TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Experimentelle Genetik

Protein Complex Analysis of Disease associated Proteins

Dipl.-Biol. Karsten Hermann Walter Boldt

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Dr. h.c. J. Bauer

Prüfer der Dissertation:
1. apl. Prof. Dr. J. Adamski
2. Univ.-Prof. Dr. B. Küster
3. Univ.-Prof. Dr. J. Buchner

Die Dissertation wurde am 25.08.2009 bei der Technischen Universität eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München am 10.02.2010 angenommen.

Meiner Frau und meiner Familie

A Abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BBS	Bardet Biedl syndrome
BCA	Bicinchoninic acid
bp	Base pair
BSA	Bovine serum albumin
CC	Interconnecting cilium
CDK2	Cell division protein kinase
CEP290	Centrosomal protein of 290 kDa
CEP97	Centrosomal protein of 97 kDa
CFC	Cardio facio cutaneous
CID	Collision induced fragmentation
CLUAP1	Clusterin-associated protein 1
COS-1	CV-1 in Origin, and carrying the SV40 genetic material
CP110	110 kDa centrosomal protein
ddH2O	Double distilled water
ddNTPs	Dideoxynucleotides
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
Erk	Extracellular signal-regulated kinase
FBS	Fetal bovine serum
FWHM	Full width at half maximum
GFP	Green fluerescent protein
GTP	Guanosin Triphosphate
HA	Hemagglutinin
HDAC6	Histone deacetylase 6
HEF1	Enhancer of filamentation 1
HEK	Human embryonic kidney

HPLC	High-performance liquid chromatography
HRP	Horse radish peroxidase
HSPB11	Heat shock protein beta-11
IFT	Intraflagellar transport
IgG	Immunoglobulin G
INL	Inner nuclear layer
IS	Inner segments
KIF7	Kinesin-like protein 7
KSR	Kinase suppressor of Ras-1
LB	Luria-Bertani
LC	Liquid chromatography
LCA	Leber congenital amaurosis
MALDI	Matrix-assisted laser desorption/ionization
MAPK	Mitogen-activated protein kinase
MAPKK	Mitogen-activated kinase kinase
MAPKKK	Mitogen-activated kinase kinase kinase
MEK	Dual specificity mitogen-activated protein kinase kinase
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NGF	Nerve growth factor
NIH3T3	3-day transfer, inoculum 3 x 105 cells
ONL	Outer nuclear layer
OPL	Outer plexiform layer
OS	Outer segments
PAGE	Poly acrylamide gel electrophoresis
PBS	Phosphate buffered saline
PC12	Pheochromocytoma cells
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PKD	Polycystic kidney disease
PMF	Peptide mass fingerprint
PVDF	Polyvinylidene difluoride

RABL	Rab-like
Rac	Ras-related C3 botulinum toxin substrate 1
Raf	Rapid accelerated fibrosarcome
Rap	Ras-related protein
Ras	Rat sarcoma
RPE	Retinal pigment epithelium
RPGR	Retinitis pigmentosa GTPase regulator
RT-PCR	Reverse transcriptase PCR
S/N	Signal to noise
SDS	Soduim dodecyl sulfate
SF-TAP	Strep-Flag-tandem affinity purification
SILAC	Stable isotope labeling with amino acids in cell culture
siRNA	Small interfering RNA
TAP	Tandem affinity purification
TAPo	Original TAP tag
TBS	Tris buffered saline
TEMED	Tetramethylethylenediamine
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TOF	Time of flight
TPR	Tetratricopeptide repeat
Tris	Tris(hydroxymethyl)aminomethane
TTC	Tetratricopeptide
WDR	WD-repeat
WT	Wild type
YFP	Yellow fluorescent protein

B Summary

Within the last years, more and more attention is laid on the analysis of protein complexes. Proteins are nowadays not longer considered to be single entities but parts of molecular machineries in form of protein complexes. Various methods were developed for the purification and the identification protein complex components. Since each method is characterized by certain advantages but also disadvantages, the aim of this study was to develop, establish and apply technology for this purpose and to decipher alterations in protein complexes.

The first step was the evaluation and optimization of the SF-TAP method, developed by us. In contrast to the original TAP tag, the tag is much smaller and there is no need for proteolytic cleavage since elution is achieved by competition for the binding sites on the affinity-resin. Evaluation of the method revealed that it is not only quick and straight forward but that it is at least as efficient as the original TAP method. For the analysis of three proteins, stably expressed in HEK293 cells, by quantitative Western blotting, I could determine an efficiency of 30-50%. Finally I applied the method to the analysis of the MAPK network. I could show that it is a highly suitable tandem affinity purification method for the analysis of protein complexes involved in cellular signaling. Combination of the SF-TAP method with direct, gel free analysis of the proteins resulted in a clearly increased number of proteins identified when compared to the classical, gel based approach.

After proof of principle by analysis of MAPK protein complexes, I applied the SF-TAP technology to the analysis of the lebercilin protein complex, to get insights into the function of this, at that time, unknown protein. Lebercilin is associated with Leber congenital amaurosis (LCA), a severe, inherited blinding disease. I could show that lebercilin is involved in a protein complex with mainly ciliary, centrosomal and tubulin associated proteins. This was the first hint towards a ciliary function of lebercilin in ciliary transport. I further analyzed the protein complex of lebercilin by quantitative complex analysis using SILAC, a mass spectrometry based method, combined with a quick, one-step purification. In the first step, the specific lebercilin complex components were detected by quantitative comparison of the lebercilin complex to the SF-TAP vector alone. In the second step, the complexes formed by wild type lebercilin and the complex

of the LCA causing P493TfsX1 mutant were quantitatively compared. The analysis revealed that lebercilin is associated with the intraflagellar transport machinery (IFT) especially with complex B of the machinery. The IFT is necessary for transport mechanisms along ciliary structures like the connecting cilium in photoreceptors. The IFT association of lebercilin is lost for the P493TfsX1 mutant. By further characterization of the IFT complex B using the SF-TAP, I could show that several more proteins, previously described as putative IFT proteins are actually part of the IFT complex B. This demonstrates the involvement of lebercilin in the transport at the connecting cilium in photoreceptors and its necessity for maintaining the outer segments of photoreceptors. Loss of this function leads to degeneration of the outer segments and eventually of the photoreceptors, resulting in blindness.

By applying the SF-TAP method and the quantitative protein complex analysis to decipher lebercilin complex components and to the comparison to the complex of the LCA associated mutant, I could clearly demonstrate the applicability of these methods. Both methods are suitable for the identification of protein complexes. Moreover, the quantitative protein complex comparison is a useful tool to identify the molecular consequence of mutations in human proteins that infer protein complex formation. The resulting knowledge can serve as the basis to characterize molecular mechanisms of disease.

C Zusammenfassung

In den letzten Jahren wurden einzelne Proteine immer mehr als Teil von molekularen Maschinen gesehen, organisiert in Proteinkomplexen, die als funktionelle Einheiten molekulare Prozesse umsetzen. Zur Analyse der Zusammensetzung von Proteinkomplexen wurden verschiedene Methoden entwickelt, eine der wohl wichtigsten ist die Doppelschrittaufreinigung von Komplexen über „tandem affinity tags". Da auch diese Methode mit inhärenten Nachteilen behaftet ist, war es Ziel dieser Arbeit verbesserte affinitätsbasierte Arbeitsprotokolle zu entwickeln und diese dann zur Analyse krankheitsassoziierter Proteinkomplexe anzuwenden.

Der erste Teil der Arbeit beschäftigt sich mit einer unter meiner Mitarbeit neu entwickelten Methode zur Tandem-Affinitätsreinigung (SF-TAP) sowie deren Optimierung und Evaluation. Es zeigte sich, dass die SF-TAP Methode sehr effizient ist und zwischen 30 und 50% des Zielproteins aufgereinigt werden können. Ein Vergleich mit der originalen TAP Methode ergab, dass die SF-TAP Methode sich durch eine mindestens genau so hohe Effizienz auszeichnet. Darüber hinaus ist die SF-TAP Methode deutlich einfacher durchzuführen: Im Zuge der Reinigung entfällt die bei der klassischen TAP Reinigung notwendige proteolytische Spaltung, bei SF-TAP werden die Proteinkomplexe ausschließlich durch Verdrängung eluiert. Um zu zeigen, dass diese Methode sich für die Analyse von Proteinkomplexen tatsächlich eignet, wurde sie für die Analyse von gut charakterisierten Proteinkomplexen im MAPK Signalweg erfolgreich angewendet.

Die SF-TAP Strategie wurde anschließend für die Analyse eines bisher unbekannten Proteinkomplexes, den des zu dieser Zeit unbekannten Proteins Lebercilin, angewandt. Mutationen in Lebercilin verursachen eine schwere Form der erblichen Blindheit, die Lebersche congenitale Amaurose (LCA). Wie ich mittels SF-TAP Analyse zeigte konnte, liegt Lebercilin in einem Komplex mit ciliären, centrosomalen und Tubulin assoziierten Proteinen vor. Dies war der erste Hinweis darauf, dass es sich bei Lebercilin um ein ciliäres Protein handelt, das am ciliären Transport beteiligt ist. Zur eingehenden Analyse des Lebercilin Komplexes verwendete ich dann SILAC, eine quantitative, Massenspektrometrie basierte Methode, in Kombination mit einer Ein-Schritt Affinitätsreinigung.

Im ersten Schritt konnte ich spezifische Komplexkomponenten identifizieren, im Anschluss daran den Proteinkomplex von Lebercilin mit dem der LCA verursachenden P493TfsX1 Mutante vergleichen. Die Analyse ergab, dass Lebercilin in einem Komplex mit Proteinen des intraflagellaren Transports (IFT) vorliegt, die Mutante hingegen keine Assoziation mit dem IFT Komplex zeigt. Der IFT ist ein essentieller Transportmechanismus an Cilien, insbesondere am Cilium in Photorezeptoren. Die Charakterisierung des IFT Komplexes B zeigte, dass dieser auch im verwendeten, humanen Zellkultursystem, wie in der Literatur vor allem für *Chlamydomonas reinhardtii* und *C.elegans* beschrieben, vorliegt. Meine Resultate zeigen, dass der Verlust von funktionellem Lebercilin wahrscheinlich zum Erliegen des Transports, dadurch zur Degenerierung der äußeren Segmente und der Photorezeptoren und letztendlich zur Blindheit führt.

Durch die Anwendung der SF-TAP und der quantitativen Methode für die Analyse und den Vergleich von Proteinkomplexen konnte ich zeigen, dass sich diese Methoden dafür eignen, Proteinkomplexe und durch Mutation erzeugte Änderungen in Proteinkomplexen zu identifizieren. Das daraus resultierende Wissen kann als Ausgangspunkt für weiterführende Analysen genutzt werden mit dem Ziel, die molekularen Mechanismen einer Krankheit zu aufzuklären.

D Introduction

1. Protein complexes

Within the last decades, the focus in analyzing the influence of proteins on disease mechanisms was mainly laid on applying methods to genetically manipulate protein expression or function. Therefore methods were developed for the knock out or knock down of specific genes like the knock out mice[1-3] or knock down by the usage of small interfering RNA which leads to the degradation of a specific mRNA[4-7]. In general, these methods, especially the generation of knock out mice, are very laborious and time consuming. An alternative strategy to get insights into the function of proteins is the analysis of protein interactions and complexes. By identifying the proteins that are associated with the protein of interest, it is possible to get insights into the functional network of a protein and thus into the molecular mechanism, the protein is involved in and suggestions on the functions it exerts.

Several decades ago, cells were considered as compartments where proteins were randomly colliding with each other due to rapid diffusion. Nowadays, it is obvious that proteins do not exert their function by random diffusion as single entities but in molecular machines, composed of several different proteins or copies of a single protein[8]. Each protein complex is an assembly of proteins, more or less stably associated or interacting in a dynamic fashion. The assembly of proteins in complexes results in limited number of possible reaction partners and thereby restricts the number of possible reactions. Additionally, the organization of the complex leads to a sequential order of reactions since one reaction depends on another one and can only take place after the first reaction was completed[9]. This creates a highly structured and regulated reaction space.

But these protein complexes do not perform a biological function as single machines; a protein complex interacts with other protein complexes and modulates their activity or is modulated by other complexes itself. As localization and reactions of proteins, the interplay of protein complexes is not randomly happening in cells. The localization of protein complexes is strictly regulated and depending on the localization, the function they perform can be different. Nevertheless, the function can not only be regulated by localization to different cellular compartments but also by

localization to certain organelles of the cell[10-13]. Ras, a small GTPase involved in cell cycle regulation as well as oncogenesis, for example can localize to different cell membrane structures. Recruitment of protein complexes to this structures results in distinct functions of these complexes, differing depending on the localization[14].

The composition of protein complexes may be highly regulated. Single or multiple complex components or even complete sub-complexes can either be recruited to or retracted from the complex[15,16]. These changes can happen within seconds or minutes, depending on the stimulation or inhibition which acts on a cell[17,18]. These rapid mechanisms enable cells to react on changes in their environment within a very short period of time. The processes of recruitment or retraction can be regulated by various mechanisms, for instance by post translational modification like phosphorylation[19,20] or by recruiting a protein that replaces a component by competing for the same binding motif[21-23]. Regulated assembly of protein complexes plays an essential role in many cellular processes. The MAPK cascade for example regulates the cell fate and can activate or inactivate other signaling cascades. C-Raf, MEK and Erk are assembled by the scaffold protein KSR upon EGF stimulation, these brings the three proteins into close proximity to enable the sequential activation of the cascade and results in proliferation of the cells[24,25]. Paxilin assembles the same proteins upon c-MET receptor activation and is phosphorylated by activated Erk. The focal adhesion kinase binds to this phosphorylation site and activates the phosphatidylinositol-3-kinase and Rac and thereby induces detachment of the cell from the substrate and migration[26,27]. These two examples show that one set of proteins, if assembled by different scaffold proteins, can lead to a completely different cellular function, to either migration or proliferation.

Another strictly regulated process, for which the assembly of a distinct protein complex is essential, is the intraflagellar transport machinery (IFT). It was discovered in *Chlamydomonas reinhardtii* flagella but plays an essential role in ciliated mammalian cells including photoreceptors. IFT is a highly conserved mechanism, composed of the same components in many species[28]. The complex consists of at least 30 proteins which can be grouped into at least three sub-complexes, the IFT complexes A, B and the Bardet-Biedl syndrome protein (BBS) complex[28,29]. The correct assembly of these sub-complexes and the entire complex is essential for the function of

the IFT. Disruption of the complex by dysfunction of a single component can lead to its dysfunction. The result on the cellular level may be and photoreceptor degeneration or even syndromic disorders, affecting various ciliated cells and tissues[30,31]. Assembly and disassembly of the IFT complex as well as the loading of the complex with cargo is a strictly regulated mechanism. This enables the highly specific transport of only certain components and is an example for highly regulated and specific assembly of protein complexes. Even though the composition of the complex and the sub-complexes is well established, it remains unclear how the loading specificity of the complex is regulated.

The above examples clearly demonstrate that the regulation of protein complexes and especially their composition can strongly alter the function of the complex and thereby not only influence or alter activation or inactivation but also induce a switch between certain functions. The modular use of proteins within dynamically regulated protein complexes increases systemic flexibility by which it can quickly be adapted to the needs of the cell in a certain state.

Until a decade ago, the analysis of protein complexes was restricted to small scale or rather artificial experiments. New complex components were usually identified by directed approaches or methods like yeast-two-hybrid, which enables the identification of direct interactions only[32]. Ten years ago, a first method for the large scale analysis of protein complexes was described. This method called tandem affinity purification (TAP) was first developed and applied for the analysis of protein complexes in yeast[33,34]. Although it was used for the analysis of protein complexes in higher eukaryotic cells as well[35-37], many efforts were undertaken to develop new tags and methods to be able to analyze protein complexes in these cells in a efficient way[38-41]. Even though the TAP was a big leap forward, several intrinsic weaknesses and limitations remain.

The next chapter will describe methods and approaches used for the purification and the analysis of protein complexes including their advantages and their disadvantages.

2. Purification and analysis of protein complexes

Protein complex analysis can basically be divided into two steps. The first step is the purification or enrichment of the protein of interest, the second step is the analysis of a protein complex. Since several methods exist for both of these steps, it is important that the combination of both methods is optimal for each specific purpose.

The purification and comprehensive analysis of protein complexes and protein interactions is an important and still demanding task. There are various methods available and all of them have several advantages and disadvantages. With a modern yeast-two-hybrid approach for example, a specific protein can be screened against complete libraries of potential binders within a reasonable period of time and high specificity[42]. The major drawbacks of yeast-two-hybrid based screens in the context of protein complex analysis are that only binary interactions can be detected. This can be an advantage for example if the structure of a protein complex is to be analyzed. Interactions depending on post translational modifications cannot be detected. Like most of the methods for protein complex and protein interaction analysis, yeast-two-hybrid screens are to some extend artificial and thereby can produce false positive results or miss important interactors. Nevertheless, its potential was demonstrated in various studies[42-47].

A number of frequently used methods for the purification or enrichment of protein complexes are the affinity-based methods. These methods either depend on the expression of tagged proteins in cell lines and purification of the tagged protein and the associated complex via the tag using affinity reagents or on the use of antibodies specific for a protein of interest. The focus of this study lies on affinity-based methods and the expression of tagged proteins. A selection of these methods will be described in the next chapters.

2.1 Purification of protein complexes

The first step in the analysis of protein complexes is the purification of these complexes. Depending on the down-stream analysis method, different levels of purity are necessary. For targeted quantitative or semi-quantitative approaches, an enrichment of the complex can be sufficient since only specific proteins are monitored and the abundance of the

proteins in the enriched complexes is compared to controls. For global, non-quantitative methods, the purity must be very high. Here it is not possible to clearly distinguish between specific and unspecific complex components. In this chapter, three affinity-based purification strategies are described. The first method is actually not a single method but describes an approach that is used in several similar versions, affinity purification via tags. The second method, the TAP method, is a combination of two tag-based affinity purification methods and the third approach is an antibody-based affinity purification approach.

2.1.1 Single-step purification via tags

To purify proteins via tags, the DNA expression construct must be modified in a way that the tag is fused to the expressed protein. This is only possible if the organism to be used for the analysis of the protein, can be genetically manipulated. Within the last few years, a huge variety of tags was developed. Commonly used tags are the Flag tag, the Strep-tag II, the His, the HA and the c-myc tag[48-54]. Table 1 gives an overview over these tags and the reagents used for binding the tagged proteins. The tags not only differ in their sequence and size, but in the binding affinity, avidity of the tag-affinity binder combination and in the extent of non-specific binding as well. The extent of non-specific binding is dependent on the tag, the affinity binder, the resin the binder is coupled to and on the method used for elution. It is also dependent on the amount of affinity capture reagent necessary for purification. Is the capacity of an capture reagent high, less reagent is necessary for the purification and thereby less non-specific binding is possible. A comparison of several commonly used tags showed that there are huge differences in the degree of contamination in the final eluates by non-specifically bound proteins[51].

The principle which is used for purifying protein complexes is the same for all tags. The expression construct for the protein of interest is manipulated so that the protein is expressed with the tag at the N- or C-terminus. In most cases, some amino acids are introduced between the protein and the tag as spacers to prevent folding or binding alterations due to the tag or the protein, respectively. Whether a protein should be tagged C- or N-terminally depends on the protein structure and function. Especially if the protein contains signal peptides or binding sites near one end, it should be preferred to attach the tag to the other end of the protein. This is also true

if a protein is cleaved and cleavage would lead to a loss of the tag. If no functional critical or interfering domains, signal peptides or cleavage sites are detectable, both versions should be tested.

Tag	Length (AA)	Sequence	Binder	Elution
Flag	8	DYKDDDDK	Antibody	Peptide
Strep-tag II	8	WSHPQFEK	Strep-Tactin	Desthiobiotin
His	6	HHHHHH	Ni2+-NTA Co2+-CMA	Imidazole low pH
Hemagglutinin (HA)	9	YPYDVPDYA	Antibody	Peptide
c-myc	10	EQKLISEEDL	Antibody	Peptide

Table 1: A selection of available affinity purification tags. The table shows a selection of available tags for the affinity purification of protein complexes. The first column shows the name of the tag, the second the number of amino acids of which this tag is composed, the third column the amino acid sequence. The affinity binder used to capture the tagged protein and the reagents used to elute the tagged protein from the binder are given in the last two columns. If peptide is given as elution reagent, the peptide sequence corresponds to the sequence given for the tag. In these cases and for desthiobiotin, the elution is accomplished by competition. In case of the His tag, there are two metal groups serving as affinity capture reagent and thereby two options for elution.

2.1.2 Tandem affinity purification

The tandem affinity purification (TAP) approach is essentially a sequential combination of two affinity purifications. The tagging of the protein is done as described for tagging with a single affinity tag but in contrast to the single-step purification method, two binding moieties are added in a sequential manner. The first TAP method described was a combination of a protein A domain and a calmodulin binding peptide (CBP) domain. The two moieties are separated by a TEV cleavage site, necessary to elute the bound protein from the IgG matrix which binds the protein A tag with high affinity (Figure 1). At first, TAP was developed for the analysis of protein complexes in yeast[34]. It was later applied in higher eukaryotes as well but with less success. Therefore, several optimized TAP methods have been developed and applied later on to purify protein complexes from different cell types and different organisms[39-41,55-58]. To optimize the TAP method, the binding moieties used in the original approach were exchanged against other binding domains. The overall principle remained the same, especially the proteolytic cleavage, was necessary for elution of the complexes from the resin used in the first step for all these novel TAP tags.

