin and recruits proteins of the Ena/VASP family (E/V) that on the other hand bind profilin (P). Profilin then promotes actin filament elongation by binding actin monomers. In addition, active Vav bound to zyxin might locally reduce actin filament capping via a pathway involving proteins of the Rho-family. Together with the Arp2/3 complex, a second function of zyxin might be the promotion of nucleation sites of new actin filaments.

Three possible ways of actin filament elongation exist: 1) Addition of actin monomers to existing filaments at their pointed ends. This is a rather slow and not favorable process. 2) Addition of actin monomers to the uncapped barbed end is fast and the preferred
mode of filament elongation. 3) Formation of new nucleation sites. The Arp2/3 complex was shown to promote this nucleation process by stabilizing the actin-trimer needed for initiation of a new filament (Zigmond, 1998). This nucleation is a very slow process, but it could be shown that ActA greatly accelerates this process by stabilizing the weak bonds present in a freshly growing actin filament. Zyxin now might either contribute to actin filament elongation in a classical way (Fig. 1.6., right) or together with the Arp2/3 complex enhance the nucleation of new filaments in a way similar to ActA (Fig. 1.6., left).

1.3.4. Possible nuclear function of zyxin

Two observations obtained with zyxin suggest a nuclear function of zyxin: First, a close look at the amino acid sequence of zyxin revealed a putative nuclear export signal just in front of the LIM-domains. Second, it was known from the very first experiments with zyxin, that this protein is not only present at the sites of focal adhesion, but also along the stress fibers, which point towards the nucleus. A recent experiment helped to shed more light onto this putative nuclear function of zyxin. Nix and Beckerle (1997) could show that zyxin shuttles between the adhesion plaques and the nucleus. This probably happens along the actin filaments. The mechanism by which this shuttling takes place is unknown. The means by which zyxin enters the nucleus remain to be identified, but the import into the nucleus has to be an active process because zyxin is too large to enter the pores passively. The nuclear export signal present in the protein is responsible for the removal of zyxin from the nucleus since a construct lacking this sequence accumulated inside the nucleus. All these findings are tempting to postulate that zyxin might be part of the signaling mechanism by which integrins pass on their signal from the extracellular space through the membrane to the nucleus. This remains to be proven of course.

1.4. Collagen VI

Collagen VI has been extensively studied in our group. It represents a major component of the ECM of most mesenchymal cells (Timpl and Engel, 1987), but is dramatically downregulated in a number of cell lines which were either transformed by RNA or DNA tumor viruses or which were derived from spontaneous tumors (Schreier et al., 1988). The activity of a single oncogene was shown sufficient for abolishing collagen VI expression. It is important to note, that a collagen VI monomer is built up from three different chains (α1(VI) 140 kD, α2(VI) 140 kD and α3(VI) 260 - 350 kD) that are
encoded by individual genes (Trueb and Winterhalter, 1986). The genes for the \( \alpha_1 \) and \( \alpha_2 \) chain lie adjacent to each other on human chromosome 21, the gene for the \( \alpha_3 \) chain on human chromosome 2. It seems that the synthesis rate for the \( \alpha_3 \) chain is the rate-limiting step when the three chains are assembled into a triple stranded monomer. A dimer is then formed when two monomers align in an anti-parallel fashion. Two such dimers assemble to form a tetramer and the tetramers in turn associate head to head to form microfilaments (Fig. 1.7B.).

 Examination of the protein sequences of the three chains revealed that they are almost exclusively built up from von-Willebrand-Factor-A (vWF-A) domains positioned around a relatively short collagen helix (Fig. 1.7A.). One tempting hypothesis postulates, that by virtue of its vWF-A domains, collagen VI links other collagens to integrins of the \( \beta_1 \)-family, thereby tethering the ECM to the cell. This explains the fact that collagen VI promotes the attachment and spreading of fibroblasts in vitro. In
accordance with this finding is the increased production of collagen VI among other proteins in wound healing processes. Because the synthesis of collagen VI appeared to be regulated at the transcriptional level, the genes for chicken collagen VI were isolated and analyzed. The promoter regions of the genes for the α1 and α2 chain exhibit a very high content of CpG dinucleotides, as found with housekeeping genes and many oncogenes (Koller et al., 1991; Koller and Trueb, 1992). Interestingly, these promoters were active in normal as well as in transformed cells when tested in transient transfection assays, although the endogenous promoters were completely inactive. This suggests that chemical modification rather than the lack of transcription factors are responsible for the altered activity of the promoter. The only known chemical modification of DNA taking place in vertebrate cells is methylation. For collagen I, methylation of the promoter has been shown to abolish expression of the protein (Guenette et al., 1992; Thompson et al., 1991; Rhodes et al., 1994). Moreover, it was also shown that the downregulation of certain proteins correlates with the upregulation of methylase expression. Experiments with the human collagen α2(VI) promoter confirmed these findings (Saitta et al., 1992; Saitta and Chu, 1994). M. Kopp in our group then set out to examine the methylation pattern of the human collagen α2(VI) promoter. By using a novel genomic sequencing strategy, the methylation state of every CpG dinucleotide in the α2(VI) collagen promoter could be established. The examination of promoters from either untransformed or transformed cells showed a dramatic difference in the methylation pattern (Kopp et al., 1997). In contrast to normal cells, almost every CpG dinucleotide in the promoter of transformed cells was methylated. Furthermore, it could be proven that the more CpG dinucleotides were methylated, the less active the promoter was. In addition, several binding sites for the transcription factor AP2 present in these promoter regions (Willmann et al., 1995) were inaccessible after methylation.

Besides its role in the transformed state, methylation plays a very important role during embryogenesis. Early in development, most of the CpG dinucleotides in the genome are methylated. Only the CpG-islands of housekeeping genes are selectively demethylated in order to enable normal cell metabolism. During embryogenesis, genes are selectively methylated or demethylated, depending on their necessity to be active or not at a certain time or place in the embryo. The methylating enzyme has been known for quite some time, but a protein exhibiting demethylating activity could not be identified. However, there is now good evidence that such a demethylating protein has recently been characterized (Bhattacharya et al., 1999). It could be identified by searching a human EST database with a sequence derived from a protein known to bind methylated CpG dinucleotides. A yet unknown cDNA coding for a small protein of about 40 kD was isolated. It could be shown that the protein, when expressed in an in vitro
transcription/translation system, demethylates mCpG dinucleotides. When this protein was expressed in transient transfection experiments, a protein with a size of 160 - 190 kD was observed under native gel electrophoresis conditions. This suggests that in vivo, the protein assembles into a tetra- or pentamer. The coiled-coil structure proposed to be present in the protein is probably responsible for this oligomerization. It was even possible to isolate a protein with demethylating activity from a human cancer cell line, although it remains to be proven that this activity represents the novel 40kD protein. The exact mechanism by which the demethylation occurs is not yet understood in detail, but it seems that the methyl group is directly removed from the cytosine without destroying the backbone of the nucleic acid. A lot of work remains to clarify the exact function of the novel protein in embryogenesis and normal cell activity.

1.5. GFP

It is interesting to note that many organisms capable of bioluminescence have developed this ability independently during evolution and that therefore a lot of different light-emitting systems exist in nature. The driving force behind the study of bioluminescence was the curiosity of researchers to lift the secret of how these different organism are able to convert chemical energy into light. During the last couple of years, this research led to the isolation of proteins that served as novel and exciting tools for researchers with quite different research interests (Wilson and Hastings, 1998). The most prominent members of this novel protein family are luciferase and GFP (Tsien, 1998).

The Green Fluorescent Protein (GFP) from the jellyfish Aequorea victoria has been known for quite some time. Shimomura et al. (1962) first discovered it as a companion to aequorin; the well characterized chemiluminescent protein. During the purifying process of aequorin, Shimomura noted, that "a protein giving solutions that look slightly greenish in sunlight though only yellowish in tungsten lights and exhibiting a very bright, greenish fluorescence in the ultraviolet of a Mineralite, has also been isolated from squeezates." Prendergast and Mann (1978) could then further purify GFP and establish a couple of chemical and physical properties. About at the same time, one big problem concerning aequorin could be solved: Living aequorea tissue shows a green bioluminescence, whereas purified aequorin exhibits a blue chemiluminescence that peaks at 508 nm. This peak matches one of the excitation peaks of GFP and the peak of the luminescence observed in vivo is very close to the one of GFP. Morise et al. (1974) showed that aequorin could transfer its luminescence energy to GFP in a very efficient way in vitro, leading to the conclusion that this also happens in vivo. But then, GFP fell into oblivion for about twenty years. Only after the gene for the aequorea GFP was
cloned (Prasher et al., 1992) and after it was shown that the functional protein could be expressed in prokaryotic or eukaryotic cells (Chalfie et al., 1994; Inouye and Tsuji, 1994), the interest in GFP rose dramatically. In very short time, expression vectors for the 238 amino acid residues containing protein became available. These vectors encoded not the wild type but a modified GFP protein: More than 120 silent codon changes were introduced in order to facilitate expression in mammalian cells. The exchange of serine to threonine at position 65 led to a small shift in the spectrum, which enabled the researcher to use a normal fluorescence microscope equipped for FITC detection for GFP experiments. Some mutations were introduced to improve folding of the protein at 37°C, a temperature where most of the applications for GFP are performed, but where the wildtype protein was not supposed to function. In the last two years, even a blue and a yellow fluorescent variant became commercially available, enabling the researcher to perform double labeling experiments (Rizzuto et al., 1996). In some cases, the extraordinary stability of GFP proved to be a disadvantage. Therefore, even a GFP expression vector containing a PEST-sequence, that shortens the half time of the protein to two hours, is available. GFP turned out to be a very effective and convenient tool for studying the technique of DNA shuffling (Cramer et al., 1996), since the effect of a mutation is easily measurable. One of the major advantages of the GFP system though is the possibility to examine living cells. This allows the study of dynamic processes in real time. GFP proved to be very efficient in studying the dynamics of the cytoskeleton (for a review see Ludin and Matus, 1998) and intermediate filaments (Ho et al., 1998).

Another very rewarding research topics are the investigation of protein transport from the ER to the Golgi and protein secretion (Kaether and Gerdes, 1995; Scales et al., 1997; Presley et al., 1997; Lippincott-Schwartz et al., 1998b; Hirschberg et al., 1998) or the protein import into the nucleus (Carey et al., 1996; Hanakam et al., 1996), where the tagging of several different proteins with GFP led to a dramatic increase in understanding these complex processes. GFP alone was used in studying chaperonin function (Makino et al., 1997). GFP fusion proteins were also used for examining the dynamics of different cell compartments, such as peroxisomes (Wiener et al., 1997) or the nucleus and ER (Subramanian and Meyer, 1997). Furthermore, this new technique allowed new insights into the function of ion-channels (Marshall et al., 1995; Lang et al., 1997) and the mobility and function of centromeres (Shelby et al., 1996).

Although GFP was first crystallized in 1974, the 3D structure was solved just recently by two groups (Ormö et al., 1996; Yang et al., 1996). GFP turned out to be composed of 11 β-strands that form a so-called β-can and a central α-helix that runs through the axis of the β-can. The chromophore is a p-hydroxybenzylideneimidazolinone built up by residues 65 to 67 (Ser-Tyr-Gly) of the central α-helix. The chromophore is buried deep inside the protein, as can be seen in Fig. 1.8. This explains the extraordinary
stability of GFP against chemical and physical influences. The only cofactor needed for correct assembly of the protein seems to be molecular oxygen. This is again an advantage over other fluorescent proteins that often need special enzymes or cofactors for correct assembly in a foreign cell. There is substantial evidence that GFP assembles into a dimer in vivo. Some researchers hope to further modify GFP and use it for quite different applications (intracellular pH-sensor, phosphorescence, generation of reactive oxygen species, photochemical cleavage of chemical bonds). Unfortunately, aequorea GFP genes are the only ones cloned up to date, although many other fluorescent proteins from several other species are known to exist. It remains to be seen if researchers and/or companies are willing to undertake the huge amount of groundwork necessary to isolate more fluorescent proteins useful for investigators.

Fig. 1.8.: 3D structure of the Green Fluorescent Protein (GFP). The protein modeling program Swiss PDBViewer 3.1 (www.expasy.ch/spdbv/mainpage.htm) was used to analyze data available under accession number 1EMA in the Brookhaven Protein Data bank. The 11 β-strands form a β-can that is threaded by a central α-helix. The chromophore built up from amino acids 65 to 67 is depicted in a ball and stick form.
1.6. Aim of the study

Our group has a longstanding interest in the identification of proteins involved in the development of the transformed state of a cell. To this end, a subtractive cDNA library containing cDNA clones coding for proteins that are down regulated in transformed cells was constructed prior to this study. Several cDNA clones coding for proteins already known for their absence in tumor cells, such as fibronectin and collagen VI, were detected in this library. Among the unknown proteins encoded in the library were a zyxin-like protein, a novel GTP-binding protein, and an ankyrin-like protein, termed p120.

The aim of this study was to learn more about the function and structure of p120. In order to accomplish this goal, the full-length cDNA coding for p120 was isolated first and sequenced. Subsequent examination of the sequence revealed similarities with already known proteins. This then gave more insight into a possible function of this novel protein and influenced the nature of the subsequent experiments performed with p120. The first experiments investigated the expression of p120 at the mRNA level in several different human cell lines and tissues. One of the most important experiments was the investigation of p120 at the protein level. For this purpose, antibodies against p120 were raised. These antibodies then helped to study p120 expression in cells and in tissues. Moreover, immunofluorescence experiments were carried out to localize p120 within cells. In order to learn more about p120 function, transfection experiments of p120 constructs in cell culture were also performed.

Results from the different experiments permitted me to propose a function for p120 and to gain first insights into the regulation of this novel protein.
2. Results

In order to identify proteins absent in tumor cells, an approach using the technique of subtractive hybridization was used. mRNA from normal fibroblasts (W138) and from their SV40 transformed counterparts (VA13) was isolated and transcribed into ds cDNA. A special adaptor containing a Sall restriction site was then ligated to the DNA derived from W138 cells, whereas the VA13 molecules were equipped with a biotinylated adaptor. These two populations were then separately amplified by PCR. The amplified cDNA from W138 cells was then hybridized with a tenfold excess of cDNA derived from VA13 cells. The resulting molecules containing biotin were then removed by a streptavidin/phenol extraction. The obtained DNA mixture was again hybridized and extracted. These DNA molecules were amplified by PCR. This "second generation" W138 DNA was then used for another round of hybridization/extraction/hybridization/extraction/PCR. After a third round, the DNA was cut with Sall and cloned into pUC13. The resulting library contained clones coding for proteins no longer expressed in VA13 cells. Besides the already known proteins present in the library (fibronectin, collagen VI and vinculin among others), a couple of novel clones were found, whose mRNA was totally absent in transformed cells. One of them, termed 120, was chosen for this study. In the following, the protein encoded by this DNA will be termed p120.

2.1. Cloning

A 207 bp long cDNA clone (120 sal), derived from the subtractive library, was used to screen a commercial cDNA library by the plaque hybridization method. This library (HL101, CLONTECH Laboratories) had been prepared from human lung fibroblasts (IMR90). After four rounds of screening, 21 independent lambda clones were obtained (Fig. 2.1.), but I had several clues that the 5' end of the cDNA was still not isolated. Therefore, the RACE technique was used to isolate more clones. In a first run, 5 new clones were isolated which reached 150 bp more upstream. In a second RACE experiment, the mRNA was denatured with methyl mercuric hydroxide prior to first strand cDNA synthesis. This resulted in the isolation of 6 additional new clones, which pushed the beginning of our clone 200 bp more upstream (Fig. 2.1.). The sequence of the RACE clones has a very high content of the nucleotides C and G and all of the clones of the second RACE experiment varied by only 2 nucleotides in length. Additionally, a lot of stop codons occur in the first 174 bp. Therefore it seems fair to assume that the 5'
end of the cDNA coding for p120 has been isolated. The DNA sequence encoded by these 32 clones has a size of 4230 bp (Fig. 2.2). The first 174 bp make up the 5' untranslated region and are especially rich in the nucleotides C and G. The first ATG at position 175-177 is followed by an open reading frame of 3357 bp. The stop codon TAG at position 3532 - 3534 is followed by 708 bp of 3' untranslated sequence which harbour two polyadenylation signals AATAAAA: The first one at position 3842-3847 and the second at position 4218 - 4223.

![Fig. 2.1: cDNA clones of p120. All together, 37 (11 from 5'RACE, 21 from lambda-phage libraries and 5 from 3'RACE) clones were obtained. All the 11 clones obtained from 5'RACE are depicted as one (5'RACE). 120 s is the original clone from the subtractive library. The ORF is shaded. Open triangles mark the place of an AATAAAA polyadenylation signal sequence.](image)

Since no poly(A)-sequence was found in the clones ending at this site, I could not be sure to have cloned the whole cDNA sequence. Therefore, I again used the RACE technique to look for further clones. Two types of clones resulted. Four clones ended exactly at the same position (4230) as all formerly isolated clones derived from the lambda phage libraries. But a single clone extended 958 bp further downstream and contained an AATAAAA signal sequence at position 5170-5175 (Fig. 2.2.). Since none of the lambda clones and only one out of 5 3'-RACE clones were found to extend further than 4230 bp, this is the site where the poly-A tail is attached. Moreover, a poly-dT primer was used for the cDNA synthesis in the 3'RACE experiment. This means that without a poly(A) tail, no clone can be isolated. The size of the band observed on a Northern Blot (see below) also supports that the AATAAAA sequence at position 4218-4223 is mostly used for polyadenylation.
Fig. 2.2.: Total cDNA and derived amino acid sequence of p120.
(databank accession number Y10601).

When the cDNA sequence of p120 was compared with all entries of the human EST database, 4 good hits were found (Tab. 2.1). It is striking, that all these ESTs were derived from human tumor samples. All the other matches found are too short (< 20) to
be of any significance. All 4 EST clones were obtained by priming mRNA with an oligo (dT) primer. The first two hits (AA873172, AA972567) most likely represent genomic clones, since the overlap with our protein is only partial (Tab. 2.1.). In addition, the consensus sequence required for an intron/exon boundary is at the expected positions. The other two hits (AI340031, AA502609) represent authentic p120 mRNA, since the overlap extends over their full length. Moreover, these two clones end exactly at the site where the poly(A) tail is attached.

<table>
<thead>
<tr>
<th>accession number</th>
<th>origin</th>
<th>length</th>
<th>overlap</th>
<th>length of overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA873172</td>
<td>two pooled kidney tumor samples</td>
<td>580 bp</td>
<td>1367-1462</td>
<td>96 bp</td>
</tr>
<tr>
<td>AA972567</td>
<td>neuroendocrine lung carcinoid</td>
<td>284 bp</td>
<td>1367-1538</td>
<td>172 bp</td>
</tr>
<tr>
<td>AI340031</td>
<td>neuroendocrine lung carcinoid</td>
<td>400 bp</td>
<td>3840-4236</td>
<td>397 bp</td>
</tr>
<tr>
<td>AA502609</td>
<td>12 pooled colon cancer samples</td>
<td>441 bp</td>
<td>3802-4241</td>
<td>440 bp</td>
</tr>
</tbody>
</table>

Tab. 2.1: Result of a BLASTn search with all entries of the human EST clone databank. The full length cDNA sequence of p120 was used as query in the BLASTn search available at the homepage of the National Center for Biotechnology Information (NCBI) at http://www.ncbi.nlm.nih.gov/.

2.2. Amino acid sequence

The amino acid sequence resulting from the translation of the ORF (open reading frame) is 1119 amino acids in length which corresponds to a weight of 127.4 kD. A database search led to the conclusion that the novel protein can be divided into two large domains (Fig. 2.3.). The N-terminal portion (ANK-domain) is related to ankyrins and ankyrin-like repeat domains (ARD) containing proteins whereas the C-terminal part (TM-domain) is related to transmembrane domains. In addition to the similarities to ankyrin and transmembrane proteins, several further interesting matches were found: From the genome project of the nematode Caenorhabditis elegans emerged one protein, which lies
on chromosome IV and shows 28% identity (37% similarity) over the full length of p120 (Fig. 2.4.). Nothing is known about the function of the C. elegans protein, since the amino acid sequence was simply predicted from the genomic DNA sequence.

Another group of proteins contains the TRP- and TRP-like proteins, which is the only protein family that contains both TM-domains and ARDs and might therefore be related to p120.

There are two regions of 60 and 90 residues at the very N- and C-terminus, respectively, which show no similarity to any known protein motif.

The ANK-domain contains 18 motifs of the so-called ARD (Fig. 2.5.), which was first identified in ankyrin but is now found in many different proteins. When the sequence XGXTPLHLAARXGHVEVVKLLLGDVGADVNXGTK, derived from several hundred ankyrin repeats (Bork, 1993), is used as consensus sequence, 15 repeats show 21-52% identity (or 27-55% similarity if conserved amino acid substitutions are included). The remaining three repeats contain only part of the ANK motif, but show the 33-amino acid length required for correct spacing. The ANK-motif can be found in proteins with quite different function and structure, namely cytoskeletal proteins (ankyrin), membrane-bound receptors (Notch), transcription factors (NF-kB, IKB, bcl-3) and even extracellular toxins (latrotoxin). Most of them contain up to 6 such motifs, only ankyrin with 24 repeats and latrotoxin with 19 contain a comparable number of ARDs.

Fig. 2.4.: Alignment of p120 with an unknown protein from C. elegans. This protein represents the only one which shows 28% identity or 37% similarity over the entire length of the two proteins. Boxed residues indicate identity, shaded ones similarity.
### Consensus ANK repeats

<table>
<thead>
<tr>
<th>Consensus</th>
<th>ANK 1</th>
<th>ANK 11</th>
<th>ANK 12</th>
<th>ANK 14</th>
<th>ANK 10</th>
<th>ANK 2</th>
<th>ANK 4</th>
<th>ANK 5</th>
<th>ANK 8</th>
<th>ANK 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VXXEPLHXAXSEGXXVXXLXGXX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK 6</td>
<td>238</td>
<td>GKITPLHILAVONGDLEMIKMOLND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK 7</td>
<td>271</td>
<td>GCCRENRAFAATTQGATEIKLMISSGS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK 3</td>
<td>131</td>
<td>NNMPLHIAVQGMNHEVMKVLV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK 5</td>
<td>197</td>
<td>WCFCPIQQAASFSGSKECMEL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK 8</td>
<td>308</td>
<td>CHETMLHRASLFDHELADY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK 15</td>
<td>579</td>
<td>QQAFLHLALHNKRKE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 2.5.:** Alignment of all ANK repeats present in p120. Identical residues are boxed. The order of the repeats is according to similarity. The number of the ANK repeats corresponds to Fig. 2.7.

The C-terminal part of p120 contains seven hydrophobic elements of 20 amino acids each which might act as transmembrane helices (Fig. 2.6.). I used the TopPred II program, developed by Claros and von Heijne (1994), to exactly predict the position of the transmembrane elements. The algorithm proposed by Kyte and Doolittle placed these elements at positions 722-742, 767-787, 806-824, 850-870, 873-893, 941-961 and 1010-1030.

**Fig. 2.6.:** Hydrophobicity plot of p120. The algorithm according to Kyte and Doolittle was used. Shaded regions indicate possible transmembrane domains.

Another algorithm developed by Engelman et al. (1986) predicts the same elements with one exception: The element at position 850-870 is placed at position 897-917. This may
lead to the conclusion that 6 of the hydrophobic elements act as transmembrane helices and that the segment at position 850-870, which contains several charged residues, might enter the membrane partially. This is reminiscent of some ion channels, which have six transmembrane domains and a pore loop that opens the channel upon activation.

p120 contains no signal peptide at its N-terminus. Therefore, an internal sequence must act as signal, as shown with the anion exchanger of the red blood cell membrane. The charged residues adjacent to the first transmembrane domain can be used to predict the topological orientation of proteins containing transmembrane domains. This suggests that our protein assumes a type II orientation, meaning that the ANK-domain lies in the intracellular space.

One conclusion of the above points is that only the two glycosylation sites NXT/S which lie in the first extracellular loop (position 747 and 753) are candidates for modification by N-linked carbohydrates. All these points taken together lead to a structure proposed in Fig. 2.7. This structure awaits experimental verification, of course.

![Possible orientation of p120 in the cell membrane.](image)

Fig. 2.7.: Possible orientation of p120 in the cell membrane. Shaded circles depict ANK repeats, open circles stand for sequences sharing only little similarity with the ANK repeat consensus sequence. White rectangles depict the putative transmembrane domains. "Y" stands for an Asn-linked carbohydrate.
2.3. Abundance of p120 mRNA

A Northern blot was loaded with total RNA derived either from WI38 cells, VA13 cells or from two different human embryonic lung fibroblast cell lines (fibroblasts 103 and 104, a gift from P. Stähli, Freiburg). This blot was then hybridized with a radioactively labeled probe of p120 cDNA. One can immediately see that p120 is totally down-regulated in VA13 cells and is therefore a good example for the idea behind our subtractive library (Fig. 2.8.). The size of the band observed with the other RNA samples is 4.6 kilobases, which is in perfect agreement with the size of 4230 bp obtained from the cloning and an assumed length of 300 bp for the poly(A)-tail. To further test whether the absence of p120 is a common feature of tumor cells, RNA from 11 human mesenchymal tumor cells were isolated. A very strong signal could be detected in liposarcoma cells (Fig. 2.8.). Only one of two leiomyosarcoma cell lines showed a signal, but all the other tumor samples (4 rhabdomyosarcomas, two fibrosarcomas, one osteosarcoma and one chondrosarcoma) showed no band of the expected size (Fig. 2.8.). This means that the p120 mRNA is absent from most but not all tumor cell lines analyzed and might therefore be of importance for the maintenance of the normal phenotype.

Fig. 2.8.: Northern blot of p120 mRNA. Total RNA from 4 different normal human fibroblasts as well as from 12 transformed cell lines was separated on a 1% agarose gel and transferred to a nylon membrane. A radioactively labeled probe for p120 (top) or GAPDH (bottom) was hybridized to this membrane. The margins at right mark the migration positions of the ribosomal RNA subunits.
I now asked the question what tissues might express our novel mRNA. For this purpose I hybridized a probe for our clone with commercially available Northern blots containing RNA from different human tissues (Northern Territory™ Human Normal Tissue Blot I + II, Invitrogen: Skeletal muscle, heart, lung, placenta, kidney, liver, pancreas, spleen, thymus, brain, prostate, testis, ovary, small intestine, colon and cultivated leukocytes). Much to my surprise, no signal stronger than background could be observed (not shown). I then examined a blot containing fetal RNA (Northern Territory™ Human Fetal Tissue Blot I: brain, liver, lung and skeletal muscle), since our clone was originally isolated from human embryonic lung fibroblasts. But also this blot did not yield any detectable signal (not shown). A control experiment performed with probes for GAPDH or β-actin showed that both blots contained similar amounts of undigested RNA. The conclusion of these experiments must be that pl20 is expressed at very low levels in most tissues and therefore cannot be detected by Northern blotting.

![Quantitative PCR](image)

**Fig. 2.9: Quantitative PCR.** The numbers in the top row stand for the copies of the standard plasmid added to the PCR which resulted in the amplification of the 410 bp band.

To circumvent this problem, a very sensitive experiment using PCR was employed to evaluate expression of pl20 mRNA in human tissues. Using primer L4061 equal amounts of poly(A)-RNA from IMR90 cells, from a 12-week old human embryo and from embryonic muscle were transcribed into cDNA. A pair of strand-specific primers (U2780, L3115) were then used for PCR. This site was chosen for amplification because during the cloning procedure, one lambda clone (1.2) was found to contain an intron/exon boundary. By positioning the PCR primers around this boundary, one could exclude the possibility that the resulting band was due to a contamination of genomic DNA, since a band originating from genomic DNA would be much larger. To quantitate the amount of pl20 mRNA molecules present in a single cell, I used a second, competitor template as internal standard. This template consisted of the SacI(3365)/PstI(2666) fragment derived from the full length cDNA cloned into pUC19. Additionally, 52 bp derived from pcDNA3.1+ (HaeIII fragment, 2003-2054) were cloned into the EcoRV site (2919) of the insert. This enabled me to distinguish the 358 bp PCR band derived from the mRNA from the 410 bp band derived from this plasmid on a polyacrylamide gel. Serial dilutions
(3-fold) of the competitor were added to the PCR. At a low cDNA to competitor ratio, amplification of the 410 bp band was observed, and vice versa. At equal concentrations of endogenous cDNA and added standard, two bands of the same intensity could be observed. This then allowed the calculation of the exact number of molecules present in the reaction. Equimolar amounts of the two bands occurred when \(3.3 \times 10^6\) template molecules were added to cDNA stemming from IMR90 cells (Fig. 2.9.). This number corresponds to approximately 500 copies of p120 mRNA per fibroblast, which means that p120 is expressed at a moderate level. In a second experiment, cDNA from a 12-week old human embryo was used in this assay. Since the 358 bp band could be observed, p120 is faithfully expressed in human tissues, although at a level 1000-fold lower than in fibroblasts (Fig. 2.9.). Experiments performed with embryonic muscle showed that p120 is expressed there at even lower level, which means that p120 is not expressed in each individual cell of a muscle, but rather in a specific cell type or under specific conditions.

### 2.4. Genomic PCR

At one point during this thesis, I had serious doubt about the origin of p120. On a Southern blot loaded with genomic DNA from IMR90 and HT1080 cells, no signal could be detected (not shown).

![Genomic PCR analysis of p120](image)

**Fig. 2.10.: Genomic PCR analysis of p120.** The staining at the very top of each lane comes from the genomic DNA present in the PCR. The arrows on the right indicate the migration positions of the \(\Phi X/HaeIII\) marker fragments.
I therefore decided to perform a PCR with genomic DNA to prove that p120 is in fact a human protein. The two primers (L4061 and U3848) used in the PCR were chosen because they are in the 3’ untranslated region and no intron/exon boundary should therefore lie in between them. In every human DNA sample tested, a band of the expected size (236 bp) was detectable (Fig. 2.10.). When these bands were excised and digested with EcoRI (pos. 3883) and HindIII (pos. 4020), fragments of the expected sizes (127 bp, 73 bp and 35 bp) were observed on a 8% PAGE (not shown). The band resulting from the PCR with the human brain is weaker than the others because this DNA was isolated from a tissue and not from cells growing in culture. This means that the DNA might not be as pure as the others and therefore might not be a very good template for PCR. Two controls, one containing no DNA at all and the other containing genomic DNA from chicken yielded no signal (Fig. 2.10.). This means that p120 is in fact a human protein and is encoded by a gene in the genome.

2.5. Localization of the p120 gene

In collaboration with Dr. Heng (SeeDNA Biotech Inc., Ontario, Canada) we used the FISH technique to localize the gene coding for p120. A biotinylated DNA probe derived from a full length clone was hybridized to metaphase spreads from human chromosomes. Of 100 chromosomes examined, 89 showed a signal on the long arm of chromosome 8. By superimposing the FISH signal with the DAPI banding pattern, the exact position could be determined as 8q13 (Fig. 2.11.). No other position was found.

![Fig. 2.11.: Localization of the p120 gene in the human genome. The FISH signal (A) and the DAPI banding pattern (B) were superimposed to obtain the exact position of the gene on chromosome 8 as seen in (C).](image-url)
Quite a number of genes are already known to lie in this region (Tab. 2.2.). Some of them are rather interesting and code for proteins which can cause diseases. Most of the proteins responsible for these diseases remain to be identified. Whether p120 has anything to do with the disorders mapping to this position will be revealed when the human genome project (HUGO) will be completed in a couple of years. It might be of interest that the gene for ankyrin 1 also lies on chromosome 8, at position 8p11.2.

<table>
<thead>
<tr>
<th>Location</th>
<th>Symbol</th>
<th>Title</th>
<th>Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>8q12-q13</td>
<td>IL7</td>
<td>Interleukin-7</td>
<td></td>
</tr>
<tr>
<td>8q13</td>
<td>CRH</td>
<td>Corticotropin releasing hormone</td>
<td>ACTH deficiency</td>
</tr>
<tr>
<td>8q13</td>
<td>DUS</td>
<td>Duane syndrome</td>
<td>Duane syndrome</td>
</tr>
<tr>
<td>8q13</td>
<td>TERF1, TRF1</td>
<td>Telomeric repeat binding factor 1</td>
<td></td>
</tr>
<tr>
<td>8q13-q21</td>
<td>FEB1</td>
<td>Febrile convulsions, familial, 1</td>
<td>Febrile convulsions, familial, 1</td>
</tr>
<tr>
<td>8q13-q21</td>
<td>GEM</td>
<td>GTP-binding protein overexpressed in skeletal muscle</td>
<td></td>
</tr>
<tr>
<td>8q13-q21.1</td>
<td>CMT4A</td>
<td>Charcot-Marie-Tooth neuropathy-4A (autosomal recessive)</td>
<td>Charcot-Marie-Tooth neuropathy-4A</td>
</tr>
<tr>
<td>8q13-q22</td>
<td>PDE7A, HCP1</td>
<td>Phosphodiesterase-7A</td>
<td></td>
</tr>
<tr>
<td>8q13-pter</td>
<td>LYN</td>
<td>Yamaguchi sarcoma viral (v-yes-1) related oncogene</td>
<td></td>
</tr>
<tr>
<td>8q13.1-q13.3</td>
<td>TTPA, TTP1, AVED</td>
<td>Tocopherol (alpha) transfer protein</td>
<td>Ataxia with isolated vitamin E deficiency</td>
</tr>
<tr>
<td>8q13.3</td>
<td>EYA1, BOR</td>
<td>Eyes absent, Drosophila, homolog of;</td>
<td>Branchiootorenal syndrome</td>
</tr>
<tr>
<td>8q13.3</td>
<td>NDUFB9, UQOR22</td>
<td>NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9 (22kD, B22)</td>
<td></td>
</tr>
</tbody>
</table>

Tab. 2.2.: Genes located on human chromosome 8 at position 8q13. This table is adapted from data available at the homepage of the National Center for Biotechnology Information (NCBI) at http://www.ncbi.nlm.nih.gov/Omim/.
2.6. Construction of a full length clone

In order to perform further experiments, a full length construct had to be available. Unfortunately, no full length clone was obtained from the screening of the lambda phage libraries. Thus, another method had to be used. In a first set of experiments, I tried to accomplish this by RT-PCR. However, this was not successful, although a lot of different conditions for the PCR were tested. This was probably due to the high CG content present in the 5' region of the mRNA. Even when the mRNA was denatured with methyl mercuric hydroxide prior to first strand synthesis, no band of the expected size was observed. A second difficulty was the quite different composition of the two primers used for the PCR, resulting from the quite different composition of the DNA at the two sites. Another approach was therefore used. First, the SacI 67/PstI 387 fragment of RACE clone 122 (or 91 for a construct lacking the first methionine) was ligated to the PstI 387/Apal 1450 fragment of clone 6.5pl (Fig. 2.12.). In a second step, a fragment ranging from the Apal site (1450) to PstI (2666) was amplified by PCR using the original lambda clone 6.5 as template. This piece was then cloned to fragment PstI (2666)/EcoRI (3883) derived from clone 3.2pl. The resulting two constructs were then ligated using the unique Apal site (1450). This full length construct ranged from nucleotide 67 to 3883 of the sequence obtained by cDNA cloning (Fig. 2.2.). The identity was checked by sequencing of both strands.