The major advantage of the TAP systems is the high purity of the obtained complexes as a result of sequential affinity purification. This allows a direct analysis of the complex by mass spectrometry and results in a high probability for detection of specific interactors. The major disadvantage is that the complexes must be very stable to be preserved during purification. Weak or transient interactions are hard to detect with a TAP method. The main reason for the loss of these interactors or complex components is the dilution that occurs when the eluate of the first purification step is transferred to the second resin. Therefore, as also observed in many cases in my own studies, a complex of stable binders to a given protein of interest can be detected, labile or transient interactors cannot. To gain insights into a basic function of a protein, this can nevertheless be very helpful. In many cases however, important interactors or complex components are missed by TAP approaches. Further, this method can principally be used to analyze protein networks as it was applied for in yeast. However, the effort is much higher in mammalian or higher eukaryotic systems when compared to yeast because genetic manipulation is more time-consuming and less cost effective.
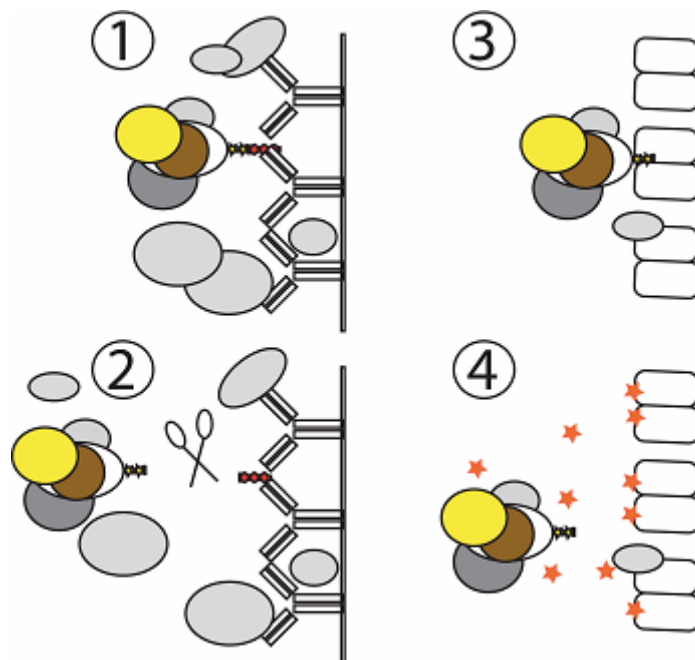
Figure 1: The principle of the original tandem affinity purification. In the first step (1), the tagged protein is bound to a protein A resin from which it is eluted by TEV protease cleavage which cuts the tag at the TEV cleavage site (2). In the third step, the tagged protein is bound to the Calmodulin resin by the Calmodulin binding peptide of the tag. (3) Elution is done by addition of EDTA which captures the calcium ions and thereby inhibits binding of the Calmodulin binding peptide to the Calmodulin resin (4).

2.1.3 Co-immunoprecipitation

Protein complex purifications by co-immunoprecipitation with antibodies specific for the protein of interest are mostly used to validate interactions detected by other methods[16] or for hypothesis-driven approaches[59]. Nevertheless, it can principally be used for the detection of complex components as well. Once combined with quantitative proteomic methods like SILAC, co-immunoprecipitations can also be used for high confidence detection of complex components (see 2.2.22.2.2)[60,61]. The main advantage of co-immunoprecipitations is that the complex of interest can be analyzed in its native tissue, under physiological or at least near physiological conditions. The main disadvantage is that the method is absolutely dependent on the availability of highly specific and efficient antibodies. In addition and in contrast to more or less standardized affinity purification methods, protocol optimizations to achieve high purities and sufficient amounts for down-stream analysis methods are essential in many cases, especially if the resulting purification is to be used for the identification of new interactors and complex components.

2.2 Quantitative protein complex analysis

As stated above, for the analysis of protein complexes it is in many cases necessary to employ quantitative or at least semi-quantitative techniques to be able to distinguish between non-specifically bound and specifically enriched proteins. Since the focus in this study lies on the combination of quantitative mass spectrometry (MS) with protein complex analysis, some of these technologies and their application to protein complex analysis will be described in the following two chapters.

2.2.1 MS-based quantification

Within the last decade, several methods for MS-based quantification were developed. The basic principle is the same for most of the methods. One protein set is labeled with a light modification, the other set with a heavy isotope labeled form of the same label. The different isotope composition of the labeling reagents creates a mass difference which is visible in the MS spectra. Thereby the peak intensities or the area under the peaks of the light or heavy labeled peptides can be compared and enable relative quantification. Most of these are based on the introduction of specific chemical, isotope coded modifications either on the protein or the peptide

level. Because of the difference in modification strategies, the level on which the modification is introduced differs. The earlier in the preparation process the samples can be combined; the lower is the risk to introduce variability during the sample preparation process. For SILAC[62] (stable isotope labeling by amino acids in cell culture) for example, a heavy isotope labeled amino acid is integrated into all proteins of a cell by its addition to the cell culture medium. Thereby, the proteins are label on the cellular level already and the samples can be combined at a very early stage. The ICAT[63] (isotope coded affinity tag) and ICPL[64] (isotope coded protein label) technologies are both based on attaching the chemical labels to the proteins in the samples on the protein level by covalent bonds. Even though the principle is similar, the two methods differ in the composition of the chemical labeling reagent used and in the reactive groups. ICAT reagents react with thiol groups of proteins and ICPL reagents with free amine groups. In contrast to ICAT and ICPL, the iTRAQ (isobaric tag for relative and absolute quantification) technology[65] is based on labeling of peptides instead of proteins. The N-terminus and the side chain amines of peptides are labeled after proteolytic cleavage of the proteins with isobaric reagents. Because isobaric reagents are bound to the peptides, the mass difference is not visible on the peptide level. During fragmentation of the peptides, the isobaric tags break apart into fragments of different sizes for the different labeling reagents and these fragments can be quantitatively compared to each other. This is the major difference to the other technologies described here. SILAC, ICPL and ICAT quantification is performed on the level of the MS spectrum of the intact peptides in contrast to iTRAQ where the quantification is done on the level of the fragment ion spectrum. An alternative technology to the isotope coded labeling strategies is the label-free quantification[66]. It is not based on the introduction of isotope coded labels. The unmodified samples are analyzed separately by LC-MS/MS. After the analysis, the LC chromatograms are aligned and the signal for a single peptide is extracted. The peak areas are integrated over the chromatographic time scale. These peak areas are then compared between the separate LC-analyses and the ratios calculated. In contrast to the other methods, the samples can not be combined before LC-MS/MS analysis. For that reason, a highly reliant and reproducible LC-MS/MS system and absolutely comparable and robust sample preparation is a prerequisite for the application of this method.

2.2.2 Application of MS-based quantification to protein complex analysis

Up to date, mainly SILAC was applied to the quantitative analysis of protein complexes and protein-protein interactions. Because the protocols for all chemical labeling strategies are optimized for labeling protein amounts of more than 50 µg, it is not trivial to use these methods for the quantification of protein complexes where the amount of protein is substantially lower (own observation and personal communication with Andreas Vogt and Alexander Schmidt).

The identification of interaction partners or protein complex components is achieved by quantitative comparison of the enriched protein complex to a negative control for proteins, non-specifically binding to the affinity matrix or the tag. Proteins, significantly enriched compared to the negative control are considered to be direct interaction partners or complex components, depending on the experimental setup[60,67]. Peptide pull downs for example can be used to identify proteins, directly interacting with a certain amino acid sequence of a protein[68]. In a combined immunoprecipitation (IP) and siRNA knock down approach (QUICK)[61], protein complexes can be identified by comparing IP eluates of normal cells versus those of cells, in which the protein of interest was knocked down. In this experimental setup, the eluates derived from knock down cells are used as negative control. Proteins significantly enriched in the IP compared to the control can be considered as specific components of the protein complex of the protein of interest. SILAC was not only used to identify protein interactions or protein complex components but for the comparison of those as well. For this purpose, instead of comparing the enriched complexes to a control, the same enrichment is performed from cells subjected to different stimuli. The impact of EGF stimulation on proteins interacting with the SH2 domain of the protein GRB2 for example was studied by quantifying proteins, binding to this SH2 domain upon EGF stimulation compared to the non stimulated state[38].

A significant advantage of the quantitative approaches is the possibility to distinguish between non-specifically bound proteins and actual protein complex components or protein interactions. This is even possible if the degree of enrichment is relatively low and the amount of non-specifically bound proteins is substantial. The additional experimental and financial

effort can however be compensated by the possibility to detect not only stable interactions but transient ones as well.

3. Deciphering disease mechanisms by protein complex analysis

When analyzing the mechanism of human diseases that are caused by dysfunction of a single or multiple proteins, it is not only important to describe the function of the proteins. It is even more important to decipher the function that is lost, or in case of gain of function mutations, the function of a protein that is gained. This is especially important when considering the development of therapies. Therefore it is essential to know which pathway or function to target, particularly when the protein exerts several functions or if the function is unknown. As described above, a rapid and straightforward way to gain insights in the function of a protein is the analysis of its protein complex. This is much faster and easier to do than for example a knock out or knock down, especially if knowledge about a protein of interest is sparse. Even if a knock down by siRNA itself is a fast way to manipulate the expression of a protein, it is not clear which read-out can be used to analyze the changes induced by the knock down. The analysis of the protein complex can be done within few weeks and can be the first step to get insights into the function of a protein. This can then be validated and further analyzed by alternative methods, like functional assays which can only detect changes in one or few functions, siRNA knock down or even by the generation and analysis of knock out model animals. The composition of the protein complex can reveal many insights into the function of the protein analyzed. For achieving knowledge on the lost or gained function of a mutated protein, the complexes formed by the wild type (WT) and the mutated protein can be quantitatively compared. This comparison can be done by expression and purification of both, the WT and the mutant protein, combined with quantitative comparison of the purified or enriched protein complexes. Given that a certain protein or even a sub-complex appears weaker or stronger associated with the mutant, this discrepancy can suggest the gain or loss of function and thus provide valuable insights into the disease's mechanism.

In some cases, the disease is caused by either weaker or no expression of the affected protein or by over expression. In such cases (gene dosis effect), it may not be possible to study the differences in protein

complexes between the disease causing gene variant and the WT protein as either the coding sequence of the gene is the same or the protein expression is completely lost. However, if mutations of the same protein are available or can be generated that are likely to impair the function of a given protein, comparative analysis of the interactome of both variants makes sense. Even if it is not possible to compare the complex of a mutant to the WT protein, the analysis of the protein complex is still a valuable tool to identify the function of a yet poorly characterized protein.

In this study, methods for purifying and identifying protein complexes as well as for the quantitative comparison have been developed and applied to the analysis of protein interactomes. The newly developed SF-TAP method was first applied to the analysis of the well known network of the mitogen activated protein kinase (MAPK) pathway and to the yet ill characterized, ciliary protein lebercilin, associated with the severe blinding disease Leber congenital amaurosis (LCA). The lebercilin protein complex was then subjected to a quantitative comparison of the interactome of the WT and a LCA associated variant. The following chapters describe the biological context in which lebercilin was described to be involved in.

4.   Lebercilin and ciliary disease

Cilia and flagella are microtubule-based, hair-like structures which project from the cell body, originating in the basal body. Cilia are present on almost every vertebrate cell type[69], in contrast to invertebrates where cilia are only present on sensory neurons and necessary for sensing chemical and mechanical signals[70]. Ciliary structures were first discovered more than 100 years ago[71] but became highly relevant for biomedical research only within the last years because ciliary dysfunction was associated with various human syndromes and diseases[72] like polycystic kidney disease, Leber congenital amaurosis[73,74], Joubert syndrome[45,75], Bardet-Biedl-Syndrome[76] and various other diseases[77]. These syndromes and diseases are summarized as ciliopathies. Although they are characterized by various and different phenotypes, they have one thing in common, the dysfunction of cilia[78].

Cilia can be grouped into two categories: 1) motile cilia and 2) non motile cilia. Motile cilia are present on many cell types in clusters like on cells of the respiratory tract where they beat in wave like concerted motion[79]. Another group of motile cilia is present on cells of the embryonic node.

Only one primary cilium is present on each cell but as in the respiratory tract, many cilia move in concerted action and thereby create a constant flow which transports signaling molecules from one side of the node to the other and thereby enabling the left right asymmetry determination[80,81]. In contrast to motile cilia, non-motile cilia, as the name suggests, are non-moving ciliary structures. One of the most important non-motile cilium is the connecting cilium of photoreceptors. It is the only connection between the inner and outer segment of mammalian photoreceptors. Additionally it is an essential filter for vesicular transport from and to the outer segments[82,83].

4.1 The structure and function of cilia

4.1.1 Cilia assembly and disassembly

Cilia assembly and disassembly is a process, closely correlated with the cell cycle. In general, cilia are assembled once cells exit the cell cycle and disassembly takes place when the cells re-enter the cell cycle and begin to proliferate. In normal proliferating cells, cilia are only present very transiently during G1 phase and the cilium is resorbed before or during the G2 phase[84-87]. The centrosome plays an important role during this assembly and disassembly process. While it functions as organizer of the microtubule-based cytoskeleton during interphase and of the bipolar spindles during mitosis, it is also essential as basis for building up the cilium[88]. The centrosome, composed of two barrel type centrioles and surrounded by pericentriolar material, migrates and docks to the membrane where one centriole serves as basal body for cilium formation[89]. The assembly of cilia is initiated by transport of specific components of the cilium towards the basal body and pre-assembly of substructures[90]. The pre-assembled components are transported by the intraflagellar transport machinery across the border of the ciliary compartment to enable the so-called compartmentalized ciliogenesis leading to the build up of the ciliary axoneme (Figure 2)[91,92]. The regulation of cilia formation is not well understood yet. However, there are some indications on how the regulation is coordinated. CP110 phosphorylation by CDK2 in G1 leads to its recruitment to the centrosome by CEP97[87]. Since knock down of CP110 or CEP97 leads to an increased proportion of ciliated cells, CP110 phosphorylation and inhibition by CDK2 could link cell cycle regulation and

cilia formation[85]. The regulation of cilia disassembly is regulated by the AuroraA kinase and HEF1. HEF1 probably stabilizes AuroraA which then initiates destabilization and disassembly of the cilia, possibly by phosphorylating HDAC6, a α-tubulin deacetylase that destabilizes the cilium by deacetylating α-tubulin[93,94]. Even though these regulatory mechanisms explain partially how ciliogenesis and cell cycle regulation are coordinated, this process in its entirety is not well understood. Many open questions remain.

4.1.2 The structure of cilia

The cilium is a highly structured organelle (Figure 2). It is composed of the basal body, a transition zone, the axoneme and the ciliar tip. The basic structure is composed of tubules, but the composition of this tubules and the organization varies depending on the type of cilium and the region in each type. The axoneme of motile cilia consists of a 9+2 structure where nine doublets of A and B tubules surround a pair of two central tubules. Inner and outer dynein arms extend from each A tubule and generate the force for cilium movement in an ATP-dependent fashion[95]. The primary, non-motile cilia are generally composed of 9 doublets without the central pair (9+0 structure). Even though this structure and this classification is generally true, there are several reports, describing exceptions with deviations from this general structure[96-99].

The complete ciliary structure is enclosed by a membrane which is a projection of the plasma membrane (Figure 2). The basal body originates from one centriole of the centrosome with a 9+3 structure. Every one of the 9 contains a triplet of A, B and C tubules in contrast to the axoneme tubules, doublets consisting of A and B tubules only. The end of the basal body is the proximal amorphous disc, a cartwheel structure followed by a middle piece that lacks appendages and the distal end with transition fibers. In the transition zone, the triplet structure of the basal body is converted into the doublet structure of the axoneme. Proximal transition fibers link every tubule that does not carry a dynein arm to the membrane. This region forms the part of the cilium where the intraflagellar transport particles accumulate and are loaded. In the distal zone, the 9+2 structure of the axoneme arises and the access control for particles to be transported towards the cilium takes place[72,100,101]. The axoneme is composed of a 9+2 or a 9+0 structure, depending on the type of cilium.

The axoneme ends in the ciliar tip where the switch between plus-end directed and minus-end directed transport takes place[28].
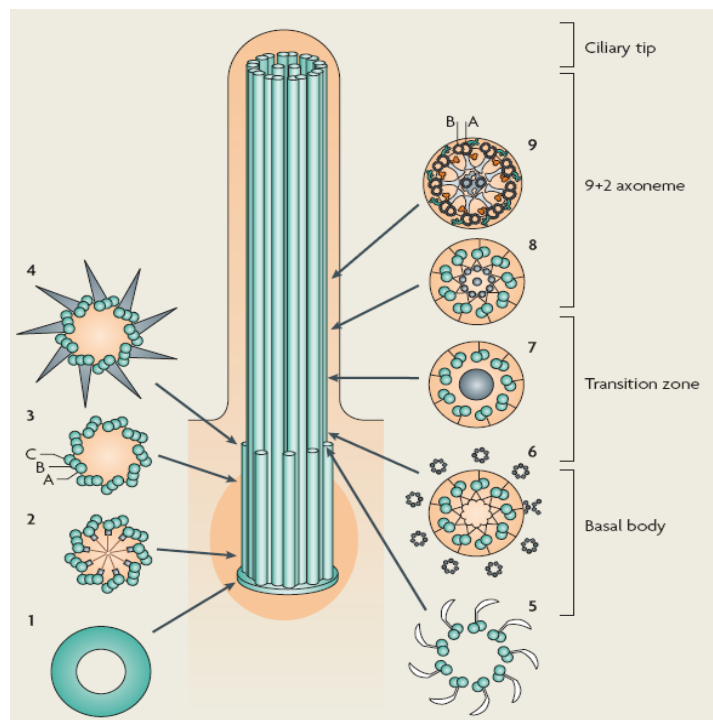
Figure 2: The structure of the cilium. The cilium is composed of the basal body, from which it originates, the transition zone, the axoneme and the ciliary tip. The basal body is composed of a 9+3 microtubular structure (2-4) and is surrounded by the pericentriolar material (dark orange). It is capped by the proximal amorphous disc (1). The transition zone is the region where the microtubular triplet structure is converted to the doublet structure of the axoneme. The proximal transition fibers connect the microtubules with the plasma membrane and thereby serve as a border to the axoneme and the region where the transport complexes are assembled (5-8). The axoneme is the core part of the cilium where the transport processes take place. It is composed of a 9+2 structure. The doublets are connected by nexins and are held in place by radial spokes which connect them to the central doublets. In motile cilia, dynein as a motor protein can attach and detach to and from the B tubules and can thereby achieve movement of cilia. This figure is reprinted with the friendly permission of the Nature Publishing Group[72].

4.1.3 Ciliary transport

The ciliar transport is mediated by the intraflagellar transport machinery (IFT). This machinery transports cargo by a mechanism that is not completely understood along the cilium in a kinesin-driven anterograde and a dynein driven retrograde transport.

The current model for the IFT proposes six functional steps[28,29,92,102]. In the first step, the IFT components are assembled at the distal part of the cilium near the transition zone of the cilium at the transition fibres[103]. Step two is the anterogade transport of vesicles along the cilium. Steps three and four are the unloading and loading steps at the tip of the cilium. This

process is not well understood yet but involves the deactivation of the kinesin motor protein, the disassembly of the IFT complex and the reassembly including the activation of the dynein motor protein. The fifth step is the retrograde transport and step six, the disassembly of the IFT complex at the transition zone of the cilium including unloading of the IFT particles (Figure 3). Although the mechanism of the IFT yet ill defined, the composition of the complex and the sub-complexes is well described even though not all complexes are experimentally confirmed in mammalian cells or tissues[28,29,92,104]. The components of the complexes as well as putative IFT proteins are shown in Table 2. The IFT complex consists of at least three sub-complexes[105], the complex A, B and the BBS complex.
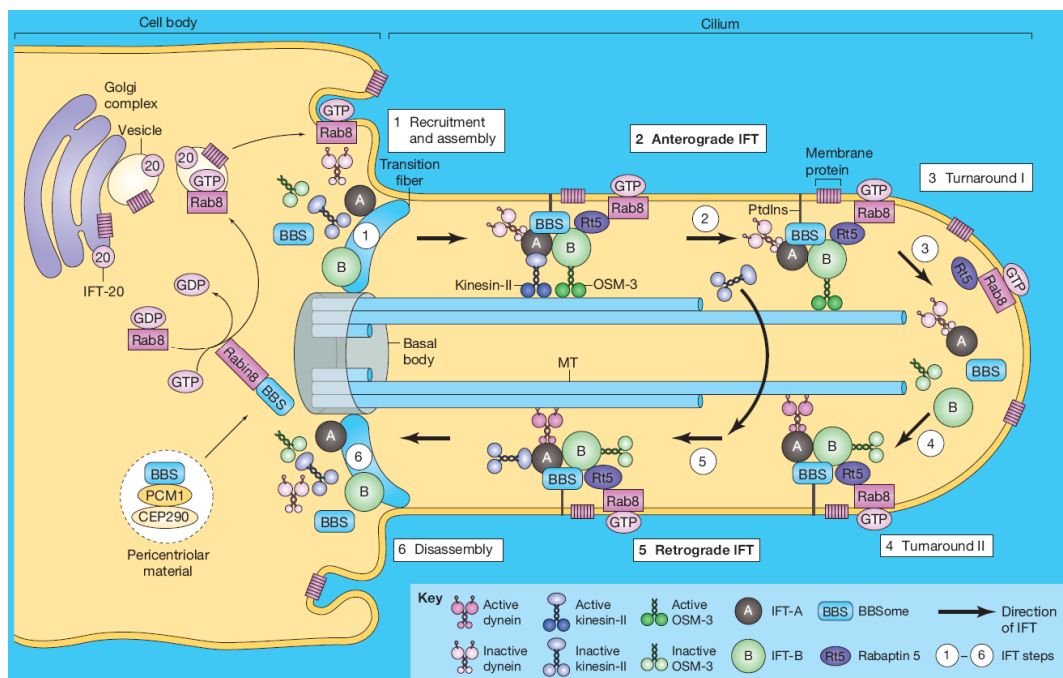
Figure 3: Model for the intraflagellar transport. The anterograde transport complexes which are composed of kinesin, the BBS complex and the complex A and B of the ciliary transport machinery are recruited and assembled (1). This complex is then transported in an ATP-dependent manner by the kinesin motor protein towards the tip of the cilium. Dynein, the motor protein for the retrograde transport is transported along with this complex (2). At the tip of the cilium, the complex for anterograde transport is disassembled (3) and the complex for retrograde transport is assembled (4). In the retrograde transport, the dynein motor protein moves the complex towards the basal body, again in an ATP-dependent manner (5). In this case kinesin is transported along with the complex. At the basal body, the complex is disassembled and the components can be used for assembling the anterograde transport machinery (6). This figure is reprinted with the friendly permission of the Company of Biologists Ltd[104].

The complex A consists of at least six different proteins, the complex B of at least 12 and the BBS complex of at least seven proteins. These complexes link the IFT cargo to the motor proteins driving the movement.

28

Up to date it remains unclear how cargo loading specificity in different cells is conferred. A possible mechanism could be mediated by specific isoforms of proteins, only present in a certain cell type.

IFT complex A		IFT complex B		BBS complex		Putative IFTs
IFT43		HSPB11		BBS1		TPR30A
IFT122A		IFT20		BBS2		CLUAP1
IFT122B		IFT27/RABL4		BBS4		WDR35
IFT139		IFT46/c11orf60		BBS5		TTC26
IFT140		IFT52		BBS7		RABL5
IFT144		IFT57		BBS8		
		IFT72		BBS9		
		IFT74				
		IFT80				
		IFT81				
		IFT82				
		IFT88				
		IFT172				

Table 2: The IFT complexes. The table shows the described IFT complexes including the proteins of which they are composed. The complexes where described in several tissues and are highly conserved among different species. Additionally, several proteins are described as putative IFT proteins which are shown in the last column. These proteins could not be assigned to the other complexes so far but where all described to be important parts of the IFT[28,29,92,105-110].

The ORF15 isoform of the retinitis pigmentosa associated protein RPGR (retinitis pigmentosa GTPase regulator) for example is predominantly expressed in photoreceptors. It localizes to the cilium and is involved in the ciliary transport[111-113]. Although it was not shown yet that RPGR is involved in determining the specificity of the IFT, it is a possible, photoreceptor specific candidate.

Knock out or knock down of one of the components of the IFT machinery leads to severe impairment or even disruption of the complete IFT and to severe phenotypes in animal or cellular models[30,31,107,108,110,114]. If one of these proteins is mutated in humans, this leads to more or less severe phenotypes as well. The same effect was described for some of the putative IFT proteins as well. These proteins are not yet described as part of one of the sub-complexes but they were shown to be part of the intraflagellar transport machinery (Table 2).

4.1.4 Functions and signaling mechanisms associated with cilia

The different functions that cilia exert take place in different cell types and tissues at different points in life of various organisms. The first function of cilia described in the life of an organism can be observed at the node of embryos where left-right patterning is organized. The model for left-right patterning describes a left directed flow of extra-embryonic fluid that is generated by concerted beating of primary cilia of the nodal cells[81]. This flow transports signaling molecules secreted by cells at one border of the node to the other border and thereby creates gradients that are needed for left-right axis determination in the embryo. Two models on how the left-right determination is established exist. One model proposes that Ca^{2+} release on the left node border is dependent on mechano-sensation of cilia, the other model proposes a mechano-sensation independent mechanism in which the nodal cilia act as antenna with receptors at or near the tip from where the signal is generated that leads to Ca^{2+} release on the left nodal border[115,116]. Both models involve cilia as both, flow creating as well as sensory organelles.

Additionally, cilia or flagella are necessary for the movement of certain cell types. Sperm movement for example is driven by motile flagella consisting of similar structures as motile cilia. Flagella consist of a 9+2 structure and dynein-driven beating leads to movement of the sperm[72,117,118]. Defects of this structure lead to male infertility due to loss of sperm movement[119,120].

4.1.4.1 *The cilium in the mammalian retina*

The mammalian retina is the structure in the mammalian eye in which light reception takes place. In the outer segments of photoreceptors, photon reception is converted into a chemical signal and further transformed into an electrical one. After integration of the signal of several photoreceptors, it is then transmitted to the brain via the optic nerve.

The retina consists of two distinct tissues, the retinal pigment epithelium and the neural retina. The retinal pigment epithelium is composed of a single layer of epithelial cells. The neural retina is composed of five different types of neurons and glial cells. Four of the five neuronal cell types are bipolar-, horizontal-, amacrine-, ganglion cells which integrate, transmit or relay the signal produced in the fifth cell types, the photoreceptors. The structure of the mammalian retina is highly organized.

30

It consists of three somatic layers, where the cell bodies of the neurons lie and two plexiform layers, where the synapses are located. The photoreceptors are the neuronal cells in which the light sensation takes place. They are connected to the bipolar cells in the outer plexiform layer (OPL). Photoreceptor nuclei are located in the outer nuclear layer (ONL) which is followed by the inner segment (IS) where the protein biosynthesis and metabolic processes take place. The IS is connected to the outer segment (OS) by the connecting cilium[121]. The connecting cilium is a very specialized ciliary 9+0 structure with doublet microtubules in the axoneme changing into a singlet structure in the proximal segment. It reaches from the cellular body of the photoreceptor (IS) deep into the outer segment where the opsin-containing membrane staples are localized and light sensation takes place. The only mechanism by which components, including rhodopsin and membrane particles can be transported from the inner to the outer segment is via IFT along the connecting cilium.
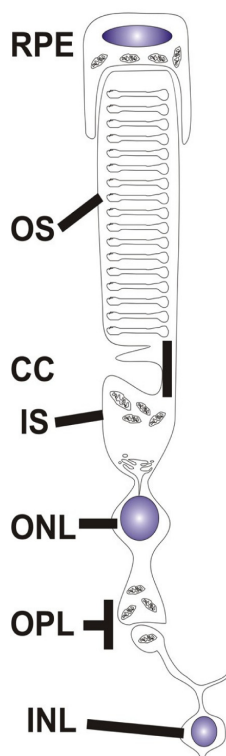
Figure 4: The structure of mammalian photoreceptors. The first segment of mammalian photoreceptors lies in the outer plexiform layer (OPL) where the photoreceptors are connected by synapses to the bipolar cells which transport the electrical signal from the photoreceptors to the ganglion cells. From these cells, the signal is transported via the optic nerve to the brain. The outer plexiform layer is followed by the outer nuclear layer (ONL) where the nuclei of the photoreceptors are localized. The protein biosynthesis and the metabolic reactions take place in the inner segment (IS). From here, the proteins destined to localize to the outer segments (OS) need to be transported to the OS via the interconnecting cilium (CC). This leads to a highly regulated and controlled transport. The reception of light takes place in the OS where membrane staples, packed with rhodopsin enable efficient light reception. The retinal pigment epithelium (RPE) is composed of pigmented cells, reflecting the light and additionally phagocyting the shedded parts of the photoreceptors. This figure is reprinted with the friendly permission of Springer[122]

The extent of transport necessary at the connecting cilium is probably the highest on all cilia known because of disc shedding, leading to the requirement of constant renewal of the outer segments from the basis. In the disc shedding process, outer segment discs are taken up by

phagocytosis, facilitated by retinal pigment epithelium cells[123,124]. This leads to loss of material in the outer segments which needs to be replaced and therefore transported from the inner to the outer segment along the connecting cilium. It is essential that this transport takes place in a very controlled fashion so that only those proteins are transported to and from the outer segments that are selected for transport.

4.2   Ciliopathies

The term ciliopathies describes a rather divergent class of human diseases with various and diverse phenotypes. However, all of these diseases have one thing in common, the involvement of ciliary dysfunction[72,78,125]. The first ciliopathy was described in 1976 when it was discovered that the phenotypes of the Kartagener syndrome (frequent infections of the respiratory tract, infertility and situs inversus) were caused by immotile and structurally abnormal cilia[126]. The important role of cilia and their association with many human disorders and syndromes became evident after mutation of IFT88 was discovered in a mouse model for polycystic kidney disease. It was described that IFT88 was required for assembly of ciliary structures[127], that cilia in the renal tube and at the embryonic node were missing or short and that IFT88 is required for left-right axis determination[127,128]. This led to the finding that the two polycystic kidney disease (PKD) proteins PKD1 and PKD2 localize to cilia[77,129] and thereby to the discovery that PKD is a ciliopathy.

The phenotypes caused by ciliary dysfunction are very diverse. They involve developmental defects like polydactyly, cranio-facial defects caused by mutations in IFT88 and IFT172 leading to impairment of hedgehog signaling due to disrupted IFT[77,130,131] or in non-canonical wnt signalling[132]. Defects in cilia can also lead to sperm defects since the sperm flagellum is a ciliary structure as well[119], loss of vision like in Leber congenital amaurosis[73,75], hearing loss (Usher syndrome[133-136]) and even obesity (Bardet-Biedl-Syndrome[76]). These diseases are generally not characterized by a single phenotype and are not restricted to a single organ but rather to several affected organs and thereby several clinical phenotypes that are overlapping between the different syndromes and diseases. However, there are also ciliopathies affecting only a single organ. One of these is Leber congenital amaurosis (LCA), only affecting the eye. LCA was first described by Theodore Leber in the year 1869[137]. It is one of the most

severe forms of blinding disease affecting children within the first days or months after birth[138]. LCA is characterized by degeneration of the outer segments of photoreceptors eventually leading to blindness. Even though not all of the genetically diverse forms of LCA are caused by mutations in ciliary proteins, a large group of the proteins associated with LCA were shown to localize to the connecting cilium of photoreceptors[73,138-141].

4.3 Lebercilin

The protein lebercilin is a newly discovered protein which was found to be associated with LCA[138]. It is encoded by LCA5 gene. The locus of the LCA5 gene was first identified in a consanguineous family of the religious community of the Old Order River Brethren[142] and linkage to this locus on chromosome 6 was also reported in an consanguineous Pakistani family[143]. Association of the C6ORF152 gene was first identified by den Hollander et al[73]. It was identified by homozygosity mapping of three affected families and of 33 consanguineous and 60 non-consanguineous affected individuals. Sequence analysis revealed a homozygous frame shift mutation in the three families (1151delC, P384fsX17) resulting in a truncated protein, a homozygous frame shift mutation in one of the individuals (1476dupA, P493TfsX1) and a homozygous nonsense mutation (835C, Q279X) in the second one. In the original family, no mutation in the coding sequence or the splice junctions of LCA5 could be identified. However, a 1598 bp deletion could be found in the promoter region of the LCA5 gene leading to absence of the transcript and thereby of the protein.

The lebercilin protein is composed of 697 amino acids and has a calculated molecular weight of 80.5 kDa. Bioinformatic analysis revealed four regions of the protein that were predicted to form coiled-coil domains. No further functional domain structure could be identified. Additionally, weak homology could be found with the centrosomal protein CEP290 but only one true homologue could be identified (C21ORF13), now called lebercilin-like protein. Expression analysis by RT-PCR showed that lebercilin is widely expressed in human tissues and in cell lines (Figure 5).
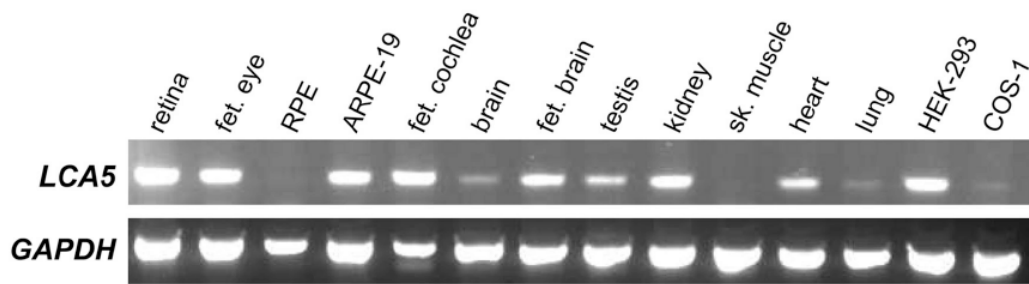
Figure 5: Expression of lebercilin in tissues and cell lines. The expression of lebercilin mRNA was investigated by RT-PCR analysis. Lebercilin shows a wide expression which is not limited to the photoreceptors or retina where disease phenotype is visible. It is strongly expressed in the fetal eye, fetal brain, testis, kidney, heart and in HEK293 and ARPE-19 cell lines. Weak expression is observed in COS-1 cells, in brain and lung. No expression is visible in the pigment epithelium and in skeletal muscles. This figure is reprinted with the friendly permission of the Nature Publishing Group[73].

Lebercilin expression was also examined at several developmental stages in mouse embryos by *in situ* hybridization. It is expressed ubiquitously at day 12.5 post coitum but was most pronounced in the eye, kidney and inner ear, regions of the central and peripheral nervous system as well as in the ciliated epithelium of the trachea, the nasopharynx and the lungs. Detailed examination of lebercilin expression in the eye showed that its expression shifted from the ganglion cell layer to the photoreceptor layer during development and was only observed in the photoreceptor layer in the adult eye. This showed that lebercilin is mainly expressed in ciliated tissues. This observation was strengthened by the fact that it is present in the ciliary proteome database[144,145].

Additionally, lebercilin was described to localize to microtubule structures, especially to the cilium and centrosome upon expression of the fluorescently labeled protein in mammalian cell lines (Figure 6). In the ciliated human ARPE-19 cell line, lebercilin localizes to the cilium, as demonstrated by co-localization with acetylated tubulin. Upon higher expression levels, lebercilin localization is not longer restricted to the cilium but it localizes to wave form like microtubular structures and the centrosome as well. In COS-1 cells a similar localization pattern was observed. The most pronounced signal of eYFP tagged lebercilin was found at the centrosome but a almost complete co-localization with α-tubulin was observed as well.
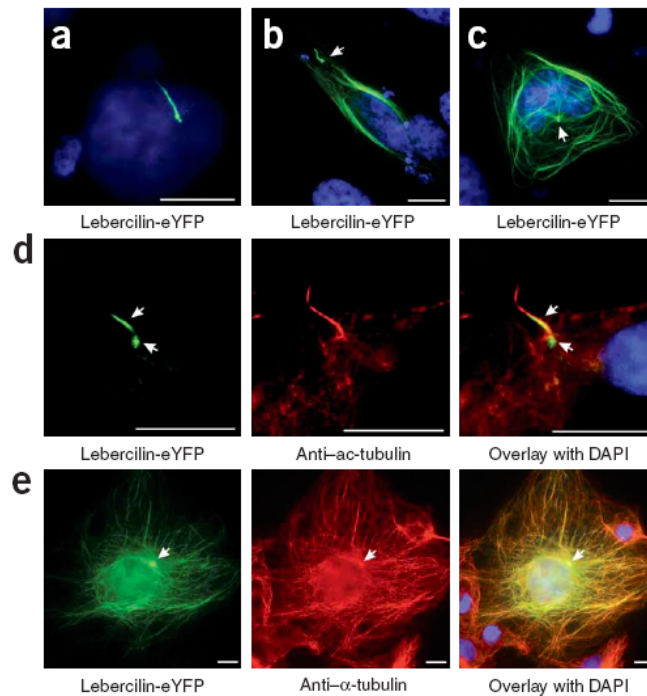
Figure 6: Localization of recombinant lebercilin in cell lines. eYFP tagged lebercilin was expressed in ARPE-19 (a-d) or COS-1 cells (e). a), b) lebercilin localizes to the primary cilium of ARPE-19 cells (arrow) and to microtubular structures when expressed at higher levels (b). If the cells do not have a primary cilium but a centrosome, lebercilin localizes to the centrosome (b, arrow) and to microtubular structures. Co-staining of the cells with an antibody, recognizing acetylated Tubulin shows co-localization at the primary cilium and additionally, that lebercilin mainly localizes to the region of the basal body and the axoneme/tip of the cilium (d). Co-staining with α-tubulin in COS-1 cells shows a almost complete co-localization of lebercilin with microtubular structures. This figure is reprinted with the friendly permission of the Nature Publishing Group[73].

The localization of lebercilin was additionally examined in mouse photoreceptors by immuno-staining and confocal microscopy as well as immuno-gold staining and electron microscopy (Figure 7). For both localization studies, a polyclonal anti-lebercilin antibody was employed. Light microscopy showed that lebercilin localizes almost exclusively to the retinal layer where the connecting cilia of photoreceptors are present (Figure 7d). High resolution and magnification images showed that lebercilin mainly localizes to the basal body, the axoneme and the tip of the connecting cilium (Figure 7e,f). By electron microscopy, the localization could be further pinpointed to the tubules of the cilium (Figure 7g).

In summary, these data strongly suggest that lebercilin is a ciliary protein which is expressed in ciliates cells and tissues and that it localizes to ciliary and centrosomal structures. However, it remains unclear which function lebercilin exerts and why mutation or absence of the protein leads to LCA.
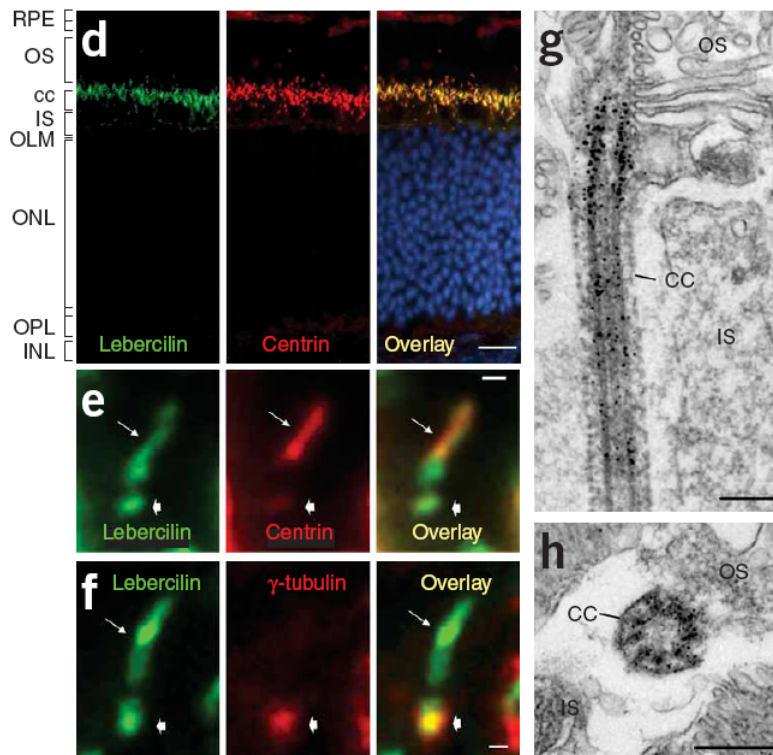
Figure 7: Sub cellular localization of endogenous lebercilin. d) radial cryosections where stained with a polyclonal rabbit anti-lebercilin antibody. Centrin was used as a marker for the connecting cilium. A complete co-localization of lebercilin and centrin is visible which shows that lebercilin is localized to the connecting cilium of photoreceptors. e) shows co-staining with centrin in high resolution and magnification. Lebercilin localizes to the basal bodies as well as to the axoneme and the tip of the connecting cilium. Co-staining with γ-tubulin, which localizes to the basal body, shows a co-localization at the basal body. g) and h) show immuno-gold staining with the lebercilin antibody and electron microscopic images of the connecting cilium. A longitudinal section of the axoneme and the tip shows that lebercilin mainly localizes to the tip of the connecting cilium and a transversal section that it localizes to the microtubular doublets but not to the central part of the ciliary structure. This figure is reprinted with the friendly permission of the Nature Publishing Group[73].

36

E Aim of the study

The analysis of protein complexes is still a field in movement and development of many tools and methods have been described during the last years. Especially, the availability of standardized methods enabling the fast and straightforward affinity purification and identification of protein complexes is limited. Current approaches have been comprised by methodological limitations: Either the protein complexes are highly purified as a result of a long lasting procedure, concomitant with the loss of transient or weak interactors or a rapid (and dirty) one step purification resulting in an enrichment of protein complexes, which makes it almost impossible to distinguish between specific and non-specific. One step methods rely on application of quantitative approaches, enabling this discrimination. Application of quantitative analytical methods to this purpose would not only enable the identification of complex components but would also make the comparison of protein complexes possible. The quantitative comparison of protein complexes composed of proteins associated with human diseases as well as the comparison of the protein complexes altered as a consequence of a mutation of a protein would then be possible. This would provide initial information on the molecular mechanism leading to a specific disease.

The aim of this study was to develop, optimize and apply methods and analytical workflows for the purification and identification of protein complexes as well as the quantitative comparison of protein complexes by mass spectrometry. Since gel based approaches are either limited in resolution like SDS-PAGE or are laborious like 2D-gels, the combination of quantitative mass spectrometry to the analysis of protein complexes offers great opportunities.

<u>Objective 1:</u>

The first aim of this work was the optimization and evaluation of protocols for the purification of protein complexes from mammalian cells by a tandem affinity purification (TAP) tag, consisting of a tandem Strep-tag II and a Flag tag (SF-TAP) previously developed by Johannes Gloeckner. Additionally, a combination of a gel-free and mass spectrometry-based approach had to be established to increase the sensitivity of the method.

For evaluation, the TAP protocols were to be applied on the well-defined network of the MAPK pathway.

Objective 2:

A second analytical approach for the analysis of transient or weakly associated complex components had to be established. The SF-TAP technology is especially useful to get first insights into protein complexes and to identify the "core complex" of a protein. However, it is limited to rather stable complexes. For this aim, a quantitative mass spectrometry-based method had to be established to ensure the highest sensitivity and specificity possible. The method needed to be compatible to the SF-TAP tag technology in order to use the same construct in both approaches.

Objective 3:

The third part of this study was the application of the technology, developed in objective 1 to the analysis of a so far unknown protein complex, the lebercilin protein complex.

Objective 4:

A quantitative mass spectrometry-based method for identifying protein complexes had to be applied to the quantitative protein complex analysis of the model system, lebercilin. The main focus was laid on the identification of alterations in the protein complex due to mutation and thereby get insights into the disease mechanism.

F Materials and methods

1. Materials

1.1 Chemicals

All chemicals were purchased from Sigma-Aldrich (Sigma, Fluka, Aldrich; Taufkirchen, Germany) or from VWR International (Darmstadt, Germany). All cell culture reagents were purchased from Invitrogen (Karlsruhe, Germany). In this study, dH_2O refers to deionized water, ddH_2O ultra-pure water (Milli-Q Biocell, Millipore, Eschborn, Germany).