![Diagram](image)

Fig. 2.12.: Cloning strategy for a full length clone of p120. The ORF is shaded. Four different DNA fragments obtained by cDNA cloning (Fig. 2.1.) were used.
2.7. Characterization of the protein

In order to further investigate our novel protein, I needed antibodies recognizing p120. Two antisera (T897, T717) were produced by injecting a synthetic peptide comprising the 15 most C-terminal amino acids (1105-1119) coupled to ovalbumin into two rabbits. Both of them recognized their antigen even at high dilution in an ELISA (not shown).

![Western blot analysis of p120](image)

**Fig. 2.13.: Western blot analysis of p120.** Crude preparations of the GST fusion protein (lanes 1 and 4), as well as membrane protein preparations from either liposarcoma cells SW872 (lanes 2 and 5) or transformed fibroblasts VA13 (lanes 3 and 6) were separated on a 3-10% gradient SDS-polyacrylamide gel and stained with Coomassie blue G250 (lanes 1 to 3) or transferred to a nitrocellulose membrane and stained with the affinity purified polyclonal antibody T897 (lanes 4 to 6).

Whole cell extracts from IMR90 cells were separated on a SDS polyacrylamide gel and transferred to a nitrocellulose membrane. Unfortunately, no signal was obtained when these blots were tested with the two antibodies. Therefore, another approach was used. A DNA sequence coding for amino acids 1063 to 1119 was cloned into the bacterial expression vector pGEX-5X. The resulting GST fusion protein (Fig. 2.13., lane 4) was then coupled to an Affigel 10 suspension. Antiserum T897 was purified over a column filled with the coupled Affigel in order to obtain an affinity purified antibody. When this antibody was tested on a blot containing total cell extracts from HTB92 cells, which had shown the strongest signal on a Northern blot, no prominent signal was obtained. But an extract enriched for membrane proteins showed a specific band of 130 kD and several minor bands smaller than 60 kD (Fig. 2.13.). The 130 kD band corresponds to the full length protein while the smaller bands probably represent unrelated proteins, since they were also obtained with other samples. As a control, membrane proteins from VA13
cells or HT1080 cells did not yield any signal. Extracts from IMR90 cells showed a very faint signal, which was too weak to be shown here.

2.8. Expression of p120 in vitro

For further functional studies of our novel protein, I tried to express it in an in vitro transcription/translation system.

<table>
<thead>
<tr>
<th>construct</th>
<th>DNA sequence</th>
<th>protein sequence</th>
<th>vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>120f</td>
<td>67-3883</td>
<td>1-1119</td>
<td>pcDNA3.1</td>
</tr>
<tr>
<td>120s</td>
<td>181-3883</td>
<td>8-1119</td>
<td>pcDNA3.1</td>
</tr>
</tbody>
</table>

Tab. 2.3.: Constructs used for expression of p120 in vitro. Two constructs (120f = full length and 120s = short length) were made. The DNA and protein sequence positions correspond to Fig. 2.2.

pcDNA3.1 was used as expression vector.

I decided to use the TNT coupled transcription/translation kit from Promega. In this system, two different extracts can be used: Rabbit reticulocyte extracts are more suitable for larger mRNA transcripts, whereas wheat germ extracts are preferably applied for shorter mRNA species. In a first set of experiments two different constructs were used (Tab. 2.3.). When these two were employed in a reticulocyte system, a lot of different bands below 60 kD were obtained (Fig. 2.14., right panel). Unfortunately, no band for a full size protein was found. Several controls were included, namely plasmids coding for luciferase, zyxin (Zumbrunn and Trueb, 1996b) and L56 protease (Zumbrunn and Trueb, 1996a). All of them yielded a signal representing the full length protein, although the luciferase sample showed a lot of smaller bands. I switched to the wheat germ system, but the results were similarly disappointing. The amount of protein produced was much smaller and for zyxin and L56 protease, no product was observed at all (Fig. 2.14., left panel).

Since p120 seems to be a transmembrane protein, I added canine microsomes to another series of experiments with reticulocyte extracts. These microsomes should support the synthesis of membrane and larger proteins as shown with collagen XIV (M. Imhof, personal communication). But again, no full length product was observed, in contrast to the two controls, luciferase and zyxin (not shown). The bands observed with the two p120 constructs corresponded to approximately 60 kD (540 amino acids (aa)), 50 kD (450 aa) and 40 kD (360 aa).
Fig. 2.14.: In vitro transcription/translation experiments. Five different constructs (two p120 constructs and 3 controls) were employed either in a wheat germ (left) or in a reticulocyte (right) system. The radioactively labeled proteins were separated on a 10% SDS-polyacrylamide gel. The dried gel was then exposed to X-ray film. The margins in the middle mark the migration positions of globular proteins.

To further investigate the initiation of the transcription and translation, two different constructs were made (Tab. 2.4.). Since these constructs yielded peptides in the range of 2 to 12 kD, the wheat germ system had to be used because the reticulocyte system contains huge amounts of globin, which is 15 kD in size and therefore disturbs the bands in this region. In addition the SDS tricine polyacrylamide gel system was employed to separate peptides of this small size.

<table>
<thead>
<tr>
<th>construct</th>
<th>amino acids</th>
<th>length</th>
</tr>
</thead>
<tbody>
<tr>
<td>122EB</td>
<td>1-106</td>
<td>106</td>
</tr>
<tr>
<td>91EB</td>
<td>8-106</td>
<td>99</td>
</tr>
</tbody>
</table>

Tab. 2.4.: Constructs used for the translation initiation experiment. The two different constructs used are listed. The numbering of the amino acids correspond to Fig. 2.2. pcDNA3.1 was used as expression vector.

When these constructs were used in an experiment, both ATGs functioned as start codon (Fig. 2.15.) and the sizes of the bands observed on the gel are in good accordance with the sizes predicted from the peptide sequences (Tab. 2.5.). When the amount of radioactivity present in a band is divided by the number of methionins present in the
peptide (\textsuperscript{35}S-methionine was used for labeling), it becomes apparent that the first ATG is only about 1.5 times more efficiently used for translation initiation than the second one (Fig. 2.15.)

![Fig. 2.15: Initiation of translation in p120. Two different constructs (see Tab. 2.4.) were used in a wheat germ transcription/translation experiment. The radioactively labeled proteins were separated on a 16.5\% SDS-tricine-polyacrylamide gel. The gel was dried and exposed to X-ray film. The arrows on the right mark the migrating positions of globular proteins.]

<table>
<thead>
<tr>
<th></th>
<th>122EB</th>
<th>91EB</th>
</tr>
</thead>
<tbody>
<tr>
<td>calc.</td>
<td>obs.</td>
<td>calc. obs.</td>
</tr>
<tr>
<td>1. ATG</td>
<td>12.2</td>
<td>11.5</td>
</tr>
<tr>
<td>2. ATG</td>
<td>11.4</td>
<td>10.8</td>
</tr>
</tbody>
</table>

This is in agreement with the following: By comparing several hundred mRNA sequences, it was found (Kozak, 1987) that the following consensus sequence stands for a good initiation of translation (Tab. 2.6.).
The analysis of the Kozak sequences of the first two methionins in p120 shows that the first ATG, in contrast to the second, is not very well suited for good initiation. This is exactly what can be seen in our experiment. An in frame start codon upstream of the effective ATG occurs in fewer than 10% of all vertebrate mRNAs. A notable exception are oncogene transcripts, where this can be observed in two thirds of all cases (Kozak 1987).

### 2.9. Expression in cell culture

In order to learn more about the function of our novel protein, I tried to express it in several different cell lines. For this purpose, a full length construct (120f, Tab. 2.3.) was transfected into human rhabdomyosarcoma (A204) and human fibrosarcoma (HT1080) cells. Positive transfectants were selected by their resistance to G418 as described. Totally, 25 positive cell lines were obtained. Unfortunately, none of the examined clones showed the expected band of 4.6 kb on a Northern blot (Fig. 2.16.). Only one of them showed a signal at all, but this was too small (<3000 b) to encode a full length protein. The same blots were hybridized with a probe for the APF resistance gene and a signal could be observed in every clone, and some of these were very strong. In a batch of freshly transfected cells, a faint band of the expected size could be detected, but it disappeared after subcloning of the G418 resistant clones.

When green monkey (COS-1) cells were used in a transient transfection experiment, similar results were obtained. Up to now, I was not able to express our protein in cell culture, which would be very important for functional studies. I do not know whether this is due to removal of the p120 coding sequence or to its selective inactivation. But it seems that expression of p120 is under tight control because it interferes with normal growth of human cells.
Fig. 2.16.: Transfection of p120 in cell culture. A full length construct for p120 (12o, Tab. 2.3.) was transfected into A204 cells. Total RNA from freshly transfected cells (lane 3) as well as from individual G418 resistant colonies (lanes 4 to 9) were analyzed on a Northern blot. The blot was either hybridized with a probe for p120 (top) or for the APH resistance gene (bottom). As control, RNA from IMR90 cells (lane 1) or from A204 cells (lane 2) was included. The arrows at right mark the migration positions of the ribosomal RNA subunits.

2.10. Immunofluorescence

I then set out to identify the location of p120 in cells. For this purpose, either IMR90 or HTB92 cells were grown to confluency in special plastic chambers and then fixed with paraformaldehyde. The affinity purified antisera T897 was used as primary antibody. As second antibody a FITC conjugated goat anti rabbit antibody was employed. Although a lot of different conditions for fixation and permeabilisation were applied, no distinct signal was obtained. Only an even staining throughout the cell could be seen. One possible explanation might be that the epitope recognized by the antibody is masked by other proteins and therefore not accessible.

2.11. GFP studies

Since the immunofluorescence experiments did not yield any new results, I tried to exploit the GFP system to obtain new insights into the function of p120. From the earlier experiments, I came to the conclusion that it would be better to clone p120 downstream of GFP since the expression of our novel protein seemed to be under tight
control. Three different GFP plasmids were constructed for the following reasons (Tab. 2.7.): p120 consists of 2 large domains and therefore it seemed reasonable to make a fusion protein containing either the ANK or the TM-domain. As a control, a full length construct was also used in the experiments.

<table>
<thead>
<tr>
<th>construct</th>
<th>amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>120TOT</td>
<td>7-1119</td>
</tr>
<tr>
<td>120ANK</td>
<td>7-716</td>
</tr>
<tr>
<td>120TM</td>
<td>696-1118</td>
</tr>
</tbody>
</table>

Tab. 2.7.: Constructs used for GFP fusion proteins. Three different constructs were made: One plasmid coded for the full length protein (120TOT), one for the ANK-domain (120ANK) and one for the TM-domain (120TM). The numbering of the amino acids corresponds to Fig. 2.2.

One big advantage of the GFP system is that one sees immediately which cells had been transfected. Nevertheless, one needs a very high transfection efficiency per cell, since the minimum concentration of GFP visible in the microscope is high (0.1 - 1 μM, this corresponds to $10^5 - 10^6$ copies per cell). Such a high expression might be contraproductive if the fusion protein is toxic to the cell. Therefore, the ideal cell type for these experiments is the COS-1 cell line.

No distinct staining was observed with the full length construct, only an diffuse fluorescence throughout the cytoplasm and the nucleus (Fig. 2.17A). Unfortunately, the control experiment with GFP alone yielded the same result (Fig. 2.17B). The findings from the ANK and TM constructs were more rewarding. When a lot of the GFP-120ANK protein was expressed in a cell, very bright spots around the nucleus (Fig. 2.17E) and sometimes throughout the cytoplasm (Fig. 2.17C) were observed. In some cells, strange duct like structures (Fig. 2.17G) could be seen. Moreover, under very high magnification, the staining turned out to be caused by small, rapid quivering and fluorescing granules.

The examination of the GFP-120TM constructs showed, that the fusion protein is mostly located around the nucleus (Fig. 2.17F,H). In contrast to the cells transfected with the GFP, GFP-120ANK or GFP-120TOT plasmids, the GFP-120TM cells were mostly small and roundish in shape (Fig. 2.17D). These GFP experiments add to the conclusion, that both the transcription and the translation of p120 is highly controlled because it plays an important role in cell regulation.
Fig. 2.17.: GFP fusion protein experiments. COS-1 cells were transfected either with constructs for GFP alone (A), GFP-120TOT (B), GFP-120ANK (C, E, G) or GFP-120TM (D, F, H). Bar 10 μm.
3. Discussion

During this thesis, I investigated a novel protein that is downregulated in a large number of tumors. A comparison of the amino acid sequence from our novel protein p120 with all entries of the protein databank of the National Center of Biotechnology Information (NCBI) revealed several similarities to already known proteins (Fig. 2.3.). Among these were a large number of proteins containing a protein motif of 33 amino acids, the so-called ankyrin-like repeat domain (ARD). Another group was the family of TRP and TRP-like proteins. In order to better discuss the results I obtained with p120, I will describe these two protein families more deeply in the following.

3.1. Proteins containing ARDs

Ankyrin was first identified in red blood cells, where it anchors (hence the name ankyrin) the spectrin-based skeleton to the integral membrane band 3 protein (Bennett and Stenbuck, 1979). Several other proteins are also known to bind to ankyrins. These include spectrin (Platt et al., 1993), the erythroid anion exchanger (Michaely and Bennett, 1995; Ding et al., 1996), the Na⁺/K⁺ ATPase (Jordan et al, 1995; Thevananther et al., 1998), the IP₃ receptor (Bourguignon et al., 1993), and the ryanodine receptor (Bourguignon et al., 1995), among others. Up to now, three different human ankyrin genes have been identified. They are located on human chromosomes 8p11 (ANK1, ankyrinR), 4q25-27 (ANK2, ankyrinB) and 10q21 (ANK3, ankyrinC). Extensive alternative splicing and differential subcellular distribution contributes to a large functional diversity in the ankyrin family. Ankyrin₉ is mostly expressed in erythrocytes, but is also found in muscle and the cell bodies and dendrites of a subset of central nervous system neurons. AnkyrinB is the prominent isoform in the nervous system. The extensive alternative splicing observed especially with ANK3 transcripts contributes to its expression in many different tissues, including kidney and the nervous system, where it is located at the nodes of Ranvier (Kordeli et al., 1995; Zhang and Bennett, 1996). But also inside the cell, ankyrinC is located at different sites, e.g. at the postsynaptic membrane, the sarcoplasmatic reticulum of skeletal muscle (Kordeli et al., 1998) or at the Golgi apparatus (Devarajan et al., 1996). This observation led to a broadening of the proposed function of the spectrin-ankyrin skeleton. This network not only organizes specialized membrane domains by selectively recruiting integral membrane proteins to their specific subcellular location (Bennett and Gilligan, 1993),
but also organizes cytoplasmic compartments and regulates intracellular traffic (Lippincott-Schwartz, 1998a; De Matteis and Morrow, 1998). However, the basic structure of all these protein isoforms encoded by the three ANK genes is the same and consists of three different domains: A highly conserved N-terminal 89-95 kD domain contains 24 ARDs that are involved in protein-protein interactions. Several experiments led to the conclusion that this N-terminal region forms four subdomains, each comprised of six ARDs (Michaely and Bennett, 1993). The central 62 kD domain is also well conserved and binds to spectrin. The C-terminal 55kD domain is the regulatory domain and changes the activity of the first two regions. As expected, malfunctioning or loss of ankyrins causes several diseases (Lambert and Bennett, 1993; Peters and Lux, 1993). A well-studied but relatively mild illness is the hereditary hemolytic anemia, caused by morphologically abnormal erythrocytes that lack the stability and deformability required in blood microcirculation. Several mutations in the ANK1 gene could be linked to one form of this disease, the hereditary spherocytosis (Eber et al., 1996).