1.2 General equipment

Analysis scales BP221S	Sartorius, Göttingen, Germany
Autoclave Bioclav	Schütt Labortechnik, Göttingen,Germany
Autoclave Systec 5075 ELV	Systec, Wettenberg, Germany
Cell culture plates 6-well/10cm/14cm	Nunc, Wiesbaden, Germany
Centrifuge 5415D	Eppendorf, Hamburg, Germany
Falcon conical tubes 15/50 mL	BD Bioscience, Heidelberg, Germany
Magnetic stirrer RH basic	IKA Labortechnik, Staufen, Germany
Microspin columns	GE Healthcare, Freiburg, Germany
MILLEX GP; syringe filter unit, 0.22 µm	Millipore, Bedford, USA
Milli-Q Biocell	Millipore, Bedford, USA
Precision scales Basic Plus BP2100	Sartorius, Göttingen, Germany
Safe lock reaction tubes 0.5/1.5/2.0 ml	Eppendorf, Hamburg, Germany
Shaker Duomax 1030	Heidolph Instruments, Schwabach, Germany
Shaker KS260 basic	IKA Labortechnik, Staufen, Germany
Sigma Laboratory Centrifuge 4K15C	Sigma Laborzentrifugen, Osterode, Germany
Sigma Laboratory Centrifuge 6K15	Sigma Laborzentrifugen, Osterode, Germany
SpeedVac SPD111V	Savant, Fisher Scientific, Schwerte, Germany
Steritop-GP filter unit, 0.22 µm	Millipore, Bedford, USA
Ultra-low temperature freezer VIPTM	Sanyo Scientific, IL, USA
Ultrasonic bath Transsonic 310/H	Elma Ultrasonic, Singen, Germany
Ultraspec 3300 pro UV/Vis photometer	GE Healthcare, Freiburg, Germany
Vortex Genie 2	Scientific Industries, VWR, Darmstadt, Germany
Water bath HRB 4 digital	IKA Labortechnik, Staufen, Germany
Zoom Stereomicroscop Nikon SMZ645	Nikon, Amstelveen, Netherlands

1.3   Protein chemistry

Agfa Curix 60 Developer Densitometer	Agfa, Cologne, Germany
GelAir dryer	BioRad, Munich, Germany
GS-710 Calibrated Imaging	BioRad, Munich, Germany
Hybond-LFP membranes	GE-Helthcare, Freiburg, Germany
Hybond PVDF membranes	GE-Helthcare, Freiburg, Germany
Mini Protean 3 for SDS-PAGE	BioRad, Munich, Germany
Mini Trans-Blot cell	BioRad, Munich, Germany
NuPAGE NOVEX Bis-Tris gels	Invitrogen, Carlsbad, USA
Power Supply Power Pac 3000	BioRad, Munich, Germany
Protean II for SDS-PAGE	BioRad, Munich, Germany
Synergy HAT microplate reader	Biotek, Bad Friedrichshall, Germany
Typhoon Trio Variable Mode Scanner	GE-Helthcare, Freiburg, Germany
Trans-Blot SD semi-dry transfer blot	BioRad, Munich, Germany
XCell SureLock Mini-Cell	Invitrogen, Carlsbad, USA

1.4   Protein complex purification

10x buffer E (Desthiobiotin)	IBA, Göttingen, Germany
Anti-Flag M2 Affinity Gel	Sigma-Aldrich, Taufkirchen, Germany
Strep-Tactin Superflow	IBA, Göttingen, Germany
Strep-Tactin Sepharose	IBA, Göttingen, Germany

1.5   Mass spectrometry

1.5.1   Special equipment

MALDI-TOF/TOF mass spectrometer ABI 4700 Proteomics Analyzer	Applied Biosystems, Foster City, USA
LTQ Orbitrap XL	Thermo Fisher Scientific, Bonn, Germany
Inserts for 2 ml vials	Sigma-Aldrich, Taufkirchen, Germany
Probot liquid handling system	Dionex, Idstein, Germany
Ultimate Nano-LC	Dionex, Idstein, Germany
Ultimate 3000 Nano-LC	Dionex, Idstein, Germany

1.5.2   Kits and special reagents

RapiGest^{TM}SF	Waters, Eschborn, Germany

1.6 Mammalian cell culture

1.6.1 Special equipment

CO_2 incubator	Sanyo, Munich, Germany
Laminar flow	BDK, Sonnenbühl-Genkingen, Germany
Liquid Nitrogen Tank Chronos	Messer, Sulzbach, Germany

1.6.2 Special reagents

$^{13}C6$ L-Arginine-HCl	Thermo Fisher Scientific, Bonn, Germany
$^{13}C6$-L-Lysine	Silantes, Munich, Germany
$^{13}C6$ $^{15}N2$-L-Lysine	Silantes, Munich, Germany
$^{13}C6$-L-Arginine	Silantes, Munich, Germany
$^{13}C6$ $^{15}N4$-L-Arginine	Silantes, Munich, Germany
D4-L-Lysine	Silantes, Munich, Germany
FBS dialyzed	PAA, Pasching, Austria
L-Glutamine	PAA, Pasching, Austria
L-Proline	Silantes, Munich, Germany
L-Proline	Thermo Fisher Scientific, Bonn, Germany
L-Arginine	Silantes, Munich, Germany
L-Lysine	Silantes, Munich, Germany
SILAC DMEM	PAA, Pasching, Austria
SILAC Protein Quantification Kit	Thermo Fisher Scientific, Bonn, Germany

1.6.3 Mammalian cell lines

Cell line name	Description	Provider
HEK293	Human embryonic kidney cells	DSMZ, Braunschweig, Germany
HEK293T	Human embryonic kidney cells	Walter Kolch, Glasgow, UK
PC12	Rat pheochromocytoma cells	Walter Kolch, Glasgow, UK

Table 3: Mammalian cell lines

1.7 Molecular biology

1.7.1 Special equipment

ABI Prism 3100 Genetic Analyzer	Applied Biosystems, Foster City, USA
Cycler PTC-225	Germany
GFL Incubator Shaker for E. coli	Burgwedel, Germany
Incubator for E. coli	Memmert, Schwabach, Germany
PCR DNA Engine Tetrad Gradient	MJ Research, BioRad, Munich,
SubCell GT chambers	BioRad, Munich, Germany
UV transiluminator UVT-40M	Herolab, Wiesloch, Germany

1.7.2 Kits and special reagents

Big Dye Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems, Foster City, USA
Gateway BP Clonase II enzyme mix	Invitrogen, Carlsbad, USA
Gateway LR Clonase II enzyme mix	Invitrogen, Carlsbad, USA
Phusion High-fidelity PCR Kit	New England Biolabs, Ipswich,
PureYield Plasmid Midiprep System	Promega, Mannheim, Germany
QIAquick Gel Extraction Kit	Qiagen, Hilden, Germany
QIAquick PCR purification Kit	Qiagen, Hilden, Germany
QIAprep Spin Plasmid Miniprep Kit	Qiagen, Hilden, Germany USA
QuikChange II XL Site-Directed Mutagenesis Kit	Stratagene, La Jolla, USA

1.7.3 *Escherichia coli* strains

DH5α	Invitrogen, Carlsbad, USA
TOP10	Invitrogen, Carlsbad, USA
XL1-Blue supercompetent cells	Stratagene, La Jolla, USA

1.7.4 Enzymes

All restriction enzymes and DNA ladders were purchased from New England Biolabs (Ipswich, USA). T4 DNA ligase and Taq DNA polymerase were purchased from MBI Fermentas (St.Leon-Rot, Germany).

1.7.5 Oligonucleotides

All Oligonucleotides were purchased from Metabion (Martinsried, Germany).

Primer name	Sequence
pcDNA3_fw	5'-GCGGTAGGCGTGTACGGTGGG-3'
pcDNA3_rv	5'-GGGCAAACAACAGATGGCTGGC-3'
pTRE-tight_fw	5'-GAGAACGTATGTCGAGGTAGG-3'
pTRE-tight_rv	5'-GTGATGCTATTGCTTTATTTGTAAC-3'
Seq_BRAF_2	5'-CTGCAAGGTGTGGAGTTAC-3'
Seq_BRAF_3	5'-CAGCGTTGTAGTACAGAAG-3'
Seq_BRAF_4	5'-CCAATGTGCATATAAACAC-3'
Seq_BRAF_5	5'-GTCTACAAGGGAAAGTGGC-3'
Seq_BRAF_6	5'-AATCATCCACAGAGACCTC-3'
Seq_BRAF_7	5'-GGATACCTGTCTCCAGATC-3'

Table 4: Sequencing primer

1.7.6   Plasmids and constructs

1.7.6.1   Plasmids

Plasmid name	Description	Resistance	Provider
pcDNA3	Mammalian expression vector	Amp	Invitrogen
pcDNA3-SF-TAP	SF-TAP expression vector	Amp	JG
pDEST-N-SF-TAP	GATEWAY SF-TAP vector	Amp	Invitrogen/JG
pDONR201	GATEWAY donor vector	Kan	Invitrogen
pTRE-tight	Tet-Off expression vector	Amp	Clontech

Table 5: Plasmids JG: Johannes Gloeckner

1.7.6.2   Constructs

Construct name	cDNA (human)	Plasmid	Provider
14-3-3ε-SF	14-3-3ε	pDEST-N-SF-TAP	KB/JG
B-Raf-SF	B-Raf	pcDNA3-SF-TAP	JG
C-Raf-SF	C-Raf	pcDNA3-SF-TAP	JG
IFT88-SF	IFT88	pDEST-N-SF-TAP	RR
MEK1-SF	MEK1	pDEST-N-SF-TAP	KB/JG
Lebercilin-P493T-SF	Lebercilin-P493T	pDEST-N-SF-TAP	RR
Lebercilin-SF	Lebercilin	pDEST-N-SF-TAP	RR

Table 6: DNA constructs. JG: Johannes Gloeckner; KB: Karsten Boldt (this study)

1.8   Antibodies

1.8.1   Primary antibodies

Anti-	Species	Dilution	Provider
14-3-3 epsilon	Rabbit, polyclonal	1:1000	Santa Cruz
B-Raf (C19)	Rabbit, polyclonal	1:1000	Santa Cruz
B-Raf (H145)	Rabbit, polyclonal	1:1000	Santa Cruz
C-Raf (C12)	Rabbit, polyclonal	1:1000	Santa Cruz
Erk1/2	Rabbit, polyclonal	1:1000	Cell Signaling
Flag-M2	Mouse, monoclonal	1:1000-1:10,000	Sigma-Aldrich
Flag-M2-HRP	Mouse, monoclonal	1:1000-1:10,000	Sigma-Aldrich
GAPDH	Mouse, monoclonal	1:5,000	Millipore
IFT74	Goat, polyclonal	1:1000	Everest
IFT88	Rabbit, polyclonal	1:1000	Doris Bengel
MEk1/2	Rabbit, polyclonal	1:1000	Cell Signaling

Table 7: Primary antibodies. Santa Cruz: Santa Cruz, USA; Cell Signaling: Danvers, USA; Sigma-Aldrich: Taufkirchen, Germany; Millipore: Billerica, USA; Everest: Everest Biotech, Upper Heyford, UK; Eurogentec: Köln, Germany; Doris Bengel: Helmholtz-Zentrum Munich, Germany.

1.8.2 Secondary antibodies

Anti-	Species	Dilution	Provider
Goat-HRP	Donkey, polyclonal	1:7500	Dianova
Mouse-HRP	Goat, polyclonal	1:7500	Dianova
Rabbit-HRP	Goat, polyclonal	1:7500	Dianova
Rabbit-Cy5	Goat, polyclonal	1:1000	GE
Mouse-Cy5	Goat, polyclonal	1:1000	GE

Table 8: Secondary antibodies: Dianova: Hamburg, Germany; GE Healthcare, Freiburg, Germany

1.9 Software and databases

1.9.1 Software

4000 series Explorer 3.5	Applied Biosystems, Foster City, USA
Adobe Illustrator CS	Adobe Systems, Seattle, USA
Adobe Photoshop CS	Adobe Systems, Seattle, USA
Bioworks Browser 3.3.1	Thermo Fisher Scientific, Bonn, Germany
GSP Explorer 3.5	Applied Biosystems, Foster City, USA
ImageJ 1.41	W. Rasband, Maryland, USA
Irfan View	I. Skiljan, Wiener Neustadt, Austria
Mascot 2.2	Matrix Science, Boston, USA
Mascot Daemon 2.2.2	Matrix Science, Boston, USA
MaxQuant	J. Cox, M. Mann, Martinsried, Germany (http://www.maxquant.org/)
MSQuant	P. Mortensen, Odense, Denmark (http://msquant.sourceforge.net/)
R	The R project (http://www.r-project.org)
Xcalibur 2.06	Thermo Fisher Scientific, Bonn, Germany

1.9.2 Databases and online tools

Ciliary Proteome Database	http://www.ciliaproteome.org/
HRPD	http://www.hprd.org/
NCBI	http://www.ncbi.nlm.nih.gov/
NCBI Blast	http://blast.ncbi.nlm.nih.gov/Blast.cgi
NCBI Nucleotide	http://www.ncbi.nlm.nih.gov/entrez/ query.fcgi?db=nucleotide
NCBI Protein	http://www.ncbi.nlm.nih.gov/entrez/ query.fcgi?db=protein
NCBI Pubmed	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi
POINeT	http://poinet.bioinformatics.tw/idGenerator.do
RZPD	http://www.rzpd.de/
STRING	http://string.embl.de/
Swissprot	http://www.expasy.ch/sprot/

Uniprot   http://www.uniprot.org/
Uniref   http://www.ebi.ac.uk/uniref/

2. Methods

2.1 Mammalian cell culture

2.1.1   Growth and maintenance of mammalian cells

2.1.1.1   Routine culture

HEK293, HEK293T and NIH-3T3 cells were routinely grown in growth medium (Dulbecco's modified eagle medium (DMEM) containing 10% fetal bovine serum (FBS), 50 units/ml Penicillin and 0.05 mg/ml Streptomycin) in 10 cm cell culture dishes at 37°C and 5% CO_2 in a cell culture incubator. The FBS was heat inactivated by incubation at 55°C for one hour before use. Cells were split three times a week at a ratio of 1:10. Therefore, the medium was removed; cells were washed once with PBS and incubated with 2.5 ml of trypsin/EDTA solution (0.5 mg/ml trypsin, 0.22 mg/ml EDTA in PBS) per 10 cm dish for 1-5 minutes. Then the same amount of growth medium was added and the cell suspension was centrifuged for 3 minutes at 700 xg. Cells were resuspended in 5 ml of growth medium and 0.5 ml of the suspension was added to 9 ml fresh medium in fresh cell culture dishes.

PC12 cells were grown in growth medium (DMEM containing 10% horse serum, 5% FBS, 100 units/ml Penicillin and 0.05 mg/ml Streptomycin) in 10 cm cell culture dishes at 37°C and 5% CO_2 in a cell culture incubator. The horse serum as well as the FBS were heat-inactivated by incubation at 55°C for one hour. PC12 cells were split twice a week. Therefore, the medium was removed, 5 ml of fresh growth medium was added and cells were rinsed off the plate. 1 ml of the cell suspension was added to a fresh 10 cm dish containing 9 ml fresh growth medium.

2.1.1.2   Generation of cryo stocks

Cells were trypsinized and resuspended in a mixture of 90% FBS and 10% DMSO, frozen at -20°C for one hour and at -80°C over night before they were transferred into liquid nitrogen for long term storage.

2.1.1.3 Collagen coating of cell culture dishes

If the cells were starved prior to harvesting, the culture dishes were coated with collagen before the cells were seeded out. Therefore, collagen was diluted in PBS to a final concentration 50 μg/ml and added into the cell culture dishes so that 5 μg of collagen was used per cm^2 of surface area. The plates were incubated for four hours at 37°C and washed once with PBS before the cells were seeded.

2.1.1.4 Growth and maintenance of SILAC cultures

SILAC cultures were cultivated in the same way as normal cultures with following modifications. In contrast to normal growth medium, $^{13}C6$-L-lysine and $^{13}C6^{15}N4$-L-arginine for double labeling were used for the heavy condition and normal amino acids for the light condition. Triple labeling experiments were done using normal amino acids in the light condition, $^{2}H4$-L-lysine and $^{13}C6$-L-arginine in the medium condition and $^{13}C6^{15}N2$-L-lysine and $^{13}C6^{15}N4$-L-arginine in the heavy condition. All SILAC amino acids were added to a final concentration of 0.1 mg/ml to L-lysine and L-arginine deficient DMEM medium. Proline was added to a final concentration of 0.23 mg/ml and the medium was sterile filtered. Proline was added to prevent arginine to proline conversion, which could impair the quantification. After sterile filtration, dialyzed FBS was added to a final concentration of 10% (v/v), 50 units/ml Penicillin and 0.05 mg/ml strepto-mycin were added. Heavy, medium and light media were prepared in the same way in parallel.

2.1.2 Transient transfection of mammalian cells

HEK293, HEK293T and NIH-3T3 cells were transfected using Effectene, Lipofectamine or the polyethylenimine (PEI) transfection method. The day before transfection, the cells were seeded onto fresh cell culture plates at a confluency of approximately 50%. Depending on the plates used, different amounts of medium, DNA as well as reagents for transfection were used (Table 9). For Effectene mediated transfection, the DNA was diluted in EC buffer, the enhancer was added, the mixture vortexed for 1 second and incubated for 2 minutes at room temperature. Then the Effectene reagent was added and the mixture vortexed for 10 seconds. After incubation at room temperature for 5-10 minutes, the medium was added and the complete mixture was added dropwise to the cells. The cells were

incubated under normal conditions. If the cells needed to be serum starved, the medium was removed, cells were washed with PBS and medium without serum was added.

Culture format	DNA (µg)	Enhancer (µl)	Final volume of DNA in buffer EC (µl)	Volume of Effectene reagent (µl)	Volume of medium to add to the complexes (µl)	Medium added to the cells (µl)
6-well plate	0.4	3.2	100	10	600	1400
10 cm dish	2	16	300	60	1000	9000
14 cm dish	4	32	600	120	2000	18000

Table 9: Effectene transfection. The table shows the reagents and volumes used for Effectene transfections for different cell culture formats.

PC12 cells were transfected using Lipofectamine 2000. The cells were seeded out at a confluence of 60-80% the day before transfection. Depending on the dishes used, the amounts of reagents were adjusted (Table 10). Fresh PC12 growth medium was added to the cells. The DNA and the Lipofectamine 2000 reagent were diluted in DMEM and incubated for 5 minutes at room temperature. The DNA and the Lipofectamine 2000 in Medium were mixed and the mixture was incubated for further 20 minutes at room temperature before it was added dropwise to the cells. For serum starvation, cells were washed once with PBS before serum-free medium was added.

Culture format	DNA (µg)	Lipofectamine 2000 (µl)	Volume of medium to add (µl)	Medium added to the cells (µl)
6-well plate	0.4	10	2x 250	1500
10 cm dish	2.4	60	2x 1500	7000

Table 10: Lipofectamine-2000 transfections. The table shows the volumes and reagents used for the transfection of cell lines with Lipofectamine-2000 for different cell culture dish sizes.

An alternative transfection procedure used was the PEI (polyethylenimine) transfection[146,147]. For this procedure, self made PEI solution was used. For the preparation of the solution, 100 µg of PEI were dissolved in 900 ml, of 100 mM sodium chloride, pH 5.5 by incubation at 80°C over night. The pH was adjusted to 7.4 by the addition of 0.01 N hydrochloric acid (HCl) before 100 mM sodium chloride was added to a final volume of 1000 ml.

The solution was then sterile filtered and stored at 4°C. The transfection procedure was as follows: The DNA was diluted in PEI solution; the mixture was mixed by vortexing for 10 seconds and incubated for 10 minutes at room temperature before it was added dropwise to the cells. The amounts of reagent and DNA used depended on the size of the cell culture dishes as shown in Table 11.

Culture format	DNA (µg)	Volume of PEI reagent (µl)	Medium added to the cells (ml)
6-well plate	0.8	66	1500
10 cm dish	3	300	7000
14 cm dish	8	1000	18000

Table 11: Reagents and volumes used for PEI transfections. The table shows the reagents and volumes used for the transfection of cell lines using the PEI transfection method for different cell culture formats.

2.1.3 Generation of stable cell lines

For the generation of cell lines, stably expressing the protein of interest, cells were transfected and further cultured under normal culture conditions 48 hours post transfection. Then cells were seeded at low density in growth medium containing 1000 µg/ml G-418. Since the pcDNA3 vectors used contain a neomycin resistance gene, stably transfected cells can survive the G-418 treatment. The cells were cultivated for 2-4 weeks and the medium was exchanged every second day. Single colonies were collected and transferred in 24-well plates. Once the cells were confluent, they were transferred into 6-well plates (2 wells per clone). Cells from one well were harvested and the expression was tested by Western blotting. Cells from the other well were further cultured or cryo stocks were generated.

2.1.4 Cell harvesting and generation of protein extracts

Before harvesting, culture medium was removed and cells were washed once with PBS. The PBS was removed and all following steps were done on ice. Lysis buffer (30 mM Tris-HCl, pH 7.4, 150 mM sodium chloride, 0.5% NP-40, protease inhibitor cocktail, phosphatase inhibitor cocktail I and II) was added to the cells. 1 ml was used for 14 cm dishes, 0.5 ml for 10 cm dishes and 0.2 ml per well for 6-well plates. Cells were scraped of the plates using cell scrapers. The lysate was incubated for 20 minutes at 4°C

under constant agitation followed by a centrifugation for 10 minutes at 4°C and 10000 x g. The pellet was discarded and the supernatant used for downstream experiments.

2.2 Protein chemistry

2.2.1 Determination of protein concentration

The protein concentration was determined either by the Bradford method[148] or by using the BCA protein assay[149].

2.2.1.1 Protein determination by the Bradford method

The Bradford method is based on complex formation of proteins with the Coomassie brilliant blue G-250 dye. Binding of the Coomassie dye leads to an absorbance shift from 465 nm in its unbound form to 595 nm in its complexed form. The reaction is mainly based on interaction of the anionic form of the Coomassie dye with arginine and only slightly on its interaction with other basic (histidine, lysine) or aromatic residues (tryptophan, tyrosine, phenylalanine). Although the color response is not entirely linear[150,151], by the use of a standard curve, run with each assay, it is possibly to achieve exact protein concentration determinations with very high sensitivity (1 µg protein/ml reaction)[148]. The reaction is not influenced by the presence of reducing reagents like DTT or β-mercaptoethanol but is sensitive to detergents like sodium dodecyl sulphate (SDS) or strong basic reagents. For the Bradford protein determination, the Bradford reagent stock solution was diluted five fold in water. A standard curve was prepared with bovine serum albumin (BSA) in a concentration range from 0.2 mg/ml to 1 mg/ml. Lysis buffer was added to the BSA standard dilutions so that the same volume of lysis buffer was present in each standard as in the samples of unknown protein concentration. ddH$_2$O with the same amount of lysis buffer was used as a reference. The standard concentrations and the samples were added to 1 ml of the diluted Bradford reagent solution. The reactions were incubated for 5 minutes at room temperature and the absorption was measured in the photometer at 595 nm. The protein concentrations were calculated according to the standard curve.

2.2.1.2 Protein determination by the BCA method

The BCA protein assay is less sensitive to detergents when compared to the Bradford method. Therefore it was especially used for samples with low protein concentrations. The protein concentration determination by the BCA method is based on a two step reaction. The first step is the chelation of copper ions by proteins. Cu^{2+} ions react with the proteins in an alkaline medium and by dissociation of OH^-, the Cu^{2+} ions are reduced to Cu^{1+}. This reaction only takes place with peptides and proteins with a length of at least three amino acids and results in a light blue complex. Since the sensitivity for this reaction is low (5 mg/ml), a second step is added in the BCA protein assay. In this second step, the chelation of one Cu^{1+} ion with two bicinchoninic acid molecules forms a purple reaction product. This second reaction is highly sensitive and results in a approximately 100 fold more sensitive detection than the first step[149].

For the determination of the protein concentration by the BCA assay, reagent A and reagent B of the kit were mixed in a ration of 40 to 1. Standard concentrations of BSA were prepared as described for the Bradford method (2.2.1.1). 20 µl plus the volume of lysis buffer corresponding to the volume of sample to be measured and the samples were added to 200µl of the BCA reagent mixture. The reaction was incubated for 10 minutes to one hour at 37°C and the absorption was measured at 562 nm. The protein concentrations of the samples were calculated according to the standard curve.

2.2.2 Protein concentration and removal of interfering substances

For diluted samples and for sample preparation for in-solution tryptic digestion, the protein samples had to be concentrated and interfering substances needed to be removed. In order to accomplish this, the samples were either concentrated using centrifugal units or precipitated using the methanol/chloroform protein precipitation method[152].

2.2.2.1 Protein concentration using centrifugal unit

Microcon centrifugal units with a cut-off of 3 or 10 kDa were used for protein concentration by centrifugation. The diluted protein samples were transferred into the Microcon centrifugal units which was then centrifuged at 14000 x g at 4°C until the desired volume was reached. Then the

50

Microcon unit was transferred to a fresh tube bottom up and the sample was recovered by centrifugation at 3000 x g for 3 minutes.

2.2.2.2 Protein precipitation

In order to precipitate proteins, the methanol/chloroform protein precipitation method[152] was applied. It is insensitive against contaminations by salts, lipids and detergents and therefore the method of choice for removal of these substances from protein samples, which are to be analyzed by mass spectrometry. Additionally, this method precipitates protein quantitatively over a wide range of protein concentrations.

The protein sample was diluted in 4-fold the sample volume of methanol, vortexed and centrifuged for 30 seconds at 9000 x g. One fold the sample volume of chloroform was added, vortexed and centrifuged for 30s at 9000 x g before three fold the sample volume of ddH$_2$O was added, vortexed for 5 seconds and centrifuge as before to achieve phase separation. The upper phase was removed and three fold the sample volume of methanol was added, vortexed and centrifuged for 2 minutes at 16100 x g. The supernatant was discarded and the pellet was air-dried for 15 minutes.

2.2.3 SDS-PAGE

Protein samples were separated according to their molecular weight by SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) for downstream analysis like Western blotting or protein staining[153]. Before the proteins were separated through application of an electric field in a polyacrylamide gel, Laemmli buffer containing SDS (sodium dodecyl sulfate) was added. SDS binds to the proteins and adds negative charges, which mask the intrinsic charge of the proteins. This leads to a charge-to-mass ratio that is almost constant for all proteins. In combination with the reduction of disulfide bonds by a reducing agent and heat, a migration speed relative to the mass of the protein only, is achieved. The discontinuous gel system used in this study consisted of a separating gel, overlaid by a stacking gel. By the use of a stacking gel which is more acidic compared to the separating gel and consists of a lower acrylamide concentration, the proteins are concentrated in a sharp lane before the separation takes place.

Gels were either casted or ready to use NuPAGE gels were used, especially for pre-fractionation prior to mass spectrometry. Gel casting was performed by the use of the Mini Protean 3 system. The volumes of solutions used are shown in Table 12. The solutions were mixed and filled between a 1 mm spacer plate and a short glass plate. After polymerization, the separating gel was overlaid with the stacking gel. The volumes of solutions used are shown in Table 12. A comb of 1 mm thickness with 10 or 15 wells was stuck between the glass plates before the gel was polymerized. The casted gels were placed in a Mini Protean 3 chamber and the chamber was filled with TGS-electrophoresis buffer (25 mM Tris, 192 mM Glycine, 0.1% (w/v) SDS). The NuPAGE gels were placed in the NuPAGE chamber and the chamber was filled with MES electrophoresis buffer. For both systems, the comb was removed and the wells were rinsed with electrophoresis buffer. Before the protein extracts were loaded in the wells, 5-fold SDS-sample buffer (5% (w/v) SDS, 250 mM Tris-HCl pH 6.8, 50% (v/v) glycerol, 500 mM β-mercaptoethanol, 0.025% (w/v) bromphenol blue) was added to the samples in a 1 to 5 ratio and the samples were incubated at 96°C for 15 minutes. The gel run was started with a constant potential of 80 V, after 20 minutes, the potential was increased to 130 V. The gel run was stopped when the bromphenol blue front reached the end of the separating gel.

Separating gel (20 ml)	8%	10%	12%
ddH$_2$O	9.4 ml	8.0 ml	6.7 ml
30% (w/v) Acrylamide:bisacrylamide (37.5:1)	5.3 ml	6.7 ml	8.0 ml
1.5 M Tris-HCl, pH 8.8	5 ml	5 ml	5 ml
20% (w/v) SDS	100 µl	100 µl	100 µl
TEMED	20 µl	20 µl	20 µl
10% (w/v) APS	150 µl	150 µl	150 µl

Stacking gel (10 ml)	
ddH$_2$O	6 ml
30% (w/v) Acrylamide:bisacrylamide (37.5:1)	1.3 ml
1.5 M Tris-HCl, pH 6.8	2.5 ml
20% (w/v) SDS	50 µl
TEMED	10 µl
10% (w/v) APS	75 µl

Table 12: Reagents and volumes used for casting SDS-PAGE gels. The tables show the reagents and volumes used for casting SDS-ÜAGE gels of different acrylamide concentrations and for casting the stacking gel.

2.2.4 Staining of SDS-PAGE gels

SDS-PAGE gels were stained with different methods. The method applied was dependent on the application and the sensitivity necessary to detect the protein bands. For prefractionation experiments, Coomassie staining was applied because of its optimal compatibility to mass spectrometric analysis. Although the silver staining method described here as well as the colloidal Coomassie staining method are in principle compatible with mass spectrometric analysis, as the sensitivity of the Coomassie staining method is sufficient for this application in most cases. If higher sensitivity was necessary, the colloidal Coomassie staining procedure was applied. For the gel-based analysis of protein complexes, silver staining was used due to its high sensitivity.

2.2.4.1 Coomassie staining

The Coomassie staining method is less sensitive when compared to colloidal Coomassie or silver staining. The detection limit is around 100 ng protein per band. It is based on the binding of the Coomassie brilliant blue R-250 dye to proteins.

After gel electrophoresis, the gel was placed in Coomassie staining solution (50% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie brilliant blue R-250) for 5 minutes to 1 hour. The staining solution was removed and fixation solution (50% (v/v) methanol, 10% (v/v) acetic acid) was added to remove background staining. The solution was exchanged several times until no background staining remained.

2.2.4.2 Colloidal Coomassie

The colloidal Coomassie staining is more sensitive than Coomassie but not as sensitive as silver staining. In acidic media containing ammonium sulphate, the Coomassie brilliant blue G-250 dye forms micro-precipitates[154,155]. Thereby the amount of free dye is very low, resulting in low background staining but high sensitivity since the colloids act as reservoirs and sufficient dye is available to occupy all binding sites on the proteins.

After electrophoresis, the gels were incubated twice for 30 minutes in fixation solution (20% (v/v) methanol), washed three times for 30 minutes in 10% (v/v) phosphoric acid and equilibrated once for 20 minutes in

equilibration solution (10% (v/v) phosphoric acid, 20% (v/v) methanol and 10% (w/v) ammonium sulphate). 1.2% (v/v) of a 2% (w/v) Coomassie brilliant blue G-250 solution in dH_2O was added to the equilibration solution and the gel was stained in this solution for 2-24 hours.

2.2.4.3 Silver staining

Silver staining is based on the reduction of silver ions to silver by formaldehyde under alkaline conditions[156,157]. The silver binds to and forms complexes with proteins, leading to the visualization of the proteins. The sensitivity of silver staining is 10-100 fold higher than Coomassie staining (1-10 ng protein per band).

After electrophoresis, the gels were fixed twice in 50% (v/v) methanol, 12% (v/v) acetic acid, 0.05% (v/v) formaldehyde (37%) for 15 minutes before they were washed three times for 10 minutes in 50% (v/v) ethanol. Followed a short incubation in 0.2 g/l sodium thiosulphate, the gels were washed three times shortly with dH_2O. Staining was performed for 20 minutes with silver staining solution (2 g/l silver nitrate, 0.075% (v/v) formaldehyde (37%)). Then the gel was washed three times shortly in dH_2O before the developing solution (60 g/l sodium carbonate, 5 mg/l sodium thiosulphate, 0.05% (w/v) formaldehyde (37%)) was added. The gel was developed for 1-10 minutes until clear bands were visible. The developing was stopped by exchanging the solution with fixation solution when the desired staining was achieved but before background staining was observed.

2.2.4.4 Drying gels

The stained gels were equilibrated for 10 minutes in preservation solution (20% (v/v) ethanol, 20% (v/v) glycerol) before they were placed air bubble free between two cellophane foils and air dried.

2.2.4.5 Digitalizing gels and films

Stained gels and films were digitalized using a GS-710 calibrated imaging densitometer and Adobe Photoshop CS. Scans were performed at a resolution of 300 dpi.

2.2.5 Western blot analysis

For detection of individual proteins in complex mixtures, the mixture was separated by SDS-PAGE and proteins were transferred onto Hybond-P polyvinylidene difluoride (PVDF) membranes[158]. This method enables the detection of proteins of interest by specific antibodies. For this purpose, the gel was incubated for 5 minutes in anode II buffer (0.3 M Tris) while the PVDF membrane was activated in methanol by incubation for 1 minute and washed with anode II buffer. The transfer was done in a semi-dry blotting apparatus. Three filter papers were wetted in anode I buffer (0.025 M Tris) and laid on the anode plate. Two with anode II buffer wetted filter papers were laid on top followed by the PVDF membrane. The gel was laid air bubble free on the membrane and three filter papers, wetted with cathode buffer (40 mM ε-aminocapronicacid, 0.1% SDS) on top of the gel. Possible air bubbles were removed and the cathode plate was placed on top of the staple. The transfer was performed for two hours with 50 mA per gel and a maximum voltage of 20 V. After the transfer, the membrane was incubated with blocking buffer for one hour. Depending on the primary antibody used for detection, the blocking buffer was either 5% non fat dry milk powder in TBST (30 mM Tris-HCl, pH 7.4, 150 mM sodium chloride, 0.1% Tween20) or 5% BSA in TBST. Then the membrane was incubated over night with the primary antibody. The antibodies were diluted in blocking buffer in different dilutions, depending on the antibody as shown in Table 7. After incubation over night at 4°C and under constant agitation, the membranes were washed three times for 15 minutes with TBST followed by incubation with the secondary, horse radish peroxidase (HRP) coupled antibody. The secondary antibodies were diluted 1:15000 in blocking buffer with non fat dry milk powder. After one hour incubation at room temperature under constant agitation, the membranes were washed three times for 15 minutes with TBST before the HRP was detected with ECL plus reagent. The chemiluminescent signals were detected by Hyperfilm ECL films which were illuminated by incubation on the membranes. Depending on the signal strength, the illumination was done between 5 seconds and one hour.

2.2.5.1 Quantitative Western blot analysis

For quantitative Western blot analysis, the samples were separated and transferred as described above with one modification. Instead of normal PVDF membranes, low fluorescent membranes (Hybond-LFP) were used. Primary antibodies were diluted as described in Table 7. Cy5 coupled, secondary antibodies were used as described in Table 8. Incubations and washing steps were performed as described for not quantitative Western blots with following modification: For all buffers PBS was used instead of TBS. After the three final washes, the membranes were washed three times with PBS without Tween to reduce background. The membranes were then air dried and the fluorescent signal digitalized using the Typhoon Trio Variable Mode Imager. The protein bands were quantified using ImageJ.

2.3   Protein complex purification

To identify components of protein complexes, these complexes must be enriched or purified. The enrichment or purification can be accomplished by different methods. Depending on the analysis method used for the identification of complex components, the need for the degree of purity of the complexes differs strongly. For direct identification of complex components by LC-MS/MS, the purity should be as high as possible. Even though controls can be used to distinguish between specific or unspecific identifications, the lack of quantitative information hinders the identification of complex components with high certainty. For quantitative approaches or directed identification of specific components, a quick enrichment with relatively low purity can be enough and even be helpful to identify less stable or transient complex components. In the following chapters, methods for enrichment and purification of protein complexes are described, mainly based on the ectopic (over-) expression of proteins, fused to a peptide tag.

2.3.1   Strep-Tactin affinity purification

The Strep-Tactin purification is a quick and efficient protein complex enrichment method that can be sufficient for quantitative approaches or directed identification of complex components.

The Strep-Tactin purification system is based on the Strep system developed by Schmidt and Skerra[52-54]. The Strep peptide was developed by stepwise engineering of a peptide binding to streptavidin[53,159]. However, this tag was only suitable for C-terminal tagging of proteins. For that reason an optimized version, the Strep-tag II was developed, overcoming this disadvantage[52]. To improve the binding affinity of the Strep-tag II to streptavdin, Voss and Skerra engineered a novel variant of streptavidin by mutation of its flexible loop thereby enhancing the relatively low affinity of streptavidin to strep ($2.7 \times 10^4 \, M^{-1}$) by at least one order of magnitude. This is still low enough to enable elution by competition with desthibiotin[160].

For the Strep-Tactin purification, 20-50 µl Strep-Tactin-Superflow beads per 14 cm dish of cells (5×10^7 cells) were washed three times with 1 ml of lysis buffer and added to the protein extracts of cells expressing the SF-TAP tagged proteins. Depending on the downstream application and the abundance of the bait protein and complex components, the amount of cells used for the purification varied between one and five 14 cm dishes, corresponding to 5×10^7 to 2.5×10^8 cells in 1-5 ml of lysis buffer. The mixture was incubated for 1 hour at 4°C under gentle agitation. Then the beads with bound protein complexes were centrifuged for 1 minute at 8000 x g before the supernatant was discarded and the beads were transferred to micro-spin columns. The beads were then washed three times with 500 µl of wash buffer (TBS containing 0.1% NP-40, phosphatase inhibitor cocktail 1 and 2). The complex was eluted by incubation of the resin in four fold the resin volume of Strep-Tactin elution buffer for 10 minutes under gentle agitation at 4°C.

2.3.2   Flag affinity purification

Flag purifications can be applied to the same applications as Strep-Tactin purifications. The efficiency and the purification time necessary are similar to the Strep-Tactin purification[51]. In case of direct in-solution digest of one-step purified protein complexes, it is the method of choice because in contrast to Strep-Tactin purifications, no desthiobiotin is used for elution. Since desthiobiotin binds to the C18 resin in reverse phase separations, it greatly reduces the capacity of this chromatography matrix. This results in decreased detection of peptides by subsequent mass spectrometry. The major difference between the Strep-Tactin and the Flag system is that the

Flag system is based on binding of the Flag tag to the anti-Flag-M2 antibody. Originally, the Flag tag and the first anti-Flag antibody (anti-Flag-M1) were engineered by Hopp and colleagues[161]. The binding of the antibody to the tag was Calcium-dependent. For that reason, a second antibody, the anti-Flag-M2 antibody was produced[48]. Its binding is not dependent on calcium; therefore the elution can not be accomplished by a chelating reagent like EDTA. Elution needs to be done either under acidic conditions or by competition using the Flag peptide.

10-25 µl of Flag-M2-sepharose resin per 14 cm dish of cells used for the purification were washed three times with 1 ml of lysis buffer and added to the protein extract of cells expressing SF-TAP tagged proteins. Depending on the downstream application and the abundance of the bait protein and complex components, the amount of cells used for the purification varied between one and five 14 cm dishes, corresponding to 5×10^7 to 2.5×10^8 cells in 1-5 ml of lysis buffer. The mixture was incubated for 1 hour under gentle agitation at 4°C. The beads with bound protein complexes were centrifuged for 1 minute at 8000 x g before the supernatant was discarded and the beads were transferred to micro-spin columns. The beads were then washed three times with 500 µl of wash buffer. The complexes were eluted by incubation of the beads in 4-fold the beads volume of Flag elution buffer (200 µg/ml Flag peptide in wash buffer) for 10 minutes under gentle agitation at 4°C.

2.3.3 Tandem affinity purification

The tandem affinity purification (TAPo) method was initially developed and applied for purification of protein complexes from yeast[33,34]. The TAP tag is composed of two IgG binding domains of protein A, a tobacco etch virus cleavage protease site (TEV), which is necessary to elute the protein from the protein A resin and a calmodulin binding peptide (CBP).

For TAPo purifications, 50 µl of IgG resin per 14 cm dish were used for expression of the TAP tagged proteins. Before lysis, the IgG resin was washed three times with lysis buffer and added to protein extracts of cells expressing TAPo tagged proteins. Depending on the down stream application and the abundance of the bait protein and complex components, the amount of cells used for the purification varied between one and five 14 cm dishes, corresponding to 5×10^7 to 2.5×10^8 cells in 1-5 ml of lysis buffer. After incubation for one hour at 4°C under gentle

agitation, the beads with bound protein complexes were centrifuged for 1 minute at 8000 x g before the supernatant was discarded and the beads were transferred to micro-spin columns. The beads were then washed three times with lysis buffer followed by three washes with TEV cleavage buffer (50 mM Tris-HCl pH 8.0, 150 mM sodium chloride, 0.5 M EDTA, 1 mM DTT). Proteins were eluted by protease cleavage with 100 units of AcTEV protease in 400 µl TEV cleavage buffer for 1 hour at 16°C under gentle agitation. The eluate was supplemented with 1.2 ml calmodulin binding buffer (10 mM Tris-HCL pH 8.0, 150 mM sodium chloride, 10 mM β-mercaptoethanol, 1 mM magnesium acetate, 1 mM imidazole, 2 mM $CaCl_2$, 0.1% NP-40, protease inhibitor cocktail w/o EDTA, phosphatase inhibitor cocktail I and II) and 6 µl of 1 M $CaCl_2$. 25 µl of calmodulin resin per 14 cm dish of cells, which were washed three times with calmodulin binding buffer, were added to the mixture. After incubation for one hour at 4°C under gentle agitation, the mixture was transferred to micro-spin columns and was washed three times with calmodulin binding buffer. Proteins were eluted four times with calmodulin elution buffer (10 mM Tris-HCl pH 8.0, 150 mM sodium chloride, 10 mM β-mercaptoethanol, 1 mM Magnesium acetate, 1 mM imidazole, 2 mM EGTA, 0.1% NP-40) by incubation for 5 minutes each at 4°C under gentle agitation.

2.3.4 SF-TAP purification

The SF-TAP procedure is a combination of a Strep-Tactin purification followed by a Flag-purification. For that reason, the tag is a combination of a Flag tag with a tandem Strep-tag II separated by linker sequences to optimize the efficiency of the single tags (Figure 8A)[162].

As the TAPo tag, the SF-TAP purification is applied for the enrichment of protein complexes to high purity. Compared to the TAPo tag (~25 kDa), it is very small (~5 kDa). Because of the combination of tags with medium affinity, no proteolytic cleavage is necessary for eluting the proteins (Figure 8B). Avoidance of this step accelerates and simplifies the procedure significantly. The calmodulin binding peptide was removed to avoid interference of the tag or the calcium used for elution with cellular signaling mechanisms and protein interactions[163,164]. Because of the high purity, resulting in a low background due to unspecific bound proteins, the eluate is suitable for direct analysis by mass spectrometry. This is the major advantage of TAP purifications compared to one step purification

methods, which exhibit a high background of non-specifically bound proteins. Therefore discrimination between specifically and non-specifically bound proteins is only possible by combining one-step purification strategies with quantitative methods[60].
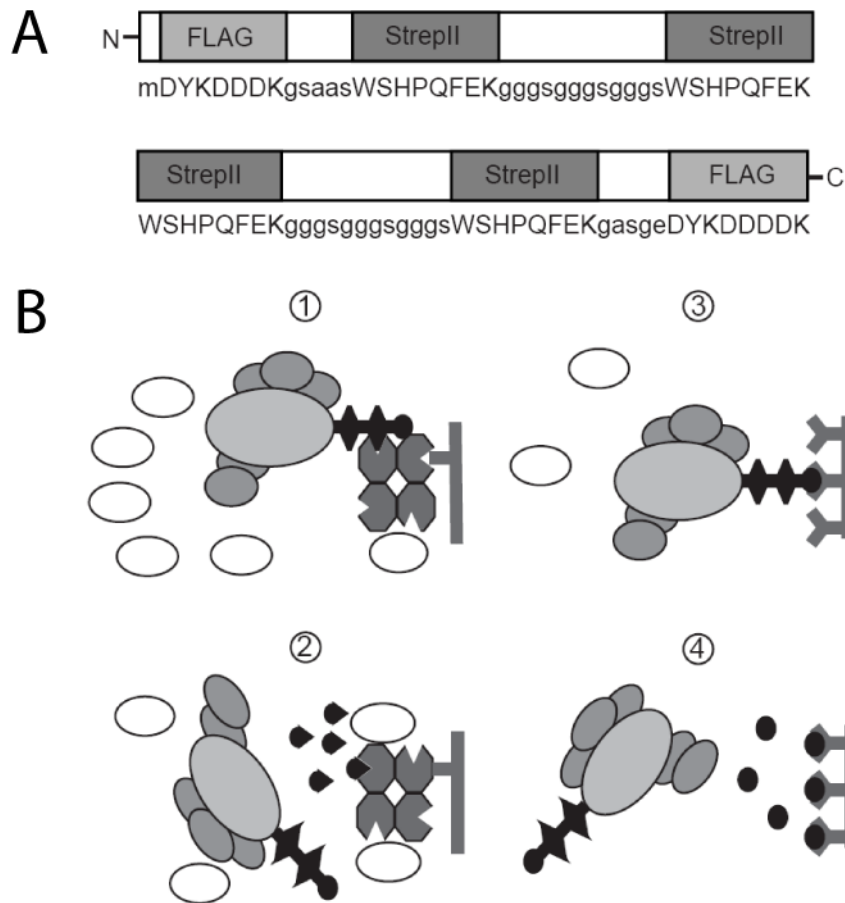

Figure 8: The SF-TAP tag and the SF-TAP method. A) Structure and sequence of the SF-TAP tag for N-terminal fusion (upper panel) and C-terminal fusion (lower panel). In both cases, the first Strep-tag II moiety of the SF-TAP tag is fused protein of interest (POI), followed by a linker sequence, the second Strep-tag II, another linker sequence and the Flag tag. B) Purification procedure for which the SF-TAP tagged POI is bound to the Strep-Tactin matrix by the tandem Strep-tag II in the first step and eluted by competition with desthiobiotin (I). The eluate is then incubated with the Flag matrix and the bound proteins are eluted again by competition using the Flag peptide (II). This figure is reprinted with the friendly permission of Springer[165].