With time, many other ARD containing proteins were characterized. In 1993, Bork found more than 650 ARDs in 91 different proteins from almost all phyla (Bork, 1993). It seems that the minimum number or ARDs present in a single protein is four. ARDs occur in such diverse proteins as enzymes (GLsk/Glsl), toxins (α-latrotoxin), channels (TRP/TRPL), membrane bound receptors (Notch), and transcription factors (IkB, SWi6). These proteins are all located intracellularly with the notable exception of the toxin of the black widow spider, α-latrotoxin. With 20 ARDs this protein is the only one besides the novel protein characterized in this thesis that contains almost as many repeats as the ankyrins. α-latrotoxin causes its toxicity by activating all types of synapses. It will be interesting to learn more about the exact mechanism by which the ARDs of this toxin accomplish this neurotransmitter release.

Another very interesting research field where ARDs play a tremendously important role is the NFκB/IκB signaling pathway. Many of the genes encoding proteins involved in inflammation are turned on by the transcription factor NFκB (Baeuerle, 1998). NFκB rests in an inactive form in the cytoplasm and is kept there by its inhibitor IκB, which strongly binds to the transcription factor. In case of an inflammation signal, IκB is first phosphorylated in its N-terminal region by a large multikinase complex, then polyubiquitinated, and finally degraded by the proteasome. The released NFκB is translocated to the nucleus, where it binds to its DNA recognition sequence and thereby initiates the transcription of several genes. The transcription of IκB is also activated by NFκB and this newly synthesized IκB enters the nucleus. The second ARD of IκB is responsible for its nuclear localization (Sachdev et al., 1998). Once in the nucleus, the inhibitor is then able to remove active NFκB from its DNA binding site, turning it into an inactive form. The NFκB/IκB complex is then transported out of the nucleus into the
cytoplasm, thereby closing the regulation cycle. The recent elucidation of the three dimensional structure of an NFκB/IK κB complex made it possible to understand the exact mechanism by which NFκB is inactivated and what role the ARDs play (Huxford et al., 1998; Jacobs and Harrison, 1998). The structure of the ARDs present in IkB is in accordance with the structure obtained from other ARD containing proteins such as 53BP2 (Gorina and Pavletich, 1996), the cyclin-dependent protein kinase inhibitors p18\textsuperscript{INK4c} (Venkataramani et al., 1998) and p19\textsuperscript{INK4d} (Luh et al., 1997), the β-subunit of transcription factor GABP (Batchelor et al., 1998), myotrophin (Yang et al., 1998), and the yeast cell cycle regulator Swi6 (Foord et al., 1999).

A single ARD is built up by two antiparallel α-helices separated by a β-turn. The C-terminal loop is perpendicular to the axis of the helices, giving the structure an L-shaped appearance (Fig. 3.1.).

![Secondary and tertiary structure of an ARD](image)

Several ARDs then align side-by-side to form a broad interface for protein-protein interactions. In IkB, the six ARDs form a slightly bent cylinder with loops protruding from the packed stack of α-helices (Fig. 3.2.). The amino acid residues located in these
loops are essential for the ability of IkB to bind to NFkB, thereby rendering it inactive. The six ARDs fulfill different tasks. The first two repeats inhibit nuclear localization by binding to a long stretch of NFkB that contains an NLS (nuclear localization sequence). The binding of NFkB to repeats 3 to 6 induces a dramatic change in the conformation of NFkB, thereby burying residues needed for DNA binding. There are still some uncertainties and contradictions in the two structures solved by Huxford et al. (1999) and Jacobs and Harrison (1999). This can be explained, at least in part, by small sequence differences in the bacterially expressed proteins used for the crystallization experiments. Additional experiments with other members of this protein family and with NFkB bound to DNA will surely help to clarify the situation.

Fig. 3.2.: Tertiary structure of the six ARDs present in IkB. The amino acids in the loops (dark gray) connecting the ARDs are responsible for the interactions with NFkB. The structure was modeled using the program Swiss PDBViewer 3.1 and the data available under accession number 1NFI in the Brookhaven Protein Data Bank.

3.2. The TRP protein family

A number of calcium influxes exist in a variety of different cells. One of them, the so-called store-operated calcium entry (SOCE) is thought to be the most important way of calcium influx in nonexcitable cells (Montell, 1997). SOCE (formerly known as capacitative calcium entry) is activated by depletion of internal calcium stores. The exact identity of such pools remains elusive, but some parts of the ER are supposed to constitute a large portion of these stores. Generally, intracellular calcium release results in the activation of tyrosine kinase receptors or G protein-coupled receptors, which in turn activate PLC-γ. The following production of IP₃ induces a biphasic rise in calcium
concentration. The first phase is transient and caused by the activation of the IP$_3$ receptor, a calcium release channel located in the calcium stores. The resulting depletion of internal calcium stores leads to the more sustained opening of calcium channels located in the plasma membrane. But not until recently, candidates for such channel proteins in the plasma membrane were cloned and characterized. Experiments with these proteins then seemed to corroborate the hypothesis that this novel protein family is essential for SOCE (Friel, 1996; Hardie, 1996; Montell, 1997, Minke and Selinger, 1996).

Probably the best system for studying SOCE is the drosophila photoreceptor cell. The very thorough examination of vision in drosophila helped tremendously in isolating proteins involved in SOCE. Almost all proteins involved in this signal transduction pathway are known. Light activated rhodopsin activates a G$_q$ protein that in turn stimulates an eye-specific phospholipase C (PLC) encoded by the norpA locus. This PLC is essential for vision, since norpA mutant flies are totally unresponsive to light. PLC then catalyzes the hydrolysis of 4,5-PIP$_2$ to IP$_3$ and DAG. IP$_3$ activates the IP$_3$ receptor that releases calcium, while DAG activates PKC. The exact mechanism by which activation of PLC leads to the opening of cation specific ion channels of the trp family is not well understood and many contradictory hypotheses exist in the literature. Initially, TRP was identified because photoreceptor cells from flies with a mutation at the trp (transient receptor potential) locus were not able to respond to bright light for a longer period. The cloning of the cDNA encoded by the trp gene showed that the proposed protein is 1275 amino acids in length and contains four ARDs in the N-terminal region and six transmembrane domains at its C-terminus. A second protein (TRP-like, TRPL) was subsequently isolated that shared about 40% sequence identity with TRP over the N-terminal 700 amino acids, but differed significantly over the C-terminal region. Experiments showed that this C-terminal region is responsible for the SOCE response in TRP, but not in TRPL. As with other channels, a putative pore-loop domain between transmembrane segments 5 and 6 is responsible for channel opening. Moreover, it is thought that the functional protein is a tetra- or pentamer. In vitro transfection experiments in cell culture performed with the two proteins helped to deepen the knowledge of their function. It was shown that TRP is a bona fide store operated channel (SOC) (Xu et al., 1997). However, not all the characteristics of the in vivo calcium currents could be imitated by these transfection experiments. It seems reasonable to postulate that some other proteins are assisting TRP in its function. A very good candidate for this task is TRPL. It was then shown that TRPL is constitutively active when transfected into cells and that it does not constitute a SOC (Xu et al., 1997). Nevertheless, TRPL is a functional ion channel, but with a less strict ion selectivity than TRP. It could be shown that heteromultimers between TRP and...
TRPL form when two plasmids coding for the two different proteins are transfected simultaneously. Additional experiments then led to the conclusion, that TRP and TRPL-homomultimers are part of the light activated current observable in drosophila photoreceptor cells (Niemeyer et al., 1996). Although there are two reports in the literature, that TRP and TRPL can form heteromultimers in heterologous systems (Gillo et al., 1996; Xu et al., 1997), there is strong evidence that this is not the case in vivo in drosophila photoreceptor cells (Scott and Zuker, 1998). The exact mechanism by which depletion of internal calcium stores is coupled to activation of the photoresponse is still not known. The signal does not seem to be calcium itself, since SOCE can be activated in the presence of calcium chelators. A couple of different hypotheses exist (Hardie, 1996, Montell, 1997). A small soluble factor, called calcium influx factor is postulated to mediate the signal. There are ambiguous reports about the isolation of such a factor and further studies need to be done to clarify the identity of such a protein. A second theory proposes that the IP₃ receptor changes its conformation upon emptying of the internal calcium stores and is thereby able to directly interact with TRP (Berridge, 1995). There are some experiments in support of this hypothesis, but substantial evidence is still missing.

The conservation of TRP in a variety of invertebrates led to the search of TRPs in mammals. Consequently, a human relative of TRP, called TRPC1 was isolated (Wes et al., 1995; Zhu et al., 1995, Zitt et al., 1996). TRPC1 shares about 40% identity with TRP/TRPL over the first 625 amino acids. The C-terminus of the human protein is quite distinct from the drosophila proteins. At least five other members of this family were identified in mice and humans (Zhu et al., 1996; Okada et al., 1998; Groschner et al., 1998; Boulay et al., 1997, Birnbaumer et al., 1996). Initial experiments performed with all these proteins showed that they are involved in SOCE. The most highly conserved regions among all members of this family are the most N-terminal amino acids containing four ARDs, the transmembrane domains, and the 60 amino acids C-terminally of these.

Although the mechanism of vertebrate vision is quite distinct from invertebrate vision, a similar signaling transduction as seen with invertebrate vision is also seen in vertebrate cells. The phosphoinositide pathway is a ubiquitous, G-protein coupled signaling pathway that not only causes the release of calcium from IP₃ sensitive stores, but also the influx of calcium across the plasma membrane. This influx is essential for a sustained cellular response and has been postulated to be involved in a variety of different processes, such as T-cell activation, mast cell degranulation, cell proliferation, and apoptosis (Putney, 1990, Berridge, 1995). The critical signal for the calcium influx seems to be the reduced calcium concentration in the internal stores. This calcium influx can be initiated experimentally by agents that deplete such stores, such as thapsigargin, a potent inhibitor of calcium-ATPases.
There exist a number of different types of store operated conductances in vertebrate cells. A very well characterized SOCE is $I_{\text{crac}}$ (= calcium release-activated calcium current). A key question concerning $I_{\text{crac}}$ is whether any of the TRP channels are responsible for the highly calcium-selective, low-conductance, store-operated current typical of $I_{\text{crac}}$. A finding from extensive experiments was that no single member of the trp family could induce a current with the properties of $I_{\text{crac}}$. This does not mean that TRPs are not involved in $I_{\text{crac}}$, but that most likely several different channels produce this current. The possibility remains, however, that the TRPs are not SOC themselves, but constitute regulatory or accessory proteins for some yet unidentified channel proteins.

Recent results suggest that not all members of this family are SOCs. The human Trp3 protein for example was shown to be activated after PLC activation, but not after store depletion (Zhu et al., 1998). This does not exclude that Trp3 forms heteromultimers with other proteins that are sensitive to store depletion.

In vivo, however, things seem to be different. For example, no calcium release from internal stores during drosophila phototransduction could be detected, which would be expected if TRPs function as SOC. Moreover, the depletion of internal calcium stores should lead to TRP activation. Up to now, this could not be shown in vivo, only in vitro. Another argument against a SOC function for TRPs is the fact that the calcium stores are located far away from the phototransduction complex and that diffusion of a signaling molecule would be too slow to account for the rapid phototransduction seen in drosophila, which is one of the fastest signaling cascades known.

It seems therefore that phototransduction in drosophila is after all much more complicated than thought a few years ago, since several experiments contradict each other (Montell, 1998).

### 3.3. p120

During the preparation of this thesis, I have cloned the cDNA coding for a novel transformation-sensitive protein, termed p120. I examined the cDNA sequence and the presence of our gene in the human genome. The protein was characterized from several different aspects. To learn more about the function of p120, I expressed it in vitro and in cell culture. The expression of p120 in different tissues and cell types on the mRNA and protein level was also looked at. Moreover, antibodies against p120 were raised and used for several different experiments, such as immunofluorescence and Western blotting. Different domains of p120 were expressed as GFP fusion proteins in transient transfection experiments to learn more about its subcellular localization.
3.3.1. cDNA and protein sequence of p120

Already the cloning of the cDNA coding for p120 turned out to be difficult. The high CG content in the 5' region of the mRNA causes strong secondary structures that disturb the enzymes used for reverse transcription. Therefore, no clones containing the 5' end of our cDNA could be isolated from lambda phage libraries. These problems could be solved by using the RACE technique and consequently a number of clones coding for the 5' end of p120 were isolated. The secondary structures present in the mRNA also seemed responsible for the inability to obtain a full-length cDNA clone by RT-PCR. The full-length clone needed for future experiments was therefore put together from several clones originating from lambda phage libraries and RACE experiments. After the initial screening of lambda phage libraries, no clone containing a poly-(A) sequence at its 3' end was found. Once more, the RACE technique was employed. Several additional clones, all containing a poly(A) tail, could be isolated. The subsequent aligning of all the clones obtained during the cloning procedure resulted in a cDNA that was 5188 bp in length. The investigation of the full-length cDNA sequence revealed some interesting facts. The first 250 bp show a very high CG content, although only the first 174 base pairs make up the 5' untranslated region. The ORF with 3357 bp in length encodes a protein of 1119 amino acid residues and is followed by a 3' untranslated region of 1658 bp. Three polyadenylation signals AATAAA could be found in the 3' untranslated region. Up to now, only a single band of 4600 bp was observed on all Northern blot experiments performed with p120 mRNA. Two findings support the idea that the second AATAAA signal is mostly used for polyadenylation. First, the size of the mRNA band obtained from Northern blots matches the length of the cDNA at this second polyadenylation signal, considering a poly-(A) tail of about 300 nucleotides. Second, all of the cDNA clones obtained during lambda phage library screening and all but one of the 3' RACE clones ended at this second signal.

Experiments with ankyrin_R mRNA showed that the three different polyadenylation signals present in this transcript are used in a tissue specific way (Birkenmeier et al., 1993). Since the 3' untranslated regions of the mouse and human ANK1 gene show an unusual high homology, this region might contain elements contributing to the regulation of expression. This regulation might take place by stabilizing the mRNA or by subcellular localization of the transcripts. Further experiments will have to show if similar mechanisms also play a role in p120 regulation.

The fact that only four human EST clones matching the p120 cDNA sequence were found in a databank search also suggests that p120 is expressed at a very low level and only in a fraction of the cells in the human body. It is interesting to note, however, that
all these EST clones are derived from tumor samples. This clearly contradicts our own findings that p120 is absent from most of the tumor cell lines that we analyzed. Only RNA from a liposarcoma cell line showed a stronger signal on a Northern blot than that obtained from IMR90 cells. A possible explanation for this seemingly contradiction might be that p120 expression is downregulated when cells are transferred from their in vivo environment to an artificial in vitro environment, such as a cell culture dish. Another finding from the examination of the cDNA sequence contributed to a possible explanation for the difficulties encountered in the expression of p120 in vitro. The considerable number of rare and unusual amino acid codons might cause the ribosome to slow down.

I did not examine the gene structure and organization of our novel protein, but it is known from studies performed with the ANK1 gene that almost each ARD is encoded by a separate exon (Gallagher et al., 1997). It can be assumed that the same holds true for our gene. Further studies performed with this novel protein could also include the isolation of its promoter region. The thorough analysis of this region might then supply some more clues that would explain the tight regulation of p120 expression.

A genomic PCR approach established that p120 is encoded in the human genome and with the help of the FISH technique the position of the gene was determined as 8q13. Interestingly, a number of diseases also map to this region. Whether p120 has anything to do with these deficiencies remains to be shown, of course.

As mentioned above, the comparison of the amino acid sequence of p120 with a protein database led to several interesting matches, including ankyrins and ARD containing proteins, members of the trp family, and a protein from C. elegans. This nematode protein shows 28% identity or 37% similarity over the full length of p120. It is therefore fair to assume that the protein from C. elegans is the homologue or at least a close relative of our novel protein. The fact that a close relative of p120 was also found in the nematode C. elegans suggests that this protein is conserved throughout the animal kingdom and is responsible for a basic and therefore important process.

### 3.3.2. RNA experiments

Up to now, the examination of Northern blots loaded with RNA derived from more than 20 different adult and embryonic human tissues did not yield any signal at all. Several explanations are conceivable to account for this observation. p120 expression could be limited to a special cell type and/or to a limited time span during development. Another possibility might be that a special signal is needed for the initiation of p120
transcription and translation. If this were the case, the search for p120 expression in vivo would be almost like searching a needle in a haystack.

The application of a very sensitive quantitative PCR approach helped to learn more about p120 expression in cell culture and in human tissues. These experiments showed that about 500 mRNA molecules representing p120 are present in cultured human lung fibroblasts, which means that p120 is expressed at a moderate level. However, experiments with mRNA from a human embryo showed that p120 expression there is about one thousand times lower. This data can be interpreted in two different ways: Either only every thousandth cell in a human embryo shows the same level of expression as cultured fibroblasts or p120 expression is extremely low in a wide variety of cells. This low occurrence would explain our failure to detect p120 expression in human tissues by Northern blotting. Human embryonic muscle yielded even a lower signal. Since this method is much more sensitive than Northern blotting, some future experiments using this PCR method could identify tissues and cell types expressing p120 at a level comparable to cultured fibroblasts. If such at tissue would be identified, other methods such as in-situ hybridization could be employed to further characterize p120 expression.