As for the Strep-purification, 50 µl of Strep-Tactin sepharose resin per 14 cm dish of cells used were washed three times with 1 ml of lysis buffer and added to the protein extract of cells expressing SF-TAP tagged proteins. Depending on the down stream application and the abundance of the bait protein and complex components, the amount of cells used for the purification varied between one and five 14 cm dishes, corresponding to

5×10^7 to 2.5×10^8 cells in 1-5 ml of lysis buffer. The mixture was incubated for one hour at 4°C under gentle agitation. The Strep-Tactin resin was collected by centrifugation at 7000 x g for 1 minute, the supernatant was discarded and the resin was transferred to micro-spin columns. In the micro-spin columns, the resin was washed three times with 500 μl of wash buffer and the bound complexes eluted with 500 μl of Strep-Tactin elution buffer by incubation for 10 minutes at 4°C under constant agitation. The eluate was transferred to a micro-spin spin column, containing 25 μl of washed Flag resin (three times in 1 ml of lysis buffer) per 14 cm dish of cells used. The mixture was incubated for one hour at 4°C under gentle agitation. The resin was collected by centrifugation at 7000 x g for 1 minute at 4°C and washed three times with 500 μl of wash buffer. The protein complexes were eluted by incubation with 200 μg/ml Flag peptide in TBS for 10 minutes under gentle agitation at 4°C.

2.4   Mass spectrometry

2.4.1   Sample preparation

Sample preparation is a critical step in analyzing samples by mass spectrometry. The buffers and reagents used for preparing the samples must be compatible with the downstream analysis method. To allow efficient analysis of the sample components, the preparation must be optimized for every analysis strategy. Following the purification of protein complexes or any other step for enrichment or modification of proteins, the sample preparation includes the proteolytic cleavage of proteins prior to mass spectrometric analysis. This is one essential step since the analysis of intact proteins is difficult and by far not as efficient as the analysis of peptides.

A second essential step in the sample preparation is the reduction of the sample complexity. The extent of reduction necessary is dependent on the initial complexity of the sample as well as the mass spectrometric analysis method used. For MALDI analysis, the complexity needs to be reduced to one or few proteins in the sample whereas for a LC-MS/MS analysis on an LTQ OrbitrapXL, tens to hundreds of proteins can be analyzed in a single run. Therefore, MALDI analysis is mostly used for in-gel digested samples of bands from SDS-PAGE or 2D-Gel spots whereas in-solution digestion in combination with LC-MS/MS analysis is used for protein complexes,

purified by TAP approaches. The pre-fractionation approach is much more laborious compared to the in-solution digestion but in many cases necessary to reduce the complexity of the sample to be able to analyze the sample in depth. The major drawback is that this procedure involves several additional sample manipulation steps, especially the in-gel-digestion leading to sample loss as well. For this reason pre-fractionation requires sufficient material to start with.

2.4.1.1 In-gel digestion

This approach is used for samples resolved by SDS-PAGE and stained by either silver or coomassie staining. Before excising the gel bands with a clean scalpel, the stained gels were washed three times for 30 minutes with ddH$_2$O to remove interfering substances like methanol or formaldehyde. The excised gel bands were cut to plugs of 1 mm^3. The gel plugs were transferred into 96 well plates for further treatment. The pieces were washed three times for 15 minutes with 200 µl ddH$_2$O and destained. The destaining procedure was dependent on the staining method used. Silver stained gels were destained by incubation in 200 µl of a 1:1 mixture of 100 mM sodium thiosulfate and 30 mM potassium ferricyanide for up to 10 minutes under gentle agitation. The destaining solution was removed as soon as the silver staining was not visible any longer. Then the gel plugs were washed three times for 5 minutes with 200 µl of ddH$_2$O under gentle agitation. Coomassie stained plugs were destained by two incubations in 200 µl of 40% acetonitrile for 15 minutes followed by incubation in 100% acetonitrile for 5 minutes. If the coomassie staining was still visible, the destaining procedure was repeated. After destaining, all gel plugs were incubated for 15 minutes at 60°C in 100 µl of 5 mM dithiothreitol (DTT), cooled down to room temperature and incubated in 100 µl of 25 mM 2-iodacetamide for 45 minutes in the dark. Then the gel plugs were washed twice for 15 minutes with 40% acetonitrile and dehydrated by incubation in 100% acetonitrile for 5 minutes. After removing the acetonitrile solution and air-drying the plugs for 15 minutes, 10-30 µl of trypsin solution (10 ng/µl sequencing grade trypsin in 50 mM ammonium bicarbonate) was added. The tryptic digest was performed at 37°C over night. The resulting peptides were acidified by addition of 5-15 µl of 1% TFA and the supernatant was transferred into a fresh 96 well plate or into 0.5 ml tubes. Then 70 µl of 40% acetonitrile, 0.5% trifluoroacetic acid

(TFA) were added to the gel plugs and incubated for 15 minutes under agitation. The second supernatant was pooled with the first and 70 µl of 99.5% acetonitrile, 0.5% TFA were added to the gel plugs. After 15 minutes of incubation, the third supernatant was pooled with the first two. The supernatants were dried in a speed vac and the samples were stored at -20°C.

2.4.1.2 In-solution tryptic digestion

In-solution digestion was used for SF-TAP purified protein complexes since the amount of proteins in these samples was low enough to allow complete or almost complete analysis of the proteins with direct analysis by LC-MS/MS. Before the proteolytic cleavage by trypsin, the protein samples were precipitated by the methanol-chloroform precipitation procedure and redissolved in 20µl of 50 mM ammonium bicarbonate containing 0.2% RapiGest. Then 1 µl of 100 mM DTT was added and the samples were incubated for 15 minutes at 60°C. The samples were cooled down to room temperature and 1 µl of 300 mM 2-iodacetamide was added before the samples were incubated for 30 further minutes in the dark. 2-4 µl of 0.5 µg/µl trypsin solution (sequencing grade) was added and the samples were incubated over night at 37°C. To cleave the RapiGest reagent, the sample was acidified by addition of 2 µl of 32% HCl and incubated for 20 minutes at room temperature in the dark. The sample was transferred to inserts and centrifuged for 15 minutes at 4°C and 16,000 xg. The interphase was transferred to a fresh tube and directly subjected to LC-MS/MS analysis.

2.4.1.3 Pre-fractionation by SDS-PAGE

To allow an efficient analysis of highly complex samples, the samples were prefractionated by SDS-PAGE. The protein samples were either precipitated and redissolved in Laemmli buffer or concentrated to a volume of 20 µl using 10 kDa cut-off centrifugal units. 5 µl of 5-fold Laemmli buffer were added. The samples were incubated for 15 minutes at 96°C and then loaded and separated on a 10% NuPAGE gels. The gel run was stopped after a separation distance of 0.5-2 cm was reached, depending on the complexity of the sample. Then the gels were stained by coomassie staining. After the staining procedure, the lanes were excised, fractionated to 3-10 bands and subjected to in-gel digestion.

2.4.2   MALDI TOF/TOF mass spectrometry

For MALDI TOF/TOF analysis, a 4700 Proteomics Analyzer was used. This mass spectrometer is composed of a matrix-assisted laser disorption ionization (MALDI) ion source and a time of flight analyzer. In addition, it is equipped with a collision cell that enables ion fragmentation.

MALDI is a soft ionization technique. For the ionization, the analytes are diluted in a matrix and co-crystallized with this matrix on a MALDI target plate. To ionize the analytes, a pulsed laser beam is subjected to the sample-matrix co-crystallization. The matrix absorbs the laser energy, is vaporized and ionized. During this process, the ionization of the analytes is mediated by the matrix ions. The ions are then accelerated in an electric field and deflected in the flight tube before they are detected by an electron multiplier detector. By using molecules of known mass and measuring the time they need to reach the detector, the mass spectrometer is calibrated. Due to this calibration, the mass of unknown molecules can be calculated by the time they need to reach the detector. For tryptic cleavages of proteins, this process leads to a defined pattern of masses which is called peptide mass fingerprint (PMF). The PMF can be correlated to a database of *in silico* tryptic cleavages of proteins and thereby the unknown protein can be identified. In addition to the PMF, selected peptides can be fragmented and the fragments can be analyzed. Therefore a single peptide ion is isolated in the timed ion selector and is fragmented by collision induced dissociation (CID). For that, a defined concentration of large nitrogen atoms is added to the peptide ions. Due to their kinetic energy, collision of the peptide and nitrogen atoms leads to fragmentation of the peptides. Depending on the breakpoint in the ion and the fragment maintaining the charge, fragments of different sizes and types are formed. N-terminal ions are named a, b and c ions, C-terminal ions x, y and z. If the breakpoint lays directly C-terminal if the amino acid residue, x and a ions are formed, if the peptide bonds break N-terminal of carboxyl group, b and y ions and if the bonds break N-terminal of the amine group, c and z ions arise. The fragments are accelerated and their mass is determined by TOF measurement. The fragmentation pattern originating from this process and the mass of the peptide ion they originate from is used for identification of the peptide sequence by correlation with a database of *in silico* calculated fragmentation patterns.

The PMF and the sequence information obtained by fragmentation are both used for identifying the unknown protein.

2.4.2.1 Analysis of in-gel digested samples

To analyze tryptic cleavages obtained by in-gel digestion, the dried samples were dissolved in 5-10 µl of 2% acetonitrile and 0.5% TFA by incubation for 15 minutes at 4°C and regular vortexing. 1 µl of each dissolved sample was spotted on a MALDI target plate and mixed by gentle up and down pipetting with 1 µl of 5 mg/ml alpha-cyano-4-hydroxy-cinnamic acid in 70% acetonitrile and 0.5% TFA. The mixture was air-dried to allow uniform crystal formation. Additionally, peptide standards were added on six spots on the MALDI target plate for calibration. Then the plate was transferred into the mass spectrometer and the positions of the spots were aligned. The laser intensities were optimized for MS and MS/MS separately to enable optimal sensitivity for PMF and fragment spectra. Then the mass spectrometer was calibrated on the 6 peptide standard spots by automatic plate model calibration. The analysis of the samples was performed in automatic data dependent mode. For each sample spot, spectra resulting from 4000 laser shots were summed up and the five most intense peptide ions were selected for CID fragmentation. For fragmentation analysis, spectra resulting from 5000 laser shots were summed up.

2.4.2.2 LC-MALDI analysis

In-solution digested samples were separated by an Ultimate nano-LC system. The nano-LC system was equipped with a nano trap column (100 µm i.d. × 2 cm, packed with Acclaim PepMap100 C18, 5 µm, 100 Å) and an analytical column (75 µm i.d. × 15 cm, Acclaim PepMap100 C18, 3 µm, 100Å). The sample was automatically injected and transferred into a 25 µl capillary loop before it was loaded onto the trap column at a flow rate of 30 µl/minute in 5% buffer B (80% acetonitrile, 0.08% TFA in HPLC grade water) and 95% buffer A (5% acetonitrile, 0.1% TFA in HPLC grade water). After 5 minutes the peptides were eluted and separated on the analytical column by a gradient from 5% to 50% of buffer B at 200 nl/minute flow rate. Depending on the complexity of the sample, 20-120 minute gradients were used. Remaining peptides were eluted by a short gradient from 50% to 100% buffer B in 5 minutes. The solution with

the eluting peptides was mixed in a 1:4 ratio with 2.5 mg/ml of alpha-cyano-4-hydroxycinnamic acid in 70% acetonitrile and 0.5% TFA % and automatically spotted onto a MALDI target plate by a Probot liquid handling system in 20 second fractions. The fractions were air dried and 6 peptide standards for calibration were spotted on the plate manually. The plate was transferred into the mass spectrometer, aligned and the mass spectrometer was calibrated by automated plate model calibration. For each spot spectra resulting from 4000 laser shots were summed up and up to 12 peptide ions were selected for fragmentation. The selected ions were isolated and fragmented by CID. Spectra resulting from 5000 laser shots were summed up.

2.4.2.3 Spectrum filtering and database searching

All MALDI data were processed using the 4000 Series Explorer and the GPS Explorer Software. For peptide mass fingerprint, cluster area signal to noise (S/N) optimization was performed and monoisotopic masses were used. The limit for peak detection was set to a S/N ratio of 5. The MS and MS/MS spectra were submitted to Mascot using the GPS Explorer software and analyzed by peptide mass fingerprint analysis. Cystein carbamido-methylation was selected as fixed modification and methionine oxidation as variable modifications. The mass tolerance for peptides was set to 65 ppm, the fragment mass tolerance to 0.2 Da. For ion fragment analysis, cluster area S/N optimization was performed and monoisotopic masses were used. A limit for peak detection of a S/N ratio of 5 was used. MS/MS database search was performed with the GPS Explorer Software in combination with Mascot[166] using the Swissprot database. The search was conducted with the enzyme trypsin allowing one missed cleavage. No fixed modifications were used while methionine oxidation and serine, threonine and tyrosine phosphorylation were allowed as variable modification. The peptide mass tolerance was set to 65 ppm and the fragment mass tolerance to 0.2 Da.

Protein identifications from analysis of in-gel digested samples were considered as significant if they were identified by a combined Mascot score of at least 66 and if at least one significant peptide (Mascot score >27) was identified by fragment analysis. Protein identifications from LC-MALDI analysis were considered if at least two peptides were significantly identified (Mascot score > 27). Protein isoforms were considered only if at least one additional, unique and significant peptide was identified.

2.4.3   LC-MS/MS on the LTQ OrbitrapXL

The LTQ OrbitrapXL is a hybrid mass spectrometer, composed of a nano spray ion source, a linear ion trap, a C-trap and an Orbitrap mass analyzer (Figure 9)[167]. The ions are produced by nano spray ionization. The solvent containing the ions is charged in the capillary needle of the nano-Spray ion source by administration of a potential to the solution, containing the peptides. Because of the charge, the solution forms an aerosol of small droplets. This process involves the formation of a Taylor cone and a jet from the tip of the cone forming a plume of small droplets. The solvent is evaporated by heat and the ions are forced closer together. Due to Coulombic forces, the ions repel each other and the droplets break apart to smaller droplet. How this eventually leads to the formation of dried ions is not absolutely clear. There are two major theories[168]. The first one is the charged residue model theory which implies that the dried ions are formed by sequential disintegration cycles due to drying of the droplets. This finally leads to a single ion per droplet which is left with the charges when the droplet is dried down completely. The second theory, the ion evaporation model suggests that once the droplets shrink down to a certain radius, the field strength at the surface becomes big enough that the ions desorb out of the droplet.
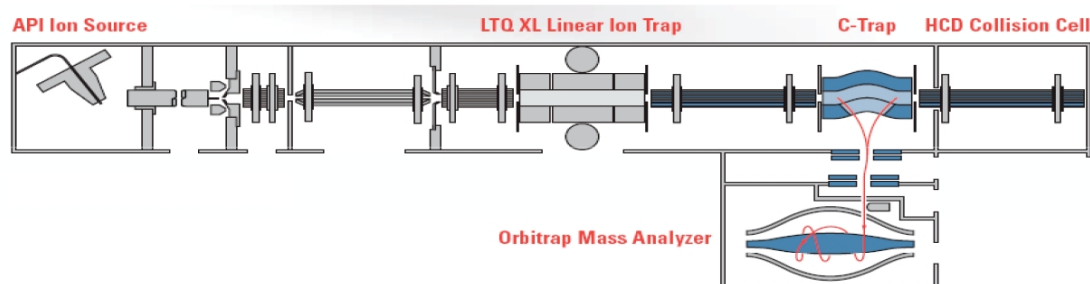

Figure 9: The LTQ OrbitrapXL. The figure shows the structure of the Orbitrap XL mass spectrometer. It is composed of an ion source where the ions are produced and from where they are transferred into the linear ion trap or via the C-trap into the Orbitrap mass analyzer. This figure is reprinted with the friendly permission of Thermo Fisher Scientific (www.thermo.com)

In a LTQ OrbitrapXL mass spectrometer operated in a top 10 acquisition mode with CID fragmentation, the ions are transferred through the linear ion trap into the C-trap from where they are injected into the Orbitrap. In the Orbitrap, a brief pre-scan is performed from which the ions are

selected for fragmentation and analysis in the linear ion trap. During the fragmentation analysis, a high resolution scan is performed in the Orbitrap. The Orbitrap itself is a novel type of mass analyzer, producing mass spectra of high resolution with high sensitivity and exhibits a high dynamic range[169-171]. In the Orbitrap, the ions injected from the C-trap, orbit around an axial electrode performing harmonic oscillation which is detected and transformed to mass spectra by Fourier transformation. For fragmentation, the selected ions are isolated in the linear ion trap, fragmented by CID and detected.

2.4.3.1 LC-MS/MS analysis

LC-MS/MS analysis was performed on an Ultimate3000 nano HPLC system online coupled to a LTQ OrbitrapXL mass spectrometer by a nano spray ion source. The Ultimate3000 nano HPLC system was equipped with a nano trap column (100 μm i.d. × 2 cm, packed with Acclaim PepMap100 C18, 5 μm, 100 Å) and an analytical column (75 μm i.d. × 15 cm, Acclaim PepMap100 C18, 3 μm, 100Å). The columns were heated to 40°C. The sample was automatically injected and loaded onto the trap column at a flow rate of 30 μl/minute in 5% buffer B (80% acetonitrile, 0.1% FA in HPLC grade water) and 95% buffer A (5% acetonitrile, 0.1% FA in HPLC grade water). After 5 minutes the peptides were eluted and separated on the analytical column by a gradient from 5% to 50% of buffer B at 300 nl/minute flow rate. Depending on the complexity of the sample, 20-140 minute gradients were used. Remaining peptides were eluted by a short gradient from 50% to 100% buffer B in 5 minutes. The eluting peptides were directly ionized by nano spray ionization. From the Orbitrap MS pre-scan, the 10 most intense peptide ions were selected for fragment analysis in the linear ion trap if they exceeded an intensity of at least 200 counts and if they were at least doubly charged. The normalized collision energy for CID was set to a value of 35 and the resulting fragments were detected with normal resolution in the linear ion trap. While the fragment analysis took place, a high resolution (60,000 FWHM) MS spectrum was acquired in the Orbitrap with a mass range from 300 to 1500 Da. To be able to calibrate every spectrum, the lock mass option was activated and the ion from ambient air with the mass of 445.120020 was used as lock mass[172]. This ion is stored in the C-trap and injected into the Orbitrap together with

the sample ions. Every ion selected for fragmentation, was excluded for 30 seconds by dynamic exclusion.

2.4.3.2   Spectrum filtering and database searching

The acquired spectra were processed and analyzed either by using the Bioworks Browser software and the Sequest[173] algorithm for database searching or by using Mascot[166]. For Sequest analysis, monoisotopic masses and full tryptic cleavage were selected. The peptide tolerance was set to maximal 10 ppm and the fragment ions tolerance to 1 Da. Only y and z ions were considered for the identification. The threshold for peak detection was set to 100 counts and the molecular weight range to 300-4500 Da. Cystein carbamidomethylation was selected as fixed modifications and methionine oxidation, serine, threonine and tyrosine phosphorylation were allowed as variable modifications with a maximum of three modifications per peptide allowed. The database used was Uniref100 from which species-specific subsets were produced using the Bioworks Browser.

For Mascot analysis, the raw files were converted to Mascot generic files using DTASupercharge which uses the extractmsn tool to generate dta files and combined them to Mascot generic files. Only LTQ CID MS^2 spectra were used. The Mascot generic files were then either directly submitted to Mascot or in case of pre-fractionation experiments, combined to a single file using the MultiRawPrepare tool. The data were analyzed using Mascot 2.2 with following search parameters: Trypsin was selected as enzyme, cystein carbamidomethylation was selected as fixed modification, methionine oxidation, serine, threonine and tyrosine phosphorylation were allowed as variable modifications. The peptide tolerance was set to 10 ppm, the MS/MS tolerance to 1 Da. The instrument type selected was ESI-Trap and the decoy option was activated. The database selected was the Uniref100 database and depending on the species the samples originated from, a species specific subset was selected.

2.4.3.3   Data analysis using Scaffold

Sequest as well as Mascot result files were analyzed by the Scaffold software. This tool uses the PeptideProphet[TM][174-176] algorithm to convert Mascot as well as Sequest scores into a single probability for peptide identifications. For Scaffold analysis, the search result files were imported

into the Scaffold software and analyzed using the same databases and modifications used for the database searches. After analyzing the data, 80% probability for peptide identifications, a minimum of two peptides and 95% as protein probability threshold were selected as minimum requirements for protein identifications.

2.4.4 Quantification by SILAC and mass spectrometry

Since mass spectrometry is not a quantitative method per se, several methods for quantifying proteins by mass spectrometry were developed[62,63,65,177]. The quantification of proteins by stable isotope labeling by amino acids in cell culture (SILAC) is one of these methods[62]. In general, it can only be applied metabolically active systems, like cell culture cells, since it involves metabolic labeling by stable isotope labeled amino acids. However, Krüger et al. developed a method by which the SILAC method was used for heavy isotope labeling of a complete organism-a mouse[178]. To enable labeling of proteins in cell culture, one or several amino acids in the growth medium are replaced by isotope labeled ones. The cells are grown in this medium for at least five generations so that virtually all of the amino acids selected for the labeling contained in the proteins are the isotope labeled ones. The major advantage of this method is that the samples can be combined at an early stage in sample processing to exclude variations due to sample manipulation.