**3.3.3. Protein experiments**

The ultimate goal of every study is to characterize the function of the investigated protein. One important tool to fulfill this task are antibodies. Therefore, polyclonal antibodies against the C-terminal region of p120 were raised. The antisera obtained from the immunization of two different rabbits were used for immunofluorescence and Western blotting experiments. To further enhance the specificity of the antibodies, they were subjected to an affinity purification.

Unfortunately, no distinct signal could be observed in immunofluorescence experiments performed with cells that showed a signal for p120 expression on Northern blots. The same even staining throughout the cells was also seen in cells that did not express p120 at all, such as VA13 cells. A possible explanation for this might be that the epitope recognized by the antibody is not accessible or even absent. Since only the last 15 amino acids of p120 were used for the immunization of the rabbits, the antibody might only detect a linear epitope and not the three-dimensional structure this region forms in vivo.

An antibody raised against another region of p120 might help in solving this problem.

The affinity-purified antibody was also used in Western blotting experiments with whole cell extracts from IMR90 or SW872 cells. Unfortunately, no signal of the expected size was observable. The quality of the antibodies could not be responsible for
these negative results, since they recognized their antigen even at high dilution, as shown with an ELISA. To further enhance the sensitivity of the experiments, not only whole cell extracts but also extracts enriched for integral membrane proteins were prepared. A band of the expected size was detected on Western blots containing membrane protein extracts from IMR90 or SW872 cells. The signal from IMR90 cells was very weak but clearly detectable. In contrast to these results, no band of the expected size could be seen in membrane protein extracts from cells that did not show any p120 mRNA signal on a Northern blot. This strongly suggests that the band observed in extracts from IMR90 or SW872 cells clearly represents p120.

Results obtained with an in vitro transcription/translation system contributed to our understanding of p120 regulation. In a first set of experiments, I investigated the initiation of transcription and translation. These experiments showed that both methionines at position 1 and 8, respectively, can function as start codons, although the first one seems to be more efficient. When only the second methionine is available, a much higher expression efficiency is observed compared to experiments where both methionines are present. This result is in agreement with findings made by Kozak (1987). She investigated the transcription start regions of several hundred different vertebrate mRNA molecules. It turned out that a guanine, mostly an A, is conserved at position -3, considering the A of the first ATG as position 0. This is exactly what can be seen with the second methionine, but not with the first one present in p120. The first methionine could function as an upstream initiation codon, as seen with proto-oncogenes. They often contain upstream ATGs that do not follow the Kozak consensus sequence. These upstream ATGs are thought to slow down the ribosome. This data is another indication that p120 expression is tightly controlled.

The failure to produce a full length product for p120 in an in vitro transcription/translation system can be explained in two different ways. As mentioned above, a large number of unusual triplett codons is present in the ORF of p120. This signifies that the pool of corresponding tRNA's in the reaction mixture is quickly exhausted and therefore translation of p120 stops. Another possibility is that additional proteins not present in the in vitro system are needed for correct expression of p120. These additional proteins could act on the transcriptional level, by stabilizing the mRNA, or on the translational level, by facilitating the folding of the protein.

The expression of p120 in a cell culture system confirmed our earlier findings regarding p120 regulation. For these transfection experiments, one plasmid containing the full length coding sequence for p120 and one lacking the first methionine were constructed. These two different constructs were then transfected either into A204 or into COS-1 cells. In order to obtain stable cell lines, the cells were treated with an antibiotic to kill all the cells that were not transfected and were therefore not expressing the resistance gene
encoded by the plasmid used for transfection. This procedure led to the isolation of about two dozen individual stable cell lines. The mRNA from these cells was examined on a Northern blot. Surprisingly, no signal representing the full-length p120 mRNA was detectable. Only one clone yielded a signal, but the observed size of this band suggests that it represents a truncated p120 mRNA. As a control, expression of the antibiotic resistance gene was also examined. Much to my surprise, a strong band of the expected size was observable in every stable cell line. This can only mean that p120 expression was selectively suppressed in these cells. It has to be noted, however, that shortly after the transfection and before the selection for individual stable cell lines was performed, a signal for p120 mRNA was observable on Northern blots. Several different modes of inactivation of p120 expression are conceivable. The inactivation has to be selective for the p120 cDNA, since the APH resistance gene is expressed. Maybe the coding sequence is excised and removed. Another possibility might be that the promoter is inactivated by methylation or by proteins that bind to this region. All these findings suggest that p120 expression appears to interfere with normal cell growth. Future experiments could involve the transfection of plasmids encoding only portions of p120. These experiments could identify the region of p120 causing the tight regulation.

3.3.4. GFP experiments

In order to learn more about the function and the subcellular localization of p120, I decided to exploit the GFP system. For this purpose, three different GFP fusion plasmids were made. One construct beared the cDNA for the full-length p120 protein (120TOT) while the other two contained either the N-terminal region made up of the ARDs (120ANK) or the C-terminal region featuring the transmembrane domains (120TM). As a control, a plasmid coding for GFP alone was used. These four constructs were transfected into monkey epithelial cells (COS-1), rhabdomyosarcoma cells (A204), or fibroblasts (IMR90). The low transfection efficiency observed with A204 and IMR90 cells did not allow an evaluation of the results obtained in these two cell lines. In contrast to this, COS-1 cells were easily transfectable. The expression of GFP alone resulted in an even staining throughout the cell, including the nucleus. The only structures that were not stained were sharp, sometimes ring like structures. The identity of these structures remains unknown. Experiments with full-length p120 showed exactly the same results as with GFP alone. This result supports the idea that p120 needs some other proteins for proper function and subcellular localization and that this partner is missing in COS-1 cells. The transfection experiments with the two
individual domains yielded some very interesting results. In some cells transfected with the ARD containing construct, bright spots throughout the cytoplasm and around the nucleus could be observed. These spots probably consist of aggregated proteins. This is quite plausible, since the ARDs are protein-protein binding modules. The reason for the aggregation might be the high expression of the GFP-fusion protein and the lack of the natural binding partners for p120. Some other transfected cells showed duct-like structures extending throughout the cells. The identity and nature of these structures remains unknown. Some very fine, but distinct, hair-like structures could easily be detected because they did not show any staining at all. It will be interesting to determine the nature of these structures. When the stained cells were looked at under high magnification, it became evident that small, quivering, and fluorescing granules caused the staining observed in all these cells. The movement of these granules seemed to be of random orientation. Again, this observation supports the idea that these GFP-ARD constructs constitute sticky proteins that aggregate into small clusters. In support of this is also the fact that these granules could only be observed in GFP constructs containing ARDs. The transfection experiments with GFP-120TM constructs yielded some rather surprising results. In contrast to normal COS-1 cells, transfected cells were mostly roundish and small. Since these constructs contained the putative transmembrane domains, I expected the fluorescence signal at the plasma membrane. This was not the case, as the staining was concentrated around the nucleus. Preliminary experiments using the technique of confocal microscopy showed that the strong staining around the nucleus was due to a higher concentration of the fluorescence signal and not simply to an accumulation of the signal because the cell is much thicker at this place. These findings support the idea that the stained compartment is in fact the Golgi network. A couple of reasons might explain why these GFP-120TM constructs are located in the Golgi apparatus. First, it could be that the normal location of p120 is the membrane of the Golgi network and not the plasma membrane. Second, because the N-terminal region of p120 is missing in these constructs, misfolding could take place that would lead to an accumulation of the GFP-120TM constructs in the Golgi. Future experiments could include Golgi-specific markers that would clearly show if the GFP-120TM fusion protein is located in the Golgi apparatus.

3.4. Conclusion and future prospects

Several findings support the idea that p120 constitutes a novel type of ion channel whose expression is tightly regulated on the transcriptional and translational level. The
N-terminal domain of p120 is made up of a large number of ARDs that are known to be involved in protein-protein interactions. It is very likely that the C-terminal region of p120 contains six transmembrane domains and one pore loop that enters the lipid bilayer partially. Such pore loops are proposed to mediate the ion selectivity of ion channels (MacKinnon, 1995). The proposed pore loop of p120, formed by amino acids 850 - 870, contains two glutamate residues. It has been shown that especially in calcium channels, such glutamate residues provide the ion specificity needed for proper functioning of the channel. If p120 constitutes such a novel ion channel, it has to assemble into a tetramer to be functional, as seen with other ion channels. There is also the possibility that p120 forms heteromultimers with other as yet unidentified partner proteins.

Considering not only the organization of the transmembrane domains but also the recent elucidation of the three dimensional structure of ARD containing proteins, I propose the following tertiary and quaternary model for p120 assembly (Fig. 3.3.). This model shows that a tremendous number of ARDs are present at the site of p120 localization. This suggests that many different proteins can bind to p120. Therefore, a possible function of our novel protein might be to recruit several different proteins as seen with the INAD protein involved in drosophila phototransduction.

Two different ways of p120 action are possible: The opening of the channel could influence the binding properties of the cytoplasmic domain. On the other hand, binding of proteins to the ARDs could lead to the opening of the channel.

A novel method might help to isolate these binding partners. During the last few years, the two-hybrid system provided researchers with a very powerful tool for studying protein-protein interactions. Experiments with ankyrinR showed that it can bind to at least seven different proteins, though not at the same time. This finding suggests that a comparable number of proteins might also bind to p120. Different portions of the N-terminal domain of p120 could be employed in the two-hybrid system in order to isolate possible binding partners.

Future experiments could also use GST-fusion proteins consisting of different parts of p120. These peptides could then be used to transfect cells or to further investigate p120 binding to proteins isolated from the two-hybrid system.

Another approach would include the construction of transgenic mice. This procedure would demand a big effort, since the murine equivalent of our novel protein needs to be identified and cloned first. This would enable us to make transgenic animals either lacking a functional p120 or expressing a mutated p120. However, these experiments are connected with a big financial and human effort.

More knowledge about the function of p120 might also come from other researchers studying proteins that are involved in a novel signaling pathway or in a deficiency.
Fig. 3.3.: Proposed tertiary and quaternary model of p120. The cytoplasmic domain of p120 is composed of 15 well and 3 less well conserved ARDs. From studies performed with other proteins such as ankyrin, it can be assumed that the cytoplasmic domain is composed of three subdomains. A subdomain is formed by the stacking of six ARDs. The two antiparallel α-helices build up the structural backbone and make contact with the other ARDs in the same and in the other subdomains. The loops connecting the separate ARDs then make up the surface of the stacks and are responsible for protein-protein interactions (see also Figs. 3.1. and 3.2.). The six transmembrane domains and the poor loop make up the C-terminal region of p120. Four such p120 monomers can then assemble into a tetramer, as seen with many ion channels. For better separation of the four monomers, they are shown in different shades of gray.
4. Methods

Unless noted otherwise, chemicals of the highest quality available were purchased from Fluka or Sigma. Many of the protocols described here are modifications of methods described by Ausubel et al. (1987) or Sambrook et al. (1989).

4.1. DNA techniques

4.1.1. Microprep (Rusconi)

When a vector without blue/white selection was used for cloning or when a large number of clones had to be looked at, the following method was used.

Lysis Buffer

1 mg/ml lysozyme
1 mg/ml RNase (DNase free)
50 mM Tris-HCl
50 mM EDTA
pH 8.0
store at 4°C

1) Resuspend either a single colony of an LB plate or the pellet of 50 μl of an overnight culture in 10 μl lysis buffer
2) Incubate at RT for 10 min
3) Add 1 μl phenol, vortex, and centrifuge for 5 min
4) Mix 5 μl of the supernatant with 1 μl 5x DNA sample buffer and load onto a 0.7% agarose gel (0.5x TBE), containing 0.5 μg/ml EtBr
5) After running of the gel, three bands can be observed under UV illumination: supercoiled, nicked-circular and linear plasmid DNA. Differences in the migration of the supercoiled DNA band indicate whether or not the construct is carrying

4.1.2. Alkali Lysis Miniprep

Alkali Lysis Buffer I

50 mM glucose
25 mM Tris-HCl
10 mM EDTA
pH 8.0
store at 4°C

Alkali Lysis Buffer II (100 ml)

60 ml potassium acetate 5 M
11.5 ml acetic acid conc.
store at 4°C
1) Centrifuge 1.5 ml of an overnight culture for 4 min at 4°C and 10,000 g
2) Resuspend the pellet in 100 μl chilled Alkali Lysis Buffer I
3) Add 100 μl 2% SDS and 100 μl 0.4 M NaOH and tip the tube several times
4) Incubate on ice for 5 min
5) Precipitate protein with 150 μl Alkali Lysis Buffer II and mix vigorously
6) Incubate on ice for 5 min
7) Centrifuge for 10 min at 4°C and 10,000 g and transfer supernatant into a new tube
8) Add 400 μl phenol and vortex thoroughly
9) Centrifuge for 3 min and transfer the supernatant into a new tube
10) Add 200 μl phenol and 200 μl chloroform and vortex thoroughly
11) Centrifuge for 3 min and transfer the supernatant into a new tube
12) Add 400 μl chloroform and vortex thoroughly
13) Centrifuge for 3 min and transfer the supernatant to a new tube
14) Add 1 ml 100% ethanol (-20°C) and centrifuge for 15 min at 4°C and 10,000 g
15) Without disturbing the pellet, remove the supernatant
16) Dry the pellet in a SpeedVac Concentrator
17) Add 18 μl TE and 2 μl RNase (1 mg/ml) and resuspend the pellet
18) Incubate at RT for 10 min
19) Store at -20°C

4.1.3. QIAGEN QIAprep Spin Miniprep

This protocol was used to isolate DNA of high quality that was suited for sequencing.

1) Centrifuge 3 ml overnight bacterial culture and resuspend the pellet in 250 μl P1 (chilled)
2) Add 250 μl buffer P2 and invert tube several times
3) Add 350 μl chilled buffer N3, mix vigorously, and centrifuge for 10 min
4) Apply the supernatant to a QIAprep Spin column and centrifuge for 1 min
5) Wash column with 500 μl buffer PB and centrifuge for 1 min
6) Wash column with 750 μl buffer PE and centrifuge for 1 min
7) Remove residual buffer PE by an extra spin
8) Soak column with 100 μl EB and wait for 1 min
9) Centrifuge tube for 1 min
4.1.4. QIAGEN Plasmid Maxi Kit

All the stock vector DNA preparations were made with this protocol. Furthermore, all the p120 DNA constructs used for transfection or in vitro transcription/translation experiments were purified with this protocol to ensure the same high quality of the DNA preparation throughout the experiments.

Buffer P1 (Resuspension Buffer)

- 50 mM Tris-HCl
- 10 mM EDTA
- 100 µg/ml RNase A
- pH 8.0
- Store at 4°C

Buffer P2 (Lysis Buffer)

- 200 mM NaOH
- 1% SDS

Buffer P3 (Neutralization Buffer)

- 3.0 M potassium acetate
- pH 5.5

Buffer QBT (Equilibration Buffer)

- 750 mM NaCl
- 50 mM MOPS
- 15% isopropanol
- 0.15% Triton X-100
- pH 7.0

Buffer QC (Wash Buffer)

- 1 M NaCl
- 50 mM MOPS
- 15% isopropanol
- pH 7.0

Buffer QF (Elution Buffer)

- 1.25 M NaCl
- 50 mM Tris-HCl
- 15% isopropanol
- pH 8.5

TE

- 10 mM Tris-HCl
- 1 mM EDTA
- pH 8.0

1) Inoculate 100 ml LB medium with 200 µl of a fresh overnight culture
2) Harvest the bacteria by centrifugation at 6000 g and 4 °C for 15 min
3) Resuspend the pellet in 10 ml P1 and add 10 ml P2
4) Mix gently by inverting several times and incubate at RT for 5 min
5) Add 10 ml chilled P3, mix by inverting several times and incubate on ice for 20 min
6) Centrifuge at 20,000 g for 30 min at 4°C
7) Transfer supernatant promptly into a new tube and recentrifuge at 20,000 g for 15 min at 4°C
8) Equilibrate a QIAGEN-tip 500 by applying 10 ml QBT buffer, and allow the column to run down by gravity flow
9) Apply the supernatant from step 7) to the QIAGEN-tip and allow it to enter the resin by gravity flow
10) Wash the column with 60 ml QC buffer
11) Elute DNA with 15 ml QF buffer
12) Precipitate DNA by adding 10.5 ml (0.7 volumes) RT isopropanol
13) Mix and centrifuge at 15,000 g for 30 min at 4°C
14) Carefully decant the supernatant, wash pellet twice with 70% RT ethanol
15) Centrifuge at 15,000 g for 10 min
16) Air-dry the pellet for 10 min and redissolve it in a suitable volume of TE buffer

4.1.5. M13 ss DNA Miniprep

PEG/NaCl

2.5 M NaCl

1) Centrifuge 1.5 ml of an overnight culture for 5 min and recentrifuge the supernatant
2) Add 200 µl PEG/NaCl, vortex and let stand at RT for 15 min
3) Centrifuge for 10 min and discard supernatant
4) Resuspend pellet in 100 µl TE
5) Add 100 µl phenol, vortex, and let stand at RT for 15 min
6) Vortex and centrifuge for 3 min
7) Transfer the aqueous phase to a new tube and add 500 µl ether
8) Vortex and centrifuge for 3 min
9) Remove the ether and add 10 µl NaAc 3 M pH 5.2 and 200 µl ethanol 100%
10) Allow the DNA to form a precipitate at -20°C for at least an hour and centrifuge for 20 min
11) Remove supernatant and wash the pellet with 300 µl ethanol 70%
12) Dissolve the dry pellet in 20 µl TE

4.1.6. M13 Complementation Test
1) Centrifuge 1.5 ml of an overnight culture for 5 min and recentrifuge the supernatant
2) Combine 20 μl of each supernatant, e.g. 1 with 2, 1 with 3, 2 with 3 etc.
3) Add 2 μl SDS 2% and 4 μl NaCl 0.5 M, vortex and incubate for 1 hour at 65°C
4) As a negative control, subject 40 μl of a single supernatant to the same procedure
5) Add 5 μl 5x sample buffer
6) Run the samples on a 0.7% 0.5x TBE agarose gel

4.1.7. Electroporation

A device constructed by Dr. M. Kopp was used as electroporator.

SOC

0.5% yeast extract
2% trypton
10 mM NaCl
2.5 mM KCl
10 mM MgCl₂
10 mM MgSO₄
20 mM glucose
pH 7.5, filtrate through a 0.45 μm filter and store at -20°C

ds DNA

1) Quickly thaw a 50 μl aliquot of electrocompetent cells and put them on ice
2) Add 1 to 5 μl ligation mix containing DNA and let stand on ice for 5 min
3) Pulse the cells at 3000 V and immediately add 500 μl chilled SOC
4) Incubate the cells at 37°C for 30 min
5) Streak out the bacteria on an appropriate agar plate and incubate at 37°C

ss DNA

1) Quickly thaw a 50 μl aliquot of electrocompetent cells and put them on ice
2) Add 1 to 5 μl ligation mix containing DNA and let stand on ice for 5 min
3) Pulse the cells at 3000 V and immediately add 500 μl chilled SOC
4) Incubate the cells at 37°C for 30 min
5) Add 3 ml melted 2X YT Top agar and invert several times
6) Pour the solution on an appropriate agar plate and incubate at 37°C

4.1.8. Electrocompetent Cells

IPTG

0.1 M isopropylthio-β-galactoside dissolved in ddH₂O and stored at -20°C
X-gal 20 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside, dissolved in DMF (dimethylformamide), stored at -20°C, shielded from light

XL-1 blue or BL-21 E. coli cells were used as electrocompetent cells.

1) Add 2.5 ml of a fresh overnight culture to 500 ml LB medium
2) Shake at 37°C and 180 rpm until the OD₆₀₀ reaches a value of 0.8
3) Place the cells on ice for 10 min
4) Centrifuge for 10 min at 4°C and 12,000 g and discard the supernatant
5) Resuspend the pellet in 10 ml ice-cold ddH₂O and add 240 ml ice-cold ddH₂O
6) Centrifuge for 20 min at 4°C and 12,000 g
7) Discard the supernatant, resuspend the pellet in 10 ml ice-cold ddH₂O and add 240 ml ice-cold ddH₂O
8) Centrifuge for 20 min at 4°C and 12,000 g and discard supernatant
9) Resuspend the pellet in 20 ml glycerin 10% and centrifuge for 10 min at 4°C and 12,000 g
10) Discard supernatant
11) Resuspend cells in one volume glycerin 10% and make 50 μl aliquots
12) Store at -80°C

4.1.9. Ligation

Mix the following:
- 50 ng vector (= 0.025 pmol of a 3000 bp vector)
- equimolar amount of DNA insert
- 1 μl 10x ligation buffer (Roche Molecular Biochemicals)
- 1 U T4 DNA ligase (Roche Molecular Biochemicals)
- add ddH₂O to 10 μl
- incubate at 15°C for at least 3 h

4.1.10. DNA gel electrophoresis

5x sample buffer (DNA) 5 ml glycerol 100%
- 1 ml TE 10x
- 1 ml SDS 10%
- 2.9 ml H₂O
- 100 μl 1% bromophenolblue

TBE 5x (1 l) 54 g Tris base
- 27.5 g boric acid
Agarose gels

Submarine agarose gels in the range of 0.7% to 1.5% agarose were used. They contained 0.5x TBE as buffer and ethidium bromide (0.2 μg/ml). The DNA solution was mixed with the required amount of 5x DNA sample buffer and loaded into the slots. The initial voltage of 70V was raised to 150 V after 5 to 10 min. Normally, the gel run was stopped when the bromophenolblue ran out of the gel. The result of a separation was observed under 365 nm UV-illumination.

PAGE (DNA)

For DNA PAGE, the mini-gel system from Hoefer (now Amersham Pharmacia Biotech) was used.

<table>
<thead>
<tr>
<th></th>
<th>5%</th>
<th>8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>40% Acrylamide (19:1)</td>
<td>625 μl</td>
<td>1000 μl</td>
</tr>
<tr>
<td>50% Glycerol</td>
<td>750 μl</td>
<td>750 μl</td>
</tr>
<tr>
<td>5x TBE</td>
<td>1000 μl</td>
<td>1000 μl</td>
</tr>
<tr>
<td>ddH2O</td>
<td>2625 μl</td>
<td>2250 μl</td>
</tr>
<tr>
<td>10% APS</td>
<td>30 μl</td>
<td>30 μl</td>
</tr>
<tr>
<td>TEMED</td>
<td>6 μl</td>
<td>6 μl</td>
</tr>
<tr>
<td>Total volume</td>
<td>5 ml</td>
<td>5 ml</td>
</tr>
</tbody>
</table>

For easy use, 5 ml gel solution aliquots were store at -20°C (without TEMED). Such an aliquot was then thawed, TEMED was added, and the solution was cast into a 0.7 mm gel chamber. The gel was allowed to polymerize for 45 min. As running buffer, 1x TBE was used.

As with agarose gels, the DNA sample was supplied with the required amount of 5x DNA sample buffer before it was loaded onto the gel. After 15 min, the initial voltage of 70V was raised to 150 - 200V. The run was stopped when the bromophenolblue dye ran out of the gel.

After the run, the gel was incubated in 100 ml 1xTBE containing ethidium bromide (0.2 μg/ml) for 5 min and analyzed under 365 nm UV-illumination.

4.1.11. Restriction digestion of DNA

Restriction enzymes were purchased from Boehringer Mannheim (now Roche Molecular Biochemicals) or from Life Technologies.
Normally, the desired DNA was digested with a tenfold excess of the chosen restriction enzyme in the corresponding buffer at the recommended temperature for at least an hour.

When a cloning vector containing only one restriction site was prepared, the free ends of the vector were dephosphorylated in order to prevent self-ligation of the vector. This was done using the BAP enzyme.

Mix the following:
- 1 μg digested vector DNA in a volume of 40 μl
- 10 μl BAP-buffer (10 mM ZnCl₂, 5 M Tris, pH 9.5)
- 45 μl ddH₂O
- Incubate at 65°C for 1 min
- Add 5 μl (= 5U) BAP (Life Technologies)
- Incubate at 65°C for 20 min
- Put on ice for 2 min
- Extract DNA by phenol-extraction and ethanol precipitation

### 4.1.12. Radiolabeling of DNA

I used the prime-a-gene labeling kit (Promega), a modification of the random-primed-oligolabeling-technique described by Feinberg and Vogelstein (1983).

1) Centrifuge DNA solution for 5 min at 14'000 g
2) Mix the following: 25 ng DNA
   - add ddH₂O to 16.5 μl
3) Denature DNA by incubation at 95°C for 5 min
4) Chill on ice for 3 min and spin briefly
5) Add the following:
   - 5 μl labeling buffer
   - 1 μl dNTPmix (without dCTP)
   - 1 μl BSA (10 mg/ml)
   - 1 μl (10 μCi) α³²P-dCTP (Amersham or Hartmann Analytic)
   - 0.5 μl (= 5U) Klenow polymerase (Promega)
6) Incubate at 37°C for one hour
7) Incubate at 95°C for 5 min and spin briefly
8) Cool on ice for 3 min
9) If not used immediately, store at -20°C

### 4.1.13. Isolation of DNA fragments

Several different methods were used to purify DNA fragments from agarose or polyacrylamide gels.
QIAEX (QIAGEN)

This method yields DNA of very high quality.
1)  Excise the desired DNA band from the agarose gel
2)  Weigh the gel slice. Add 3 volumes of buffer QX1 to 1 volume of gel (100 mg =
     100 pl) for DNA fragments of 100 bp to 4 kb, for larger DNA fragments use 3
     volumes QX1 plus 2 volumes ddH2O
3)  Add 5 µl QIAEX II slurry and mix
4)  Incubate at 50°C for 10 min. Mix every 2 min
5)  Centrifuge tube for 30 sec and remove supernatant
6)  Wash pellet with 500 µl buffer QX1
7)  Centrifuge tube for 30 sec and remove supernatant
8)  Wash pellet twice with 500 µl of buffer PE
9)  Centrifuge tube for 30 sec and remove supernatant
10) Air-dry pellet for 15 min
11) Add 20 µl 10 mM Tris-HCl, pH 8.5
12) Resuspend pellet and incubate for 5 min at RT (for DNA fragments > 4 kb
     incubate at 50°C for 5 min)
13) Centrifuge tube for 30 sec
14) Transfer supernatant to a new tube and store at -20°C

Low melt agarose

This method yields DNA of low quality.
1)  Under UV illumination (365 nm) cut out the desired DNA band
2)  Transfer the slice to a tube and add TE to 500 µl
3)  Let the agarose dissolve at 65°C for at least 10 min
4)  Cool the tube to RT
5)  Extract the aqueous phase twice with phenol, once with phenol/chloroform and
     once with chloroform
6)  Perform an ethanol precipitation
7)  Dissolve pellet in 20 µl TE and store at -20°C

Electroelution

This method is used to isolate small DNA fragments (< 300 bp) from polyacrylamide

gels.
1)  Rinse a dialysis bag thoroughly with 0.5 x TBE
2)  Close one end of the bag
3)  Excise DNA band with a sterilized scalpel and transfer gel slice into the bag
4)  Add 500 µl 0.5 x TBE to the bag
5) Close bag without trapping air bubbles
6) Position the bag the following way in an electrophoresis chamber:

8) Apply 150 V (10 V/cm) for 1 h
9) Reverse current for 2 min and incubate bag for 5 min in ddH₂O
10) Transfer contents of bag into tube and perform an ethanol precipitation
11) Resuspend pellet in 10 µl TE and store at -20°C

Note: If fragments smaller than 100 bp were extracted, a Spectrapor dialysis bag was used and the elution time was reduced to 30 min

Preparation of dialysis bags

1) Incubate 25 cm long pieces of dialysis bags for 10 min in 1.5 l boiling water containing 30 g NaHCO₃ and 0.56 g EDTA x 2H₂O
2) Thoroughly rinse the bags with ddH₂O
3) Cut the 25 cm pieces into 4 equal parts
4) Autoclave bags in 2 l ddH₂O containing 0.74 g EDTA x 2H₂O
5) Store the bags at 4°C

4.1.14. DNA ethanol precipitation

3M NaAc

3M Na acetate
adjust pH to 5.2 with glacial acetic acid

1) Add 0.1 volumes of 3M NaAc to DNA containing solution
2) Add 2 volumes of ethanol (-20°C)
3) Mix thoroughly
4) Centrifuge large DNA fragments immediately, smaller ones need to be incubated at -20°C for two h or overnight
5) Centrifuge for 30 min at 4°C
6) Remove supernatant with a drawn-out pasteur pipette
7) Wash pellet with 300 µl ethanol (70%, -20°C) and centrifuge for 5 min
8) Remove supernatant carefully
9) Dry pellet in a "SpeedVac Concentrator" for 5 min
10) Resuspend pellet in the desired volume of TE and store solution at -20°C
4.1.15. Sequencing

The T7 Sequenase V2.0 kit from Amersham was used for sequencing. This kit uses the dideoxy chain termination method developed by Sanger et al. (1977).

**Denaturation**

When ds DNA was used as template, it had to be denatured first:

1) 1 µg ds DNA in 20 µl
2) Add 2 µl of the following mixture (49 µl NaOH 2M, 1 µl EDTA 0.1M)
3) Incubate at 37°C for 5 min
4) Cool on ice for 3 min
5) Add: 7 µl ddH₂O, 7 µl NaAc 3M pH 5.2, 75 µl ice-cold ethanol
6) Incubate at -20°C for an hour
7) Centrifuge at 4°C and 14,000 g for 30 min
8) Wash pellet with ice-cold ethanol 70%
9) Centrifuge at 4°C and 14,000 g for 15 min
10) Resuspend the dry pellet in 6 µl ddH₂O

When M13 ss DNA was sequenced, 3 µl of a M13 miniprep were diluted with 3 µl ddH₂O

**Annealing**

1) Add 2 µl annealing buffer and 2 µl (= 2 pmol) sequencing primer to the 6 µl of denatured or ss DNA
2) Incubate at 65°C for 5 min
3) Cool to RT over 30 min
4) Centrifuge briefly

**Sequencing reaction**

1) Add the following: 1 µl DTT, 2 µl labeling mix 1:10 (long runs 1:5), 0.5 µl (α-³⁵) dATP (1000 Ci/mmol) (Amersham or Hartmann Analytic)
2) Add 2 µl Sequenase (diluted 1+7 in sequenase dilution buffer)
3) After the course of 1 min (long runs 5 min), add 3.5 µl aliquots of the above solution to each of 4 tubes containing either 2.5 µl ddG, ddA, ddT or ddC termination mixes and preincubate at 37°C
4) Let the 4 tubes remain at 37°C for 5 min
5) Add 4 µl stop solution to each of the 4 tubes
6) Store at -20°C if not used immediately
Sequencing gel

The Sequencing Gel Electrophoresis System S2 from Life Technologies was used to run sequencing gels.

1) Wash 2 glass plates, two spacers (0.4 mm) and combs (shark tooth) thoroughly with 2% SDS, rinse with ddH₂O and let stand to dry
2) If necessary, silanize the glass plates prior to step 1) as follows:
   spread 1 ml dimethyldichlorosilane (4% in CHCl₃) all over a glass plate. After evaporation of the chloroform, rinse the plate thoroughly with ddH₂O. Repeat step 1)
3) Assemble glass plates and spacers and seal with a rubber
4) Prepare 60 ml of a 6% (standard) or 5% (long runs) gel solution
   | Percentage | 5% | 6% |
   | Acrylamide | 40% | 7.5 ml | 9.0 ml |
   | Urea (to 6M) | 25.2 g | 25.2 g |
   | 5x TBE buffer | 6 ml | 6 ml |
   add ddH₂O to a final volume of 60 ml
5) Stir until urea has dissolved
6) Induce polymerization by addition of 360 µl APS 10% and 36 µl TEMED
7) Pour gel with a 60 ml syringe fitted to a 0.45 µm filter
8) Insert comb with the continuous site about 5 mm between the two glass plates
9) Add weights to the gel and let polymerize for at least 3 hours
10) After polymerization, remove rubber and comb
11) Place gel in electrophoresis device and fill electrode chambers with 0.5 x TBE
12) Insert comb into the gel until teeth penetrate the gel about 2 mm
13) Prerun the gel at 60 W for 30 min
14) Denature samples at 77°C for 4 min
15) Centrifuge briefly
16) Load 4 µl (ss DNA: 2 µl) per lane
17) Run the gel at 60 W
18) Stop short runs after the bromophenolblue dye leaves the gel, intermediate runs after the xylene cyanol dye. Obtain a long run by reloading the samples after an intermediate run and stopping the gel after the xylene cyanol dye of the second series runs out of the gel
19) Remove comb, spacers and one glass plate
20) Fix gel five times for 3 min in 12% methanol, 10% acetic acid
21) Remove fixing solution and place a Whatman 3MM paper onto the gel
22) Dry gel for 1 hour at 80°C under vacuum
23) Expose gel to a BioMax MR film (Kodak) overnight
4.1.16. Bacterial cultures

LB medium (1 l)
- 10 g bacto-trypton (Life Technologies)
- 10 g NaCl
- 5 g bacto-yeast extract (Life Technologies)
- add ddH₂O to 1 l, pH 7.5, autoclave

LB plates
- 1 l LB-medium
- 15 g bacto-agar (Life Technologies)
- autoclave
- let stand at RT until dry, then store at 4°C

LB amp plates
- as LB plates, but addition of 2 ml 500x Amp (50 mg/ml) after cooling to 55°C

LB amp glucose plates
- as LB plates, but addition of 2 ml 500x Amp and 20 mM glucose after cooling to 55°C

LB kan plates
- as LB plates, but addition of 1 ml 1000x Kan (10 mg/ml) after cooling to 55°C

LB top agar
- 1 l LB medium
- 7 g bacto-agar
- autoclave

Overnight culture

1) Transfer an inoculum of a bacterial stock or a single colony from a agar plate to a 15 ml tube containing 3 ml LB medium and the required antibiotic
2) Shake overnight at 37°C and 280 rpm

Bacterial stock

1) Add 250 μl glycerol 90% to 750 μl of a fresh overnight culture and mix
2) Store at -80°C

4.2. RNA techniques

4.2.1. Isolation of total RNA and mRNA
Total RNA used for the preparation of the tumor blot was isolated by T. Schenker using the SDS/Proteinase K method. mRNA used for the quantitative PCR approach was first isolated by T. Schenker and D. Belluccio from a 12 week old human embryo using the QIAGEN RNeasy Midi kit and then purified using the Oligotex purification system from QIAGEN.