For the quantification, the acquired MS spectra are used. The isotope-labeled amino acids cause a mass difference. This mass difference is dependent on the isotopes and amino acids used. A heavy isotopic form of lysine, $^{13}C6$ lysine, for example is composed of 6 carbon atoms which are ^{13}C atoms in the heavy condition instead of 6 ^{12}C atoms in the lysine used for the light condition medium. This creates a mass difference of 6 Da and the peptides can be quantified by comparing the peak area in the spectrum (Figure 10). For quantification, the area under all peaks for one peptide is then summed up for the heavy isotope-labeled peptide and the light isotope-labeled peptide respectively and the ratio between heavy and light peptides is calculated. This ratio corresponds to the difference in amounts of peptides in the samples and from the peptide ratios; the protein ratios can be calculated.
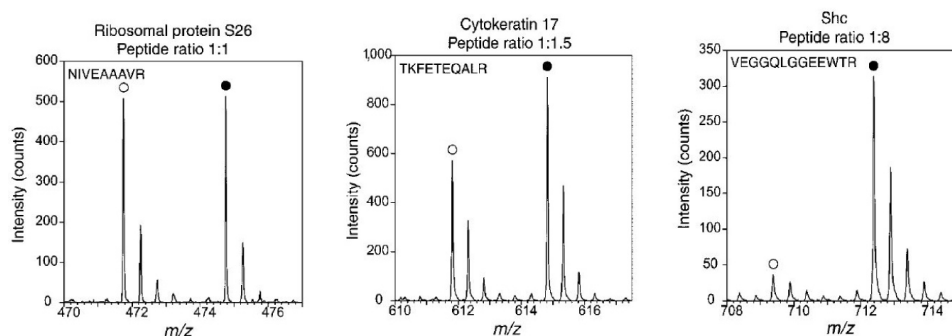
Figure 10: Examples for SILAC pairs in MS spectra. The figure shows three examples for SILAC pairs. The left panel shows an example where both, the light and heavy peptide are equally abundant and therefore show up as equally strong isotope pattern. The middle panel shows a SILAC pair where the heavy peptide is 1.5 fold more abundant than its light variant. The right panel shows a 8 fold difference which shows up as a 8 fold increased area under the peaks of the isotope pattern of the heavy isotope labeled peptide. This figure is reprinted with the friendly permission of the Nature Publishing Group[38]

Since trypsin cleaves either c-terminally of lysine or arginine, every tryptic peptide contains at least one lysine or arginine. To be able to quantify all tryptic peptides, labeled lysine as well as arginine were added to the growth media. Data analysis was done with Mascot with following modifications: additionally to the variable mutations described, heavy lysine and heavy arginine were allowed as variable modifications. ^{13}C6 lysine adds 6.020129 Da to the monoisotopic mass of lysine and ^{13}C6^{15}N4 arginine 10.008269 Da. After the database search, the HTML result file was exported as peptide report with following options: require bold red was selected; the significance threshold was set to 0.01 and the ions score cut-off was set to 25. The HTML file and the raw files were loaded into the MSQuant software[179] which was used for quantification analysis. All protein quantifications were inspected manually and then exported to Excel where the results of different analysis were combined and medium ratios were calculated.

Although initial experiments were quantified using MSQuant, they were all re-analyzed by the MaxQuant software because it was especially designed for the quantification of large datasets of high resolution and mass accuracy data, acquired with an Orbitrap[180]. The MaxQuant software is composed of two programs, the quant.exe and the identify.exe. The acquired raw files are first analyzed by the quant.exe. In a first step, it performs three-dimensional peak detection along the elution peak of all peptide signals and then automatically detects SILAC pairs which are then

used for ratio estimation. By the detection of the SILAC pairs, the number of lysines and arginines is determined. The calculated ratio is then normalized on all peptide ratios estimated in each LC run. Additionally, peptides with the same mass but different charge are detected. Before the data are prepared for Mascot database search, a non-linear mass calibration is done which is then used, in combination with the amino acid composition, to select the correct identifications in the second step (identify.exe). The Mascot search is done with the automatically prepared parameter files and the corresponding data file by using the Mascot Daemon. The database used includes a decoy database as well as common contaminants to determine the false discovery rate and to exclude false positive hits due to contamination by proteins from different species. The decoy database contains reversed sequences for all proteins present in the species specific database. Therefore, peptides identified in the decoy database must be false positive identifications and can be used for the calculation of false discovery rates. Following Mascot analysis, the Mascot result files together with the raw files, the FASTA database file, used for Mascot analysis and an experimental design file in which the experimental setup was defined were subjected to the identify.exe. In this process, the filtering of the Mascot identification using the amino acid composition takes place in a first step before linear mass calibration is performed and the identifications are filtered according to the determined mass accuracy. Following this step, the peptide false discovery rate is determined and the peptide identifications are filtered according to the threshold set in the program. If no peptide pair is detected, re-quantification uses the peptides identified by Mascot to quantify SILAC pairs, not detected in the first step. The reason for missed SILAC pairs can be extreme ratios. In case of protein complex analysis for example, the protein used as bait and complex components are strongly enriched in one sample and in many cases not or only weakly present in the control sample. Is the peak not detected in the control sample, because it is below the detection threshold, no SILAC pair is identified by MaxQuant. If the Mascot algorithm identifies such peptides including the SILAC modification, MaxQuant is again looking for the partner and can thereby quantify these peptides in many cases. After the detection and the re-quantification of these peptides, the peptide identifications are combined to protein identification groups, including all possible identifications based on the peptide identification and protein

ratios as well as significances are calculated. The protein ratios are calculated as the median of the logarithmic ratios of the peptide quantifications. The calculated significance is the outlier significance score for log protein ratios corresponding to the probability that a protein ratio is an outlier of a normal distributed sample set (significance A). In a second step, the protein identifications are grouped according to their intensity into groups of 300 protein groups and the significance B is calculated to consider the effect that the ratio variability is smaller for higher intensities when compared to low intensity identifications[181]. This second significance is only calculated if the number of identified protein groups is higher than 600 because only then more than one group can be build. The parameters used for MaxQuant analysis are shown in Table 13. The resulting protein intensities, ratios and significances were then used to visualize the result for each experiment by plotting the ratios against the intensities by a self written R-script (supplemental files 1-2). To highlight the significant altered proteins, color coding was used as shown in the legends of each plot.

Quant.exe

Parameter	Value
Instrument	Orbitrap/FT Ultra
SILAC	Doublets
Heavy labels	K-6, R-10
Max. labeled Aas	3
Variable modifications	Oxidation (M) Acetyl (Protein N-term)
Static modifications	Carbamidomethyl (C)
Database	MQ_IPI_human
Enzyme	Trypsine/P
MS/MS tol.	0.5 Da
Max. missed cleavage	2
Top MS/MS peaks per 100 Da	6

Identify.exe

Parameter	Value
Peptide FDR	0.01
Protein FDR	0.01
Max. peptide PEP	1
Min. peptide length	6
Min unique peptides	2
Min peptides	2
Reverse string	REV_
Contaminant string	CON_
Protein quantification	Use razor and unique peptides
Min. ratio count	3
re-quantify	√
Keep lower scoring matches of identified peptides	√

Table 13: Parameters used for MaxQuant analysis. The table shows all parameters selected in the Quant.exe as well as in the Identify.exe parts of the quantitative MaxQuant analysis for all experiments described.

2.5   Molecular biology

2.5.1   *Escherichia coli* cultures

Depending on the application, *Escherichia coli (E. coli) bacteria* were cultured either in liquid cultures or as plating cultures.

2.5.1.1   Liquid cultures

Liquid cultures were used for expanding clones for plasmid amplification. For this reason, single clones were transferred to 200 µl of LB-medium (Luria-Bertani; 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride), supplemented with the appropriate antibiotic for 8 hours before they were transferred into 5 ml antibiotics containing LB-medium for small size plasmid preparation or 200 ml for large scale plasmid preparations. The cultures were incubated over night at constant agitation.

2.5.1.2   Plating cultures

Plating cultures were used for clonogenic selection. The *E. coli* solution was streaked out on LB-plates (LB-medium, containing 1.5% agar, supplemented with the appropriate antibiotic) and incubated over night at 37°C before single clones were picked. The plates were stored at 4°C for several weeks.

2.5.1.3   Generation of cryo-stocks

500 µl of over night cultures were gently mixed with 500 µl of 50% (v/v) sterile glycerol and stored at -80°C.

2.5.1.4   Generation of chemically competent E.coli

E.coli of the strain DH5α were expanded as liquid culture over night in 2.5 ml LB-medium without antibiotics before they were diluted 1:100 in LB-medium, supplemented with 20 mM magnesium sulfate. The culture was then grown until a optical density (OD_{600}) of 0.4-0.6 was reached and collected by centrifugation at 5,000 x g for 5 minutes at 4°C. The pelleted bacteria were resuspended in TFB1 buffer (30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM magnesium chloride, 15% (v/v) glycerol, pH was adjusted to 5.8 with acetic acid). The suspension was incubated for 5 minutes at 4°C before it was centrifuged for 5 minutes at 5,000 xg and 4°C. The pelleted bacteria were resuspended in TFB2 buffer (10 mM MOPS [3-(N-

morpholino)propanesulfonic acid], 75 mM calcium chloride, 10 mM rubidium chloride, 15% (v/v) glycerol, pH was adjusted to 6.5 with potassium hydroxide solution). The suspension was incubated for 15 minutes on ice before it was portioned in 100 µl aliquots and frozen in liquid nitrogen. The aliquots were stored at -80°C.

2.5.1.5 *Chemical transformation of E.coli*

For chemical transformation, competent *E.coli* were thawed on ice. 10-30 ng of DNA or 5 µl of a ligation reaction were added to 50 µl suspension of competent *E. coli*. The mixture was incubated for 30 minutes on ice before it was incubated for 45 seconds at 42°C to initiate the uptake of the DNA complexes and then cooled down on ice for 1 minute. 350 µl of SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% sodium chloride, 20 mM glucose) were added and the suspension incubated under constant agitation for one hour at 37°C. Then the suspension was plated out.

2.5.2 Plasmid DNA preparation

Small scale plasmid preparation was done with the Plasmid Miniprep Kit. For preparative purposes, either the Pure Yield Midiprep Kit or the Endo-Free Plasmid Maxiprep Kit was used. For small scale preparations, 5 ml over night culture was used, for preparative purposes 200 ml. The preparations were done according to the manufacturer's instructions. All kits are based on alkaline lysis of bacteria and binding of the DNA to a resin. For the Miniprep and the Midiprep kits, the DNA was eluted by addition of hot water (60°C) and incubation for 1 minute (Miniprep) or 5 minutes (Midiprep). For the Endo-Free Maxiprep Kit, the DNA was eluted and precipitated according to the manufacturer's instructions before it was resuspended in 500 µl sterile ddH$_2$O. The concentration and purity of DNA preparations was determined by photometric determination of the absorbance at 260 and 280 nm. Therefore the DNA preparation was diluted 1 to 50 or 1 to 100 in HPLC grade water. A 260/280 nm ratio of 1.8 indicated a pure DNA preparation. The concentration was calculated as follows: DNA concentration [µg/µl] = absorption at 260 nm x 50 x dilution factor.

2.5.3 Agarose gel electrophoresis

DNA fragments were separated and purified by agarose gel electrophoresis. 1% (w/v) agarose was dissolved in TAE buffer (40 mM tris-acetate, 1 mM ethylenediaminetetraacetat (EDTA), pH 8) by boiling in a micro wave before the solution was poured in the gel tray and supplemented with 0.5 µg/ml ethidium bromide and the comb was inserted. After solidification of the gel, the comb was removed and the gel transferred into the electrophoresis chamber and overlaid with TAE buffer. The DNA samples were treated with loading buffer (6x: 0,25% bromphenol blue, 40% (w/v) sucrose) and loaded onto the gel. The gel run was performed with 50 V until the bromphenol blue front reached the end of the gel. The DNA bands were visualized by UV-light and digitalized. DNA was extracted from gel bands using the QIAquick Gel Extraction kit according to the manufacturer's protocol.

2.5.4 DNA sequencing

DNA sequencing was done by using the BigDye-Terminator v3.1 Sequencing Kit. Here, the DNA is amplified by a DNA polymerase using sequence specific primers. The dNTP mixture includes 4 dideoxynucleotides (ddNTPs), labeled with different fluorescent dyes, one for each nucleotide. This results in one label for each amplified DNA fragment. The label is dependent on the ddNTP that leads to disruption of the polymerase reaction. The fragments are separated by capillary electrophoresis and the fluorescence is detected. In this way, the fluorescent signal can be correlated to the nucleotide sequence.

For the sequencing reaction, 300-500 ng template DNA was mixed with 2µl of the BigDye kit, 0.5 µl primer solution (10 µM) and 2 µl 5x sequencing buffer which is included in the kit. ddH$_2$O was added to a final volume of 10 µl. The sequencing reaction was performed in a thermo cycler with the following program:

1.	96°C	2 min	denaturation of the double strand template
2.	96°C	30 s	denaturation of the double strand template
3.	50°C	15 s	annealing of the primer
4.	60°C	4 min	elongation of the fragment by polymerase
5.	60°C	4 min	final elongation

76

Steps 2 to 5 were repeated until 30 cycles were reached. Then the DNA was precipitated by addition of 90 µl 70% (v/v) ethanol and 15 minute centrifugation at 16,000 xg. The supernatant was removed and the pellet washed with 200 µl 70% ethanol, centrifuged again for 10 minutes at the same speed and the supernatant removed. The pellet was air dried and resuspended in 50µl HPLC grade water before the fragments were analyzed on an automated sequencer. The resulting spectra were analyzed by using Vector NTI Suite 9.0 software package.

2.5.5 DNA restriction digest

2 µg of each construct was used for restriction digests. The construct was mixed with 30 U of the two appropriate restriction enzymes. For each enzyme combination, the optimal buffer and temperature conditions were used as described by the manufacturer. The mixture was incubated for 2 hours before it was separated on an agarose gel and the fragment of interest was excised and extracted by the QIAquick Gel Extraction Kit according to the manufacturer's instructions.

2.5.6 DNA ligation

DNA fragments resulting from restriction digests were mixed in different ratios. As a starting point, equal molar amounts of the DNA fragments were used. In some cases, the ratios needed to be optimized to get sufficient results. In these cases, the amount of insert was varied from 3 to 100 fmol. 200 U T4 DNA-ligase were added to the DNA-fragment mixture and water was added to a final volume of 10 µl. The ligation reaction was done over night at 14°C.

2.5.7 Gateway cloning

The GATEWAY cloning technology is based on the recombination system of the phage λ which uses attachment (*attB*) sites in the target DNA to integrate its DNA into the target DNA by recombination. In the GATEWAY system, these *attB* sites are added to the DNA sequence of interest by a PCR reaction before the fragment is recombined into the Entry vector by the first reaction, the BP reaction. This vector serves as universal donor vector from which the sequence of interest can easily be transferred into any GATEWAY compatible expression vector by the LR reaction.

To circumvent possible problems with long primers, the attachment of the *attB* sequences to the sequence of interest was done in two PCR reactions. In the first reaction, primers specific for the sequence of interest and a part of the *attB* sequence was used and in a second PCR reaction, universal primers, adding the rest of the *attB* sequence were used. 10 ng of the DNA were mixed with 5 µl HF-Reaction buffer, 1.3 µl of forward and reverse primers (10 µM), 0.5 µl dNTP mix (10 µM) and 0.5 µl of the Phusion Tag Polymerase. ddH$_2$O was added to a final volume of 25 µl. The program used for the PCR reactions was as follows:

1. 98°C   2 min
2. 96°C   30 s
3. 55°C   40 s
4. 72°C   1 min / 1000base pairs sequence length

Steps 2 to 4 were repeated until 16 cycles were performed. The second reaction was performed as the first PCR with following modifications: the generic primers for the *attB* sequence were used and the number of cycles was increased to 25. The amplified fragment was separated by agarose gel electrophoresis and extracted before it was subjected to the BP reaction. For the BP reaction, 90 ng of the entry vector DONR201 were mixed with 3 µl of the purified PCR product and 1 µl of the BP Clonase II mix. ddH$_2$O was added to a final volume of 5 µl and the reaction was incubated at 25°C for 2 hours before 0.5 µl Proteinase K was added and incubated for 10 minutes at 37°C. The reaction mixture was then transformed into DH5α and amplified in liquid cultures before the plasmids were prepared by using the MiniPrep kit. The insert was verified by sequencing.

Verified constructs were used in the LR reaction to transfer the coding sequence into expression constructs like the pDEST-N-SF-TAP. For this reaction, 90 ng of the insert containing DONR201 plasmid was mixed with 90 ng of pDEST-N-SF-TAP vector and 1µl of LR Clonase II mix. ddH$_2$O was added to a final volume of 5 µl. The reaction was incubated for 2 hours at 25°C before 0.5 µl Proteinase K was added. The mixture was further incubated for 10 minutes. The plasmids were transformed into DH5α and amplified in liquid cultures. The plasmids were prepared using the MiniPrep kit and verified by restriction digest and agarose gel analysis.

2.5.8 Site-directed mutagenesis

Point mutations were generated using the QuikChangeII Site-directed Mutagenesis Kit. Primers were designed in a way that both the sense and anti-sense primers contained the nucleotide exchange in the middle of the primer sequence, ended 3' and 5' with either C or G and had a melting temperature higher than 78°C. The plasmid which was to be mutated was amplified by PCR using the PfuUltra High-Fidelity DNA polymerase using mutation-specific primers which led to the introduction of the mutation into the plasmid sequence. 20 ng of the plasmid were mixed with 125 ng of sense and anti-sense primer, 5 µl of reaction buffer, 1 µl dNTP mixture and 1 µl PfuUltra polymerase. ddH$_2$O was added to a final volume of 50 µl. The mutagenesis reaction was done in a thermo cycler with following program:

1.	96°C	30 s	denaturation of the double strand template
2.	96°C	30 s	denaturation of the double strand template
3.	55°C	15 s	annealing of the primer
4.	68°C	1 min/kb	elongation of the fragment by polymerase
5.	68°C	1 min/kb	final elongation

The cycle from step 2 to 4 was repeated until 16 cycles were achieved. After the mutagenesis reaction, the non mutated, methylated DNA template was selectively degradated by addition of *DpnI* endonuclease and incubation for one hour at 37°C. 5 µl of the reaction were then transformed into XL1-Blue Supercompetent Cells, plating cultures were used to select single colonies which were amplified in solution culture and the plasmids were prepared using Miniprep kits. The mutations were verified by sequencing.

G Results

1. Evaluation and optimization of the SF-TAP method

1.1 Optimization of the SF-TAP method

1.1.1 Time optimization

To increase the throuput of the SF-TAP system and to increase the probability to identify more transient protein complex components, the time necessary for incubation of the samples with the resin was optimized. For this reason, SF-TAP tagged 14-3-3ε, B-Raf and MEK-1 stably expressed in HEK293 cells, were purified by SF-TAP as described with some modifications. After generation of the cell lysates from 1×10^8 cells for each bait protein, the lysates were divided into two samples of equal volume. The lysates were all incubated with the same amount of resin in both purification steps. One half of the sample was incubated for one hour in both steps, the other half for two hours in both purification steps. The resulting samples were analyzed by quantitative Western blotting using the anti-Flag-M2 antibody and the Cy5 labeled secondary, anti-mouse antibody. The protein bands were quantified using ImageJ.

The results showed that the efficiency of the SF-TAP procedure was only slightly lower if the incubation steps were reduced to one hour. In case of 14-3-3ε, the first Strep purification was only 3% less efficient with the shorter incubation, but 15% lower after both steps. A similar result was observed for MEK-1. The Strep purification was 12% less efficient but overall the efficiency was 15% lower compared to the two hour incubation. For B-Raf, the Strep purification was clearly less efficient (81% efficiency). In contrast to the first step, the second step was more efficient when the incubation time was reduced and the overall efficiency was at 94% (Figure 11). Even though the efficiency was slightly lower when the incubation times were reduced, the reduction of the time necessary for the overall purification clearly increases the possible throughput as well as the probability to identify rather transient complex components.
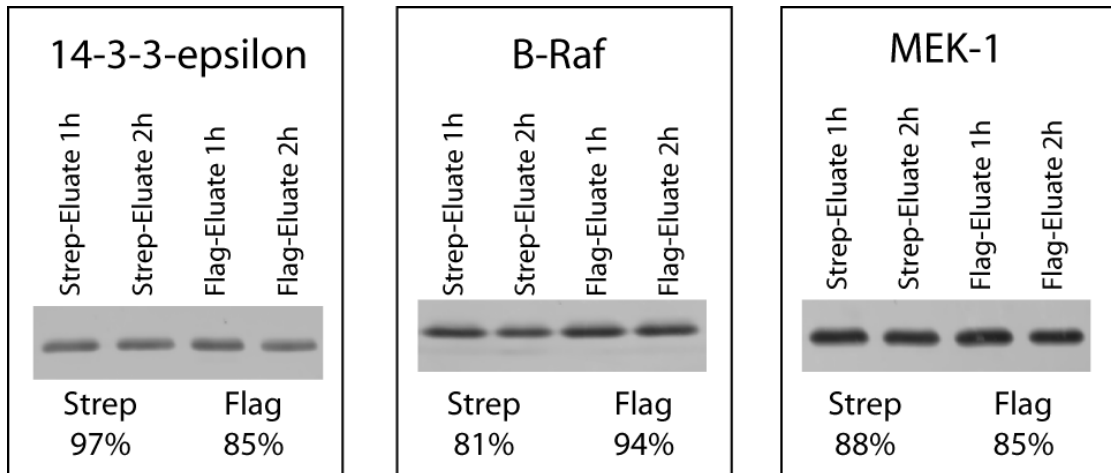
Figure 11: Time optimization of the SF-TAP purification. 14-3-3ε, B-Raf and MEK-1 were purified by the SF-TAP method from HEK293 cell lines stably expressing the bait proteins. Two experiments were done in parallel for each protein, one with incubation times of one hour for each step and one with an incubation time of two hours for each step. Samples were taken after the first, the Strep purification and after the second, the Flag purification and analyzed by quantitative Western blot using Cy5-labeled secondary antibodies. The bands were quantified using ImageJ. The quantification showed that there were only minor differences between one or two hour incubation times.

1.2 Evaluation of the SF-TAP method

1.2.1 One step purification versus SF-TAP

The effort to do TAP purifications is clearly higher when compared to single step purifications. Even though the SF-TAP method is quite simple and straight forward, a one step purification is still faster, easier to do and more efficient. However, the TAP method has a clear advantage considering the purity of the resulting complexes, making them much easier and faster to analyze. To demonstrate this advantage, a one step Strep purification was compared to the SF-tandem affinity purification (SF-TAP) (Figure 12). B-Raf was purified from HEK293 cells, stably expressing SF-TAP tagged B-Raf. The protein extract from 1.4×10^8 cells was divided into two equal samples. One sample was subjected to a one step Strep purification, the other to SF-TAP. The purified protein complexes were resolved by SDS-PAGE and its components were visualized by silver staining. The result clearly demonstrated that after the first purification step, the protein complex is highly contaminated by non-specifically bound proteins. After SF-TAP, no contaminating proteins were visible in the control lane. The loss of efficiency was acceptable for the bait protein (Figure 12: band around 97 kDa) and stable interactors like 14-3-3 (Figure 12: 2 bands around 30 kDa). However, when considering less stable

interactors that show high association and dissociation rates like MEK1/2, the loss was substantial in the second step (Figure 12: band around 50 kDa).
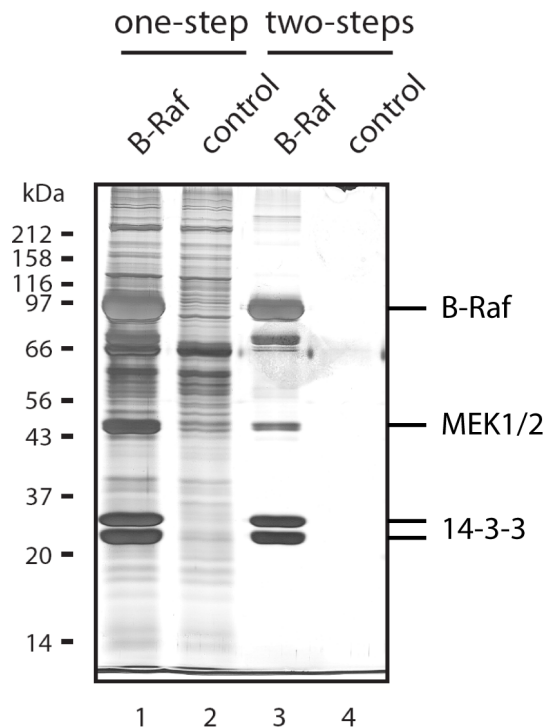

Figure 12: One step versus two step purification. SF-TAP tagged B-Raf was purified from HEK293 cells, stably expressing the tagged protein. In one experiment, the eluate of the first purification, the Strep purification was used and compared to the control. As control, HEK293 cells expressing the SF-TAP tag alone were used. In the second experiment, the eluate of both purification steps, Strep- followed by Flag purification, was again compared to the control. All four eluates were analyzed by SDS-PAGE followed by silver staining. The result shows clearly that after one step, the eluate is highly contaminated by non-specifically bound proteins and that after the second purification step; the non-specific background is below detection by silver staining. This figure is reprinted with the friendly permission of Wiley-VCH Verlag GmbH & Co. KGaA[162].