For isolation of total RNA from cells grown in culture, the QIAGEN RNeasy midi or mini kit in combination with the QIAshredder was used.

4.2.2. Northern blot hybridizing

<table>
<thead>
<tr>
<th>Component</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPC treated ddH₂O</td>
<td>Add 0.1% diethylpyrocarbonate (DEPS) to ddH₂O, shake intensely and incubate at 37°C over night. Autoclave to destroy DEPC. Every aqueous solution used for RNA experiments was made with such water.</td>
</tr>
<tr>
<td>Running buffer 10x (1 l)</td>
<td>20.93 g MOPS</td>
</tr>
<tr>
<td></td>
<td>1.86 g EDTA</td>
</tr>
<tr>
<td>Hybridization solution (50 ml)</td>
<td>7 ml ddH₂O</td>
</tr>
<tr>
<td></td>
<td>25 ml formamide</td>
</tr>
<tr>
<td></td>
<td>12.5 ml SSPE 20x</td>
</tr>
<tr>
<td></td>
<td>0.5 ml SDS 10%</td>
</tr>
<tr>
<td></td>
<td>5 ml Denhardt 50x</td>
</tr>
<tr>
<td></td>
<td>0.5 ml salmon sperm DNA (heat for 10 min at 95°C and then keep on ice for 10 min. Prior to storage at -20°C, press solution through a syringe for a couple of times)</td>
</tr>
<tr>
<td>Denhardt 50x (1 l)</td>
<td>10 g Ficoll</td>
</tr>
<tr>
<td></td>
<td>10 g polyvinylpyrrolidone</td>
</tr>
<tr>
<td></td>
<td>10 g BSA</td>
</tr>
<tr>
<td>RNA sample buffer 1.33x</td>
<td>6740 μl formamide</td>
</tr>
<tr>
<td></td>
<td>2160 μl formaldehyde 37%</td>
</tr>
<tr>
<td></td>
<td>1000 μl 10x buffer</td>
</tr>
<tr>
<td></td>
<td>40 μl EDTA 0.5 M pH 7.0</td>
</tr>
<tr>
<td></td>
<td>570 μl glycerol 87%</td>
</tr>
<tr>
<td></td>
<td>100 μl bromophenolblue 1%</td>
</tr>
<tr>
<td></td>
<td>100 μl SDS 10%</td>
</tr>
</tbody>
</table>
SSC (20x) 3 M NaCl
0.3 M Na3 citrate
pH 7.0

Wash solution I 450 ml ddH2O
50 ml SSC 20x
5 ml SDS 10%

Wash solution II 475 ml ddH2O
25 ml SSC 20x
5 ml SDS 10%

1) For a 1% formaldehyde agarose gel, mix the following:
   0.8 g agarose
   60 ml ddH2O
   8 ml 10x running buffer
   100 µl EDTA 0.5 M pH 8.0
2) Heat until agarose has dissolved and the let gel mix cool down a little bit
3) Add 6.5 ml formaldehyde
4) Add ddH2O to 80 ml
5) Pour gel and let polymerize
6) Mix 10 µl (=10 to 20 µg) RNA with 30 µl RNA sample buffer and add 1 µl
   (=0.02 µg) ethidium bromide
7) Incubate at 65°C or 10 min
8) Apply samples to gel and let it run with 1x running buffer
   The initial voltage of 70V was raised to 150V after 15 min
   Exchange the buffer of the two electrodes every 15 min
9) Transfer RNA contained in the gel to a GeneScreen nylon membrane (NEN) with
   the VacuGene vacuum blotting system by Amersham Pharmacia Biotech
10) Crosslink RNA present on the membrane twice with a UV-crosslinker by
    Stratagene (autocrosslink program)
11) Under UV-illumination, mark positions of the 18S- and 28S-rRNA
12) Wet blot in 6xSSC
13) Transfer blot into tray containing hybridization solution
14) Incubate at 42°C for 2 hours
15) Add radioactively labeled probe and mix
16) Incubate at 42°C for at least 12 h
17) Carefully transfer radioactive solution into a special container for radioactive
    waste
18) Shake blot for 30 min in wash solution I at RT
19) Wash blot twice for 30 min in wash solution II at 50°C
20) Dry blot on a Whatman 3MM paper
21) Wrap a saran foil around the blot
22) Expose to BioMax MS (Kodak) or X-OMAT (Kodak) film overnight at -80°C

4.2.3. cDNA synthesis

The cDNA synthesis kit from Roche Molecular Biochemicals was used.

1) Mix the following:
   - 5 μl cDNA synthesis primer\(^1\) (2 pmol/μl)
   - 5 μl poly(A) mRNA (250 ng)

2) Incubate at 65°C for 5 min
3) Cool to RT for 5 min
4) Add the following:
   - 4 μl 5x buffer
   - 2 μl dNTP mix
   - 1 μl RNasin
   - 1 μl AMV
   - 2 μl ddH₂O

5) Incubate for 1 h at 42°C
6) Stop reaction by addition of 1.7 μl EDTA 0.2M

\(^1\) For production of the cDNA used in the quantitative PCR experiment, primer L4061 was used.

4.3. PCR techniques

The chemicals used for the PCR were purchased from Perkin Elmer or Boehringer Mannheim. Three different thermostable enzymes were used: For standard applications, TaqPlus (Stratagene) was used. In cases where excellent amplification was essential, the AmpliTaq (Perkin Elmer) enzyme was used. Since both these enzymes to not possess a proofreading capacity, the Pfu (Stratagene) enzyme was applied in reactions where the PCR was used to facilitate the cloning of DNA fragments.

4.3.1. 3'RACE

The Marathon cDNA Amplification Kit from CLONTECH was used for performing a 3'RACE. The protocol recommended by the supplier was exactly followed. In brief, the following steps were performed: The first strand cDNA was synthesized from 1 μg poly(A) mRNA derived from IMR90 human lung fibroblasts and with the supplied Marathon cDNA synthesis primer (Fig. 4.1). After second strand synthesis, a blunt ended ds cDNA molecule was obtained. A so-called cDNA adaptor was then ligated to
the ds cDNA. This cDNA was diluted to a total volume of 500 µl. The resulting molecule could be used as template in a PCR. Two different sets of primers (GSP1 and AP1, GSP2 and AP1) were used for two different PCRs.

Marathon cDNA Synthesis Primer (52-mer)
5' TGGTAAATTGCAACCCGCGCGC(T)N,N 3'

Marathon cDNA Adaptor
5' CTAACGGGACTCACTATAGGCTCTAGGCGCGCGCGCGAGG 3'

Adaptor primer 1 (AP1, 27-mer)
5' CCAACCTAAAGCACTCACTATAGGCG 3'

GSPl (31-mer)
5' TTCTAGAATTCAGCGGCCGC(T)N,N 3'

GSP2 (26-mer)
5' TGGCTACCTAAGCATGCTCTAGGCG 3'

Clone Pfu buffer #1
2 µl

Clone Pfu buffer #2
2 µl

Pfu polymerase (5 U/µl)
0.2 µl

1) Mix all ingredients except the Pfu polymerase in a 0.5 ml microcentrifuge tube
2) Overlay with 50 µl mineral oil
3) Incubate tube for 1 min at 95°C
4) Add Pfu polymerase
5) Perform the following cycles:
   30x: 45 sec 95°C, 45 sec 54°C, 2 min 72°C
   1x 10 min 72°C
6) Clone the DNA bands obtained from these two PCRs
4.3.2. Quantitative PCR

1) Prepare the following mixes:
   - 2 µl 10x buffer
   - 2 µl MgCl₂ (25 mM)
   - 2 µl dNTP mix (2.5 mM each)
   - 1 µl primer L3115 (10 µM)
   - 1 µl primer U2780 (10 µM)
   - 0.5 µl cDNA
   - desired number of standard plasmid
   - add ddH₂O to 19.8 µl
2) Overlay with 30 µl mineral oil
3) Incubate at 94°C for 1 min
4) Add 0.2 µl AmpliTaq (1 U)
5) Perform the following PCR:
   - 35X: 1 min 94°C, 1 min 50°C, 45 sec 72°C; 1x 10 min 72°C
6) Apply samples to an 8% DNA PAGE

4.3.3. Genomic PCR

1) Prepare the following mixture:
   - 1 µg genomic DNA (isolated by Dr. M. Kopp)
   - 2 µl Low Salt Buffer
   - 0.5 µl dNTP Mix (2.5 mM each)
   - 1 µl primer L4061 (10 µM)
   - 1 µl primer U3848 (10 µM)
   - 13.3 µl ddH₂O
2) Overlay with 30 µl mineral oil
3) Incubate at 94°C for 1 min
4) Add 0.2 µl TaqPlus (1 U)
5) Perform the following PCR:
   - 35X: 1 min 94°C, 1 min 56°C, 1 min 72°C
6) Remove mineral oil with chloroform
7) Add 5 µl 5x DNA Sample Buffer to aqueous phase
8) Load samples onto an 8% DNA PAGE gel

4.3.4. Primers

Oligonucleotides used for sequencing or as primers in a PCR were designed by using the OLIGO program and synthesized by Microsynth.
The following primers were used for PCR applications mentioned in this chapter (Numbers in parenthesis correspond to Fig. 2.2.).

- **L4061** (cDNA synthesis, genomic PCR)
  5' (4083) GGGTGGGGCTATTCATCACAGTA 3' (4061)
- **L3115** (quantitative PCR)
  5' (3137) TTCTCTAAGCTGGTATGAAGTTC 3' (3115)
- **U2780** (quantitative PCR)
  5' (2780) CTTTGGTGAAGCTACAGGTGA 3' (2802)
- **U3848** (=GSP1, 3'RACE, genomic PCR)
  5' (3848) TGAAGCTTAAGTGCATAGATGGTCA 3' (3878)
- **U3979** (=GSP2, 3'RACE)
  5' (3979) GGTCTACCTAACCTAAGCTGCTTC 3' (4002)

### 4.4. Protein techniques

#### 4.4.1. Preparation of membrane protein extracts

**Buffer 1**

137 mM NaCl  
20 mM Hepes  
2 mM MgCl₂  
1 mM EDTA  
pH 7.2

**Lysis buffer**

20 mM Tris base  
137 mM NaCl  
1 mM MgCl₂  
10% Triton X-100  
add fresh:  
0.1 mM DTT  
1 mM PMSF  
10 μg/ml leupeptin  
pH 7.4

1) Wash cells twice with 1x PBS  
2) Resuspend cells in buffer 1 (1ml/10⁷ cells)  
3) Freeze thaw cells three times in liquid N₂  
4) Centrifuge at 14'000 g for 2 min  
5) Remove supernatant  
6) Resuspend cell debris in lysis buffer (12.5 µl/10⁶ cells)  
7) Incubate on ice for 30 min  
8) Vortex and centrifuge for 2 min at 14'000 g
9) Transfer supernatant to new tube
10) Store at -20°C

4.4.2. Protein determination

The method described by Lowry et al. (1951) was used.

Reagent A: 2% Na₂CO₃ in 0.1 M NaOH
Reagent B: 1% CuSO₄
Reagent C: 2% Na or K Tartrate
Reagent D: 98 ml A + 1 ml B + 1 ml C
Reagent E: 50% Folin-Ciocalteu phenol reagent in H₂O

Mix the following:
- 0.2 ml protein solution (5 - 100 μg protein)
- 1 ml reagent D, mix, let stand at RT for 10 min
- Add 100 μl reagent E, mix and wait for 30 min, measure OD at 750 nm
- A bovine serum albumin solution was used as standard

4.4.3. Protein PAGE

Two different systems were used. One was the SDS-PAGE system developed by Laemmli (1970), the second was developed by Schagger and von Jagow (1987) and is especially suited for separation of peptides in the range of 2 to 30 kD.

Before the samples containing the required amount of protein sample buffer were loaded onto the gel, they were incubated at 95°C for 4 min.

SDS-PAGE

SDS gel buffer 10x

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>30.3 g</td>
</tr>
<tr>
<td>glycine</td>
<td>144.2 g</td>
</tr>
<tr>
<td>SDS</td>
<td>20 g</td>
</tr>
</tbody>
</table>

Add ddH₂O to 1 l

Protein sample buffer 2x (10 ml)

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDS 10%</td>
<td>2 ml</td>
</tr>
<tr>
<td>glycerol 87%</td>
<td>2.3 ml</td>
</tr>
<tr>
<td>Tris 1 M pH 6.8</td>
<td>1 ml</td>
</tr>
<tr>
<td>β-MeSH</td>
<td>200 μl</td>
</tr>
<tr>
<td>EDTA 0.5 M pH 8.0</td>
<td>80 μl</td>
</tr>
<tr>
<td>bromophenolblue 1%</td>
<td>1000 μl</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>3450 μl</td>
</tr>
</tbody>
</table>
Staining solution  
25% isopropanol  
10% acetic acid  
0.1% coomassie brilliant blue R-250

Destaining solution  
25% methanol  
10% acetic acid

The following gel mixtures were used for standard applications

<table>
<thead>
<tr>
<th></th>
<th>5%</th>
<th>10%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide 30% (37.5:1)</td>
<td>500 µl</td>
<td>1670 µl</td>
<td>2500 µl</td>
</tr>
<tr>
<td>Tris 1 M pH 6.8</td>
<td>375 µl</td>
<td>1250 µl</td>
<td>1250 µl</td>
</tr>
<tr>
<td>Tris 1.5 M pH 8.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDS 10%</td>
<td>30 µl</td>
<td>50 µl</td>
<td>50 µl</td>
</tr>
<tr>
<td>EDTA 0.5 M pH 8.0</td>
<td>12 µl</td>
<td>20 µl</td>
<td>20 µl</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>2050 µl</td>
<td>1960 µl</td>
<td>1960 µl</td>
</tr>
<tr>
<td>APS 10%</td>
<td>30 µl</td>
<td>50 µl</td>
<td>50 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td>3 µl</td>
<td>5 µl</td>
<td>5 µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>3 ml</td>
<td>5 ml</td>
<td>5 ml</td>
</tr>
</tbody>
</table>

Aliquots of the above solutions were stored at 4°C (without 10% APS and TEMED). 10% APS and TEMED were added just before the gel was polymerized.

A gradient gel system was used for the separation of protein extracts enriched for integral membrane proteins.

<table>
<thead>
<tr>
<th>Gel type</th>
<th>Stack</th>
<th>Run</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>3%</td>
<td>3%</td>
<td>10%</td>
</tr>
<tr>
<td>Acrylamide 30% (37.5:1)</td>
<td>1.2 ml</td>
<td>1.2 ml</td>
<td>4 ml</td>
</tr>
<tr>
<td>Tris 1.5 M pH 8.8</td>
<td>2.4 ml</td>
<td>2.4 ml</td>
<td></td>
</tr>
<tr>
<td>Tris 1.5 M pH 6.8</td>
<td>1.5 ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDS 10%</td>
<td>240 µl</td>
<td>240 µl</td>
<td>240 µl</td>
</tr>
<tr>
<td>EDTA 0.5 M pH 8.0</td>
<td>48 µl</td>
<td>48 µl</td>
<td>48 µl</td>
</tr>
<tr>
<td>Glycerol 80%</td>
<td>1.354 ml</td>
<td>1.354 ml</td>
<td></td>
</tr>
<tr>
<td>ddH₂O</td>
<td>8.9 ml</td>
<td>6.8 ml</td>
<td>4 ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>7 µl</td>
<td>7 µl</td>
<td>7 µl</td>
</tr>
<tr>
<td>APS 10%</td>
<td>120 µl</td>
<td>120 µl</td>
<td>120 µl</td>
</tr>
<tr>
<td>Total volume</td>
<td>12 ml</td>
<td>12 ml</td>
<td>12 ml</td>
</tr>
</tbody>
</table>
These gels were run with 1x SDS gel buffer as running buffer. The initial voltage of 70V was raised to 150V after 15 min. The run was stopped after the bromophenolblue dye had ran out of the gel. If the gels were not used for Western blotting experiments, they were incubated in the staining solution for 30 to 60 min. Excess dye was then washed off by incubation in destaining solution until no background was observable. The gel was dried between pre-wet cellophan membranes for 1 h at 80°C.

**Tricine SDS PAGE**

As mentioned above, the Tricine SDS PAGE system was used to separate small peptides in the range of 2 to 30 kD.

<table>
<thead>
<tr>
<th>Gel type</th>
<th>Run</th>
<th>Spacer</th>
<th>Stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>16.5%</td>
<td>10%</td>
<td>4%</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>2000 µl</td>
<td>210 µl</td>
<td>160 µl</td>
</tr>
<tr>
<td>Gel buffer</td>
<td>2000 µl</td>
<td>330 µl</td>
<td>500 µl</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>1000 µl</td>
<td>460 µl</td>
<td>1340 µl</td>
</tr>
<tr>
<td>Glycerol 87%</td>
<td>1000 µl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APS 10%</td>
<td>30 µl</td>
<td>5 µl</td>
<td>16 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td>3 µl</td>
<td>0.5 µl</td>
<td>1.6 µl</td>
</tr>
</tbody>
</table>

**Acrylamide solution**
- 49.5% acrylamide
- 3% bis-acrylamide

**Anode buffer (10x)**
- 2 M Tris HCl
- pH 8.9

**Cathode buffer (10x)**
- 1 M Tris HCl
- 1 M tricine
- 1% SDS
- pH 8.3

**Gel buffer**
- 3 M Tris HCl
- 0.3% SDS
- pH 8.45

**Protein sample buffer (Schagger) 2x**
- 5 ml SDS 16%
- 2.75 ml glycerol 87%
94

1 ml Tris 1 M pH 6.8
400 µl β-MeSH
200 µl bromophenolblue 1%
650 µl ddH₂O

Staining solution
50% methanol
10% acetic acid
0.025% coomassie brilliant blue G-250

Destaining solution
25% methanol
10% acetic acid

The following protocol was used to run these gels:

1) Assemble gel apparatus
2) Cast gel in this order: running, spacer, and stacking gel
3) Incubate protein samples containing protein sample buffer (Schagger) for 4 min at 95°C
4) Fill cathode chamber with 1x cathode buffer, anode chamber with 1x anode buffer
5) Load samples onto gel
6) Start gel run with a voltage of 40V for 1 h
7) Raise voltage to 80V and let gel run until bromophenolblue dye runs out of the gel (up to 16 h)
8) Stain gel for 2 h in staining solution
9) Incubate gel for 4 h in destaining solution
10) Dry gel for 1 h at 80°C
11) Expose gel to BioMax MR (Kodak) film, X-OMAT (Kodak) film, or to a phosphorimager screen.

The phosphorimager screen was analyzed using a STORM 840 phosphorimager (Molecular Dynamics).

4.4.4. Western blotting

Transfer buffer 10x
30.3 g Tris HCl
144.2 g glycine
add ddH₂O to 1 l

Ponceau red
0.2 % Ponceau S in 3% trichloroacetic acid

TBS
150 mM NaCl
50 mM Tris HCl  
ph 7.5

BSA/TBS  
3% BSA in TBS

Color reaction buffer (CRB)  
100 mM NaCl  
100 mM Tris HCl  
5 mM MgCl₂  
ph 9.5

Nitroblue tetrazolium (NBT)  
500 mg in 7.5 ml ddH₂O, 7.5 ml DMF

5-Brom-4-chlor-3-indolyl phosphate (BCIP) 500 mg in 10 ml ddH₂O, 5 ml DMF

1) Depending on the type of polyacrylamide gel used, two different protocols were employed
   a) Normal mini SDS polyacrylamide gels were employed in the TRANS-BLOT SD semi-dry Transfer Cell (BioRad). The following sandwich was assembled (starting from the anode), Whatman paper, nitrocellulose membrane (PROTRAN BA85, Schleicher&Schuell), polyacrylamide gel, and Whatman paper (all soaked in 1x transfer buffer containing 20% methanol). Transfer proteins for 25 min (0.75 mm thick gels) or 40 min (1.5 mm thick gels) at 20 V (350 - 200 mA).
   b) Proteins separated on larger gels were subjected to a different method. The same sandwich as above was assembled in a submarine fashion in a tray containing 1 x transfer buffer and 20% methanol. The proteins were then transferred overnight at 4°C and 20 V (200 mA) in a tank containing 1 x transfer buffer and 20% methanol.

2) After the transfer, disassemble sandwich

3) Control transfer efficiency by staining of the membrane for 10 min with Ponceau red

4) Wash blot with water until individual bands become visible

5) Mark positions of the protein standard and the outline of the individual lanes

6) Incubate membrane with BSA/TBS for 2 h at RT

7) Add appropriate amount of primary antibody (see below *¹) and incubate for 2 h at RT

8) Discard solution and wash blot several times thoroughly with TBS

9) Incubate blot in BSA/TBS and add secondary antibody (see below *²)

10) Incubate for 1 h at RT
11) Wash blot thoroughly three times with TBS and once with CRB
12) Incubate blot in CRB and add NBT (50 μl/5 ml CRB) and BCIP (25 μl/5 ml CRB)
13) Stop reaction by washing blot with water.

*1 - The polyclonal antibody raised against p120 was used at a 1:50 dilution.
*2 - Sometimes a goat anti-GST antibody was used as primary antibody at the dilution recommended by the supplier.

4.4.5. Bacterial protein expression (GST system)

The GST Gene Fusion System (Pharmacia Biotech) was used for the production of bacterially expressed fusion proteins. For this purpose a cDNA fragment of p120 (nucleotides 3361 - 3888, corresponding to amino acids 1063 - 1119) was cloned into the bacterial GST expression vector pGEX-5X-1. This plasmid was then transfected into electrocompetent BL-21 cells. In a first set of experiments, the exact conditions for production of the GST-p120 fusion protein were optimized.

1) Add 120 μl of a fresh overnight culture to 12 ml LB medium, containing 24 μl 500x ampicillin and 120 μl 2 M glucose
2) Incubate at 37°C and 280 rpm in a 50 ml tube
3) When the culture reaches an OD600 of 1.0, add 12 μl 0.1 M IPTG to induce expression of the fusion protein
4) Incubate at 37°C and 280 rpm for two hours
5) Pellet cells by centrifugation for 5 min at 3'000 g
6) Resuspend cells in 250 μl chilled PBS
7) Lyse cells by sonification (1 min, 0.5 sec pulses)
8) Centrifuge for 10 min
9) Discard supernatant
10) Resuspend pellet in 50 μl PBS and 50 μl 2x protein sample buffer and store at -20°C

For the production of the amount of fusion protein needed for coupling to an affinity column, the above protocol was slightly modified.

1) Add 2.5 ml of a fresh overnight culture to 250 ml LB medium, containing 500 μl 500x ampicillin and 2.5 ml 2 M glucose
2) Incubate at 37°C and 280 rpm in a 1 l flask
3) When the culture reaches an OD_{600} of 1.0, add 250 μl 0.1 M IPTG to induce expression of the fusion protein.
4) Incubate at 37°C and 280 rpm for two hours.
5) Pellet cells by centrifugation for 5 min at 3000 g.
6) Resuspend cells in 10 ml MOPS (0.1 M, pH 7.5) and make 20 aliquots at 500 μl.
7) Lyse cells by sonification (2x 90 sec, 0.5 sec pulses).
8) Combine sonicated aliquots and centrifuge for 10 min.
9) Discard supernatant.
10) Wash pellet three times with 2% NP-40/MOPS.
11) Wash pellet three times with MOPS.
12) Dissolve pellet in SDS solution (final concentration 2% SDS).
13) Store at -20°C.

Note: According to the Lowry test, 15 mg fusion protein were obtained.

### 4.4.6. Affinity column

An affinity column was made in order to purify polyclonal antibodies.

**PBS 1x (1L)**

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>8 g</td>
</tr>
<tr>
<td>KCl</td>
<td>0.2 g</td>
</tr>
<tr>
<td>Na₂HPO₄x12H₂O</td>
<td>3.63 g</td>
</tr>
<tr>
<td>KH₂PO₄</td>
<td>0.24 g</td>
</tr>
<tr>
<td>pH 7.5</td>
<td></td>
</tr>
</tbody>
</table>

1) Add 2 ml GST-p120 fusion protein (~ 7 mg protein) to 1 ml Affi-Gel 10 suspension (BioRad).
2) Shake gently for 4 h at 4°C.
3) Add 1 ml Affi-Gel 10 suspension.
4) Shake gently at 4°C overnight.
5) Transfer the coupled matrix to a column.
6) Wash column with 20 ml PBS.
7) Store column at 4°C.

### 4.4.7. Affinity purification of polyclonal antibodies

1) Apply 1.5 ml antiserum (T3 897) to the affinity column.
2) Wash column with 2 x 20 ml PBS.
3) Elute antibody with 10 ml glycine buffer (0.1 M, pH 2.5).
4) Collect fractions of 0.5 ml each in tubes that already contain 33 μl neutralization solution (1.5 M Tris HCl, 15 mg BSA/ml, pH 8.8).
5) Analyze individual fractions in an ELISA.
6) Wash column with 20 ml glycine buffer and regenerate it with 30 ml PBS.
4.4.8. ELISA

The ELISA was used to assess the quality of our antisera raised against p120. The actual measurement was done with a Multiskan MCC/340 96-well plate reader from Titertek.

1) Apply 100 μl TBS containing 25 μg/ml p120 peptide (amino acids 1105 - 1119) to all wells of an 96-well plate
2) Incubate overnight at 4°C
3) Discard the peptide containing solution and wash plate three times with wash solution (0.15 M NaCl, 0.05% Tween-10 (w/v))
4) Apply 100 μl primary antibody solution in two-fold dilution steps to the wells.
5) Incubate at RT for 4 h
6) Repeat step 3)
7) Add 100 μl goat anti-rabbit IgG antibody (conjugated to BAP) solution at the concentration recommended by the supplier to each well.
8) Incubate at RT for 1 h
9) Repeat step 3)
10) Add 100 μl color reaction buffer (1 mg/ml p-nitrophenyl phosphate, 1 M diethanolamine, pH 9.8) to each well
11) Read absorbance at 405 nm

4.4.9. In vitro transcription/translation

The TNT Coupled Reticulocyte Lysate or the TNT Coupled Wheat Germ Extract System from Promega was used for producing peptides encoded by the constructs 120f, 120s, 122EB, and 91EB. As expression vector, pcDNA3.1 was used.

Reticulocyte system

1) Remove reagents from storage at -70°C. Rapidly thaw the TNT Reticulocyte Lysate by hand warming and place on ice
2) Mix the following:
   • 12.5 μl TNT Reticulocyte Lysate
   • 1 μl TNT Reaction Buffer
   • 0.5 μl TNT RNA polymerase
   • 0.5 μl amino acid mixture minus methionine (1mM)
   • 2 μl 35S-methionine (1000 Ci/mmol) at 10 mCi/ml (Hartmann Analytic)
   • 0.5 μl RNasin Ribonuclease Inhibitor (40 U/μl)
• 0.5 μg DNA template
• nuclease-free H₂O to 25 μl

3) Incubate at 30°C for 60 min
4) Store at -20°C

Wheat germ system

1) Remove reagents from storage and place them on ice
2) Mix the following:
   • 12.5 μl TNT Wheat Germ Extract
   • 1 μl TNT Reaction buffer
   • 0.5 μl TNT RNA polymerase
   • 2 μl amino acid mixture minus methionine (1mM)
   • 1.25 μl 35S-methionine (1’000 Ci/mmol) at 10 mCi/ml (Hartmann Analytic)
   • 0.5 μl RNasin Ribonuclease Inhibitor (40 U/μl)
   • 1.875 μl potassium acetate (1 M)
   • 0.5 pg DNA template
   • nuclease-free H₂O to 25 μl
3) Incubate at 30°C for 60 min
4) Store at -20°C

In the case of the two constructs 120f and 120s, the samples were then run on 10% SDS polyacrylamide gels. The two short constructs 122EB and 91EB were run on a 16.5% Tricine SDS polyacrylamide gel. Both gel types where then dried and exposed to BioMAX MR film (Kodak) or to a Phosphorimager Screen.

4.4.10. Immunofluorescence

Paraformaldehyde 4% (100 ml)

Add 4 g paraformaldehyde to 50 ml ddH₂O, heat to 60°C and add a few drops of 1 M NaOH until dissolved, let cool to RT and add 50 ml 2x PBS. Prepare fresh every day.

Mounting solution

10 mg/ml n-propyl gallate
90% glycerol
PBS, pH 7.4

1) Grow cells to the desired cell density
2) Aspirate growth medium and wash cells three times with PBS
3) Fix cells with 4% paraformaldehyde for 10 min at RT
4) Rinse fixed cells gently three times with PBS
5) Permeabilize cells with 0.2% Triton X-100 for 10 min at RT
6) Rinse cells carefully four times with PBS
7) Incubate cells for 2 hours with BSA/TBS at RT
8) Add primary antibody in the recommended dilution
9) Incubate for 2 h at RT
10) Wash slides carefully three times with 1% PBS over 5 min
11) Add secondary antibody (FITC-labeled goat anti rabbit IgG) 1:50
12) Incubate for 1 h at RT in the dark
13) Wash cells three times with 1% Triton X-100 in PBS
14) Mount slides with a drop of mounting solution and a cover slip
15) Analyze slides with a microscope equipped with epifluorescence optics (Zeiss Axiovert)

4.5. Cell culture

Every cell line used in this thesis was cultivated in a 37°C incubator at more than 95% humidity and under an atmosphere containing 5% CO₂. As medium, Dulbecco's modified Eagle medium (DMEM) supplemented with 100 U penicillin/ml and 100 µg streptomycin/ml and 9% FCS was used. All these solutions were purchased from Life Technologies.

4.5.1. Splitting of cells

1) Remove medium from cell layer
2) Wash cells once with 1xPBS
3) Add trypsin-EDTA solution (1 ml /25 cm²)
4) Keep in incubator until cells detach from the surface of the flask
5) Add medium to 10-15 ml and centrifuge at 200 g for 5 min
6) Carefully remove supernatant
7a) Resuspend cells in medium and distribute to new flasks (normally, cells were split 1:3 or 1:5)
7b) If cells will be frozen, resuspend cells in 1 ml freezing medium (90% FCS, 10% DMSO)
8) Transfer cells to a cryotube (Nunc)
9) Put tube into a tight sealed styro-foam box filled with cotton wool
10) Keep box in an -80°C freezer for one day
11) Transfer tube to a tank containing liquid nitrogen
4.5.2. Transfection of eukaryotic cells

Lipofectin (Life Technologies) (pcDNA3.1 constructs)

1) 24 h before transfection, seed $10^5$ cells in a 60 mm dish
   On the day of transfection, cells should be 30-50% confluent
2) Prepare the following solutions:
   Solution A: For each transfection dilute 2 µg DNA into 100 µl Opti-MEM 1 medium (Life Technologies)
   Solution B: For each transfection dilute 6 µl Lipofectin into 100 µl Opti-MEM 1 medium and incubate at RT for 45 min
3) Combine the two solutions, mix gently, and incubate at RT for 30 min
4) Wash cells once with 2 ml Opti-MEM medium
5) Add 1.8 ml Opti-MEM 1 medium to the combined solutions
6) Add the resulting solution to cells
7) Replace medium after 8 hours with normal growth medium
8) After 48 h, subculture cells 1:10 into medium containing the appropriate concentration of G418 (Life Technologies)
9) Continue growing until clones of G418 resistant cells occur
10) Isolate stable clones and continue to propagate them individually

Fugene 6 (Roche Molecular Biochemicals) (GFP-constructs)

1) 24 h before transfection, seed $10^5$ cells in a 35 mm dish (50-80% confluency)
2) Add 4 µl Fugene 6 reagent to 96 µl Opti-MEM 1 medium
3) Incubate at RT for 5 min
4) Add 1 µg DNA to a new tube
5) Add Fugene 6 solution dropwise to DNA
6) Mix gently and incubate for 15 min at RT
7) Add solution dropwise to cells growing in normal medium
8) 24 - 48 h after transfection, analyze the cells

4.5.3. GFP expression

The pEGFP-C3 expression vector (CLONTECH) was used for the construction of GFP-p120 fusion proteins. To facilitate cloning of the cDNA regions coding for the N-terminal (120ANK) or C-terminal (120TM) domain, they were cloned by PCR, using the Pfu polymerase. These plasmids where transfected into COS cells with the Fugene 6 reagent (see above). The cells were analyzed with a Zeiss Axiovert fluorescence microscope equipped with a FITC filter set.
5. Literature


Heterogeneity at the 5'-untranslated region generated by an alternate exon.” J Biol Chem 267(9): 6188-96.


Vu, T. H., Shipley, J. M., Bergers, G., Berger, J. E., Helms, J. A., Hanahan, D.,


Curriculum Vitae

Personal Data

Name Daniel Jaquemar
Date and place of birth 6.5.1970 in Bregenz, Austria
Citizenship Austrian

Education

1975 - 1980 Primary School in Bregenz, Austria
1980 - 1989 Liechtensteinisches Gymnasium (LG) in Vaduz, Liechtenstein
1989 Matura Type B at the LG in Vaduz
1989 - 1994 Studies in Biochemistry and Molecular Biology at the ETH in Zürich, Switzerland
1994 Master thesis in Biochemistry in the laboratory of PD Dr. B. Trueb at the ETH Zürich
1995 - 1999 Preparation of the PhD thesis in the laboratory of PD Dr. B. Trueb at the University of Berne, Switzerland

List of publications