1.2.2 Determination of the SF-TAP efficiency

The efficiency of the SF-TAP method was determined by comparing the bait input in the protein extract to the purified bait protein by quantitative Western blotting using the anti-Flag-M2 antibody and the Cy5-labeled anti-mouse secondary antibody. Protein bands were quantified using ImageJ. Proteins were purified using the SF-TAP method from 7×10^7 HEK293 cells, stably expressing the bait proteins. Because the shape of protein bands in SDS-PAGE differs depending on the complexity of the sample, the purified protein complexes were spiked back into HEK293 protein lysate so that the amount of protein and the sample complexity was equal in all lanes. Additionally, dilution series were conducted for each bait protein to prevent saturation effects. In the linear range for each protein, 2 fold dilution series were done for the purified protein and the medium intensity of all dilutions was calculated. The ratio between the medium intensities for the purified protein and the input was calculated. For the three bait proteins B-Raf, 14-3-3ε and MEK-1 three independent samples were generated and quantified. The medium and standard deviation was calculated. Figure 13

shows the efficiencies determined by this method. For B-Raf, the efficiency was 30% ±2.5%, 47.7% ±2.3% for MEK-1 and 39.5% ±1.0% for 14-3-3ε.
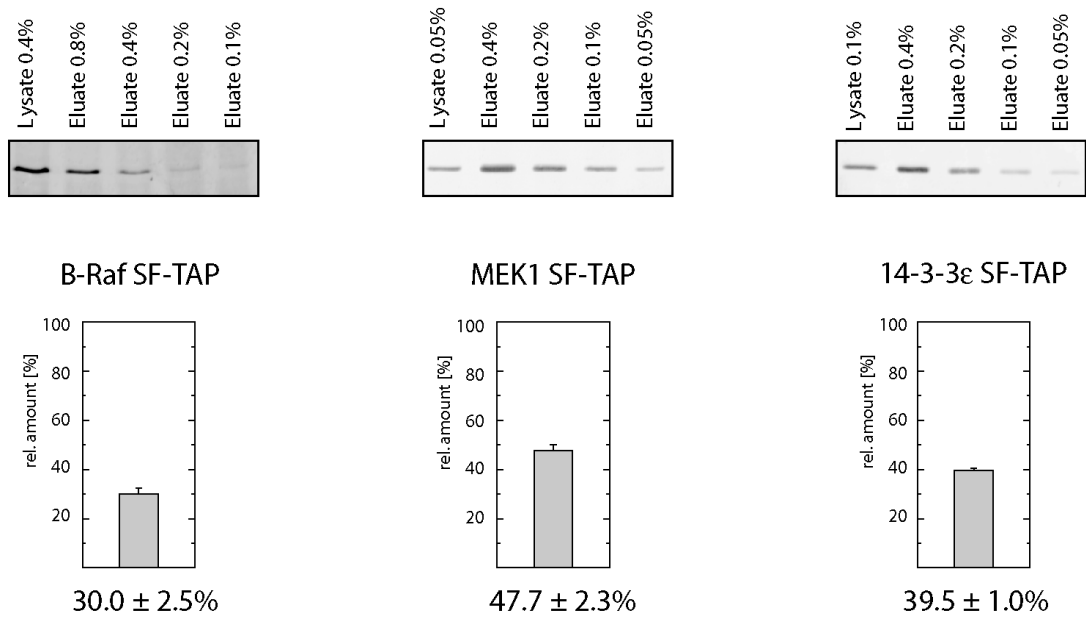

Figure 13: Efficiency of the SF-TAP method. SF-TAP tagged B-Raf, MEK-1 and 14-3-3ε were purified from HEK293 cells, stably expressing the bait proteins. The eluates of the SF-TAP procedure were spiked back into HEK293 cell lysate corresponding to the protein amount in the lysate to which the eluate was compared. The eluates were loaded in dilution series to assure quantification within the linear range of the system. The proteins were separated by SDS-PAGE and analyzed by Western blotting using Cy5-labeled secondary antibodies. The quantification was done by using ImageJ. The experiments were done in triplicates. The efficiency was calculated as the percentage of bait protein, quantified in the eluates versus the protein quantified in the eluate. This figure is reprinted with the friendly permission of Wiley-VCH Verlag GmbH & Co. KGaA[162].

1.2.3 Expression of SF-TAP tagged versus TAPo tagged proteins

The SF-TAP method is a new method for tandem affinity purification that significantly differs from the original tandem affinity purification (TAPo) method[34,162]. Therefore, before comparing the efficiencies of the both methods, it was essential to achieve equal expression levels of the bait proteins. 14-3-3ε and B-Raf were tagged with either the TAPo tag or the SF-TAP tag. For all baits, 3 different DNA preparations were used to exclude expression differences due to different DNA quality.

The expression levels for 14-3-3ε and B-Raf in HEK293 cells, transiently expressing the bait, are shown in Figure 14. Prior to Western blotting, the protein extracts were subjected to TEV cleavage. This was necessary because the original tag contains a protein A domain which strongly binds antibodies via their Fc region and would therefore lead to a falsified result

due to overestimation of the amount of TAPo-tagged bait protein in the protein extract. 10 µg of protein extract were loaded in each lane, separated by SDS-PAGE and subjected to quantitative Western blot using the B-Raf (H145) antibody and the anti-14-3-3ε antibody. For 14-3-3ε, there was no significant difference in expression levels. For B-Raf on the other hand, the expression was clearly lower for the TAPo tagged protein.
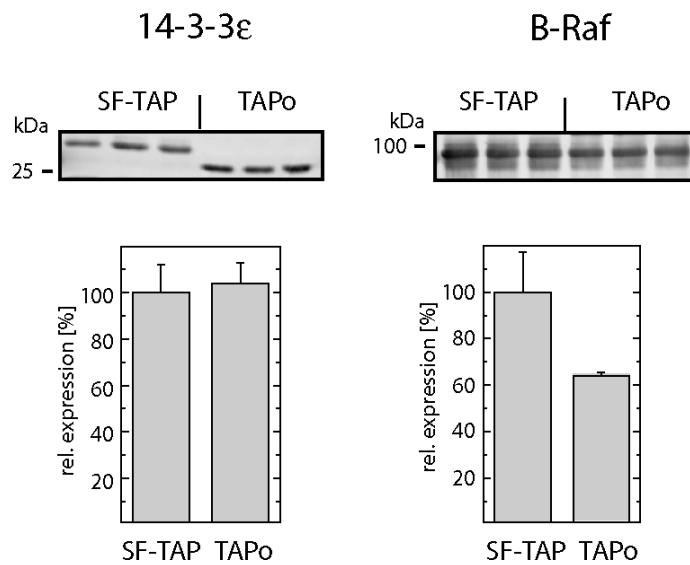

Figure 14: Expression of SF-TAP tagged and TAP tagged proteins. SF-TAP and TAP tagged 14-3-3ε and B-Raf were purified from HEK293 cells, transiently expressing the bait proteins. The TAP lysates were subjected to TEV cleavage before the expression of the proteins was compared to the SF-TAP tagged counterparts. The quantification was done by Western blotting using Cy5-labeled secondary antibodies and ImageJ. The experiment was conducted three times. The expression of 14-3-3ε was equal for both tags, for B-Raf the expression was significantly higher for the SF-TAP tagged protein. The molecular weight for TAPo tagged 14-3-3ε is lower compared to the SF-TAP tagged version because the protein A tag was cleaved of before western blot analysis. This figure is reprinted with the friendly permission of Wiley-VCH Verlag GmbH & Co. KGaA[162].

1.2.4   Comparison of the SF-TAP and the TAPo method

The comparison of the original TAP tag and the SF-TAP tag methods was done for 14-3-3ε and B-Raf. The expression levels for B-Raf differed between SF-TAP tagged and TAPo tagged (Figure 14) but by increasing the amount of DNA used for transfection to 1.5 fold the amount of SF-TAP tagged B-Raf led to similar expression levels. For all purifications, protein extracts from 1×10^8 HEK293 cells, transiently expressing the tagged bait proteins were subjected to SF-TAP or TAPo. The purified protein complexes were separated by SDS-PAGE and the gels were stained by silver staining.

The major protein bands were excised and subjected to in-gel-digestion. The tryptic peptides were analyzed by MALDI-TOF/TOF. The resulting band patterns did not differ significantly between SF-TAP and TAPo purified protein complexes. For 14-3-3ε as well as for B-Raf, the major bands observed are the same, except the BSA band (Figure 15) which was only observed as contaminant in the TAPo purified 14-3-3ε complex. The purification efficiency was comparable for 14-3-3ε. For B-Raf, the efficiency was slightly higher for the SF-TAP method. Even if the amount of DNA for transfection for TAPo tagged B-Raf was increased, the amount of purified bait as well as the amount of co-purified interactors was higher for the SF-TAP tagged and purified B-Raf. This shows that the efficiency of the SF-TAP method is at least comparable to the efficiency of the original TAP method.
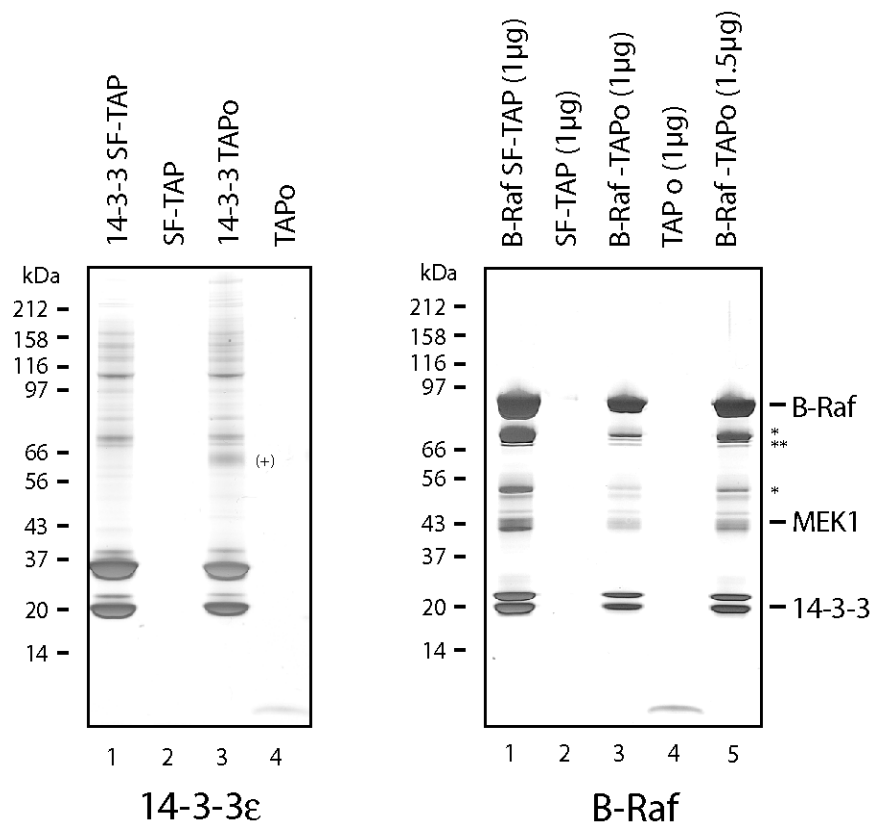
Figure 15: Efficiency of the SF-TAP and the TAP method. SF-TAP and TAPo tagged 14-3-3ε and B-Raf were transiently expressed in HEK293 cells and purified by either the SF-TAP or the TAPo method. The eluates were separated by SDS-PAGE and the gels stained by silver staining. For B-Raf, the TAP tagged protein was additionally expressed using 1.5 fold the amount of DNA for transfection to achieve similar expression levels as for the SF-TAP tagged protein. The efficiency for 14-3-3ε was slightly higher for the SF-TAP method. The same is true for B-Raf. Legend: +: BSA, *: Degradation product of B-Raf, **: HSP70 chaperones. This figure is reprinted with the friendly permission of Wiley-VCH Verlag GmbH & Co. KGaA[162].

1.2.5 Analysis of MAPK protein complexes

The results in the previous chapters show that the SF-TAP method is an efficient tool for purifying protein complexes from mammalian cells. To further demonstrate the applicability of the method, HEK293 cell lines, stably expressing SF-TAP tagged proteins were used for a protein network analysis of the MAPK pathway. The MAPKKKs B-Raf and C-Raf, the MAPKK MEK-1 and the adaptor protein 14-3-3ε were selected for this demonstration. These proteins are all part of MAPK pathway and the complexes they form are well defined[10,35,36]. Therefore these proteins are a good model to show that the SF-TAP tag is a useful tool to analyze protein complexes and networks. In a first approach a classical, gel-based separation combined with in-gel-digestion and analysis by MALDI-TOF/TOF was applied to the purified complexes. In a second approach, the purified complexes were subjected to in-solution-digestion and analyzed by LC-MALDI.

1.2.5.1 Gel based analysis

SF-TAP tagged B-Raf and C-Raf were purified from $3x10^8$ HEK293 cells, MEK-1 and 14-3-3ε from $1x10^8$ HEK293 cells, stably expressing the bait proteins. All purified protein complexes were separated by SDS-PAGE and the gels stained by silver staining. The excised protein bands were subjected to in-gel-digestion and analyzed by MALDI-TOF/TOF. The silver stained gels and the identified protein bands are shown in Figure 16. Each protein shown was identified by MS and MS/MS identifications.

For B-Raf, MEK-1 and MEK-2 as well as several 14-3-3 isoforms could be identified as complex components (Figure 16, first panel). C-Raf binds only weakly to the two MEK isoforms[182]. Therefore this interaction was not detected. HSP90, Tubulin, CDC37 as well as 14-3-3 isoforms are known interactors of C-Raf and could be identified by the gel-based approach (Figure 16, second panel). The MEK-1 SF-TAP analysis revealed only B-Raf, 14-3-3 isoforms as well as chaperones as complex components (Figure 16, third panel). The main downstream target of MEK, Erk, could not be identified. The reason for that could be that binding of Erk to MEK is characterized by high dissociation rates of Erk from MEK[183]. The band pattern observed for the 14-3-3 complex was highly complex when compared to the other protein complexes analyzed. This suggests that the complexes formed by 14-3-3 proteins are rather stable and that 14-3-3

proteins are involved in many cellular functions (Figure 16, fourth panel). For all proteins, chaperones could be identified as well. HSP90 and CDC37 are known regulators of B-Raf but association of the HSP70 proteins are probably due to the over expression of the bait proteins which leads to misfolding and thereby to the decoration with chaperones. Even though the purified protein complexes result from stable cell lines, these proteins were still over expressed when compared to the endogenous protein levels. The example of 14-3-3ε shows clearly the need for the second approach used: the LC-MALDI approach. In a SDS-PAGE-based approach, the resolution was not high enough to identify all protein complex components present in the SF-TAP eluate, especially with respect to the 14-3-3ε complex. Nevertheless this method was sufficient to show that 14-3-3 proteins form complexes with cellular signaling proteins like AKTs1, TSC2 and MARK2 as well as with proteins involved in cytoskeleton organization like KLC2 and KLC3, KHC and the adapter protein PARD3.
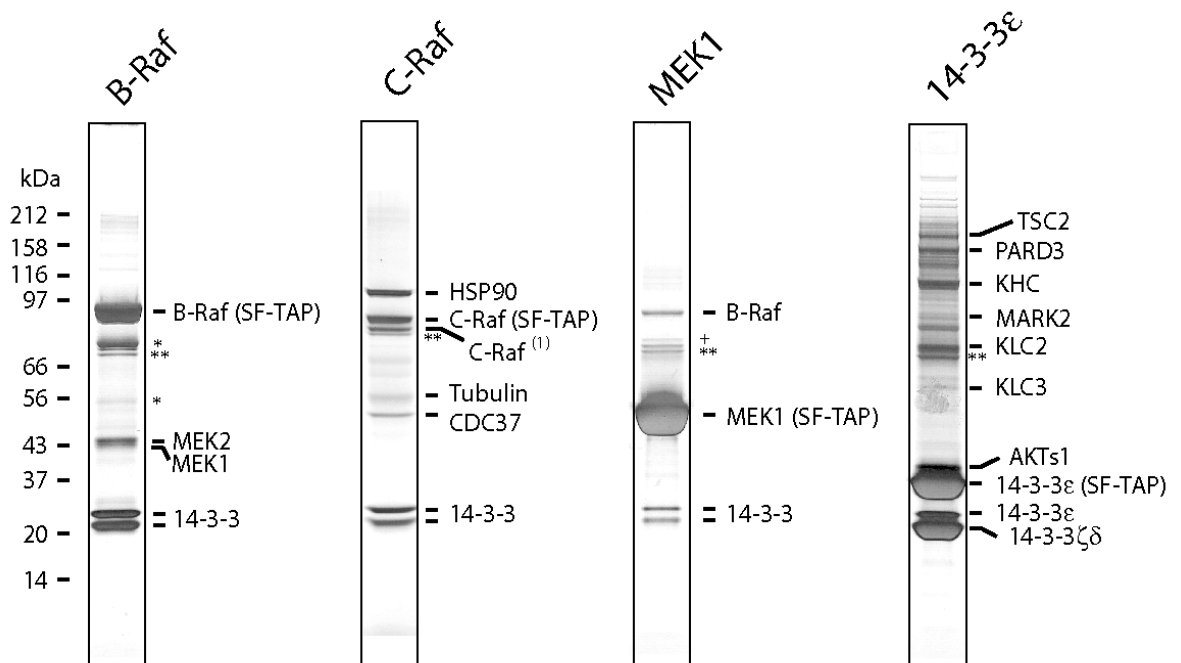

Figure 16: Gel-based analysis of MAPK complexes. B-Raf, C-Raf, MEK-1 and 14-3-3ε were purified by SF-TAP from HEK293 cells, stably expressing the SF-TAP tagged proteins. The eluates were analyzed by SDS-PAGE followed by silver staining. Protein band were subjected to in-gel digestion and the tryptic peptides were analyzed by MALDI-TOF/TOF analysis. Legend: *: typical degradation product of B-Raf, *: HSP70 chaperones, **: Bip chaperones. This figure is reprinted with the friendly permission of Wiley-VCH Verlag GmbH & Co. KGaA[162].

The interaction of 14-3-3ε and MEK-1 with B-Raf as well as with C-Raf was confirmed by SF-TAP analysis of MEK-1 and 14-3-3ε combined with Western blotting using B-Raf and C-Raf antibodies (Figure 17). In case of B-Raf this was already shown by the gel-based approach of the MEK1 SF-TAP eluate but C-Raf could not be identified in this case. The identification of B-Raf in the MEK1 SF-TAP eluate shows that not only substrates of bait proteins can be identified but upstream components of signaling cascades as well.
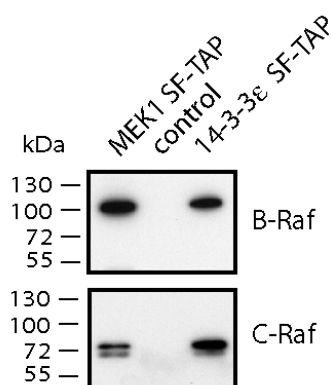

Figure 17: WB analysis of C-Raf and B-Raf binding. MEK-1 and 14-3-3ε SF-TAP eluates were separated by SDS-PAGE and analyzed by Western blotting using antibodies specific for B-Raf and C-Raf. Both, C-Raf and B-Raf could be clearly detected in the eluates and no signal was observed in the SF-TAP control. This figure is reprinted with the friendly permission of Wiley-VCH Verlag GmbH & Co. KGaA[162].

1.2.5.2 LC-MALDI based analysis

The identical samples used for the gel-based approach were used for the LC-MALDI based analysis of the protein complex as well. The purified protein complexes were subjected to in-solution tryptic digestion and the peptides were analyzed by LC-MALDI. For the separation 120 minutes gradients were used to enable the identification of several tens of proteins. All proteins identified by at least two significant peptides are shown in supplemental table 1. The complexity of the B-Raf, C-Raf and MEK-1 protein complexes was relatively low. For that reason, the LC-MALDI approach did not result in a much higher number of identifications when compared to the gel-based approach. Nevertheless, by the LC-MALDI approach, all three Raf isoforms (A-Raf, B-Raf and C-Raf) were identified as components of the MEK-1 complex as well as the 14-3-3 protein complex. MEK1 was not identified in the gel-based approach as part of the C-Raf complex but could be identified by the LC-MALDI based analysis.

KSR1 was identified as additional complex component of MEK-1 as well as 14-3-3. KSR1 is a scaffold protein which assembles the MAPK signaling complexes[24,25]. Especially the LC-MALDI analysis of the 14-3-3 SF-TAP eluate showed the power of the method when protein complexes of high complexity are to be analyzed. By the gel-based analysis mainly cytoskeleton associated proteins and components of cellular signaling machineries were identified. By the LC-MALDI approach, many more proteins of the 14-3-3 protein complex could be identified. Among these were a large number of kinases and proteins involved in G-protein signaling and apoptosis. Most of the complex components identified by SF-TAP analysis were already described before (Figure 18, supplemental table 1)[35,36]. Overall, the identified complexes demonstrate that the SF-TAP method is a highly useful tool for the identification of protein complexes in cellular signaling and that the gel-free LC-MALDI approach increases the sensitivity of the overall system.
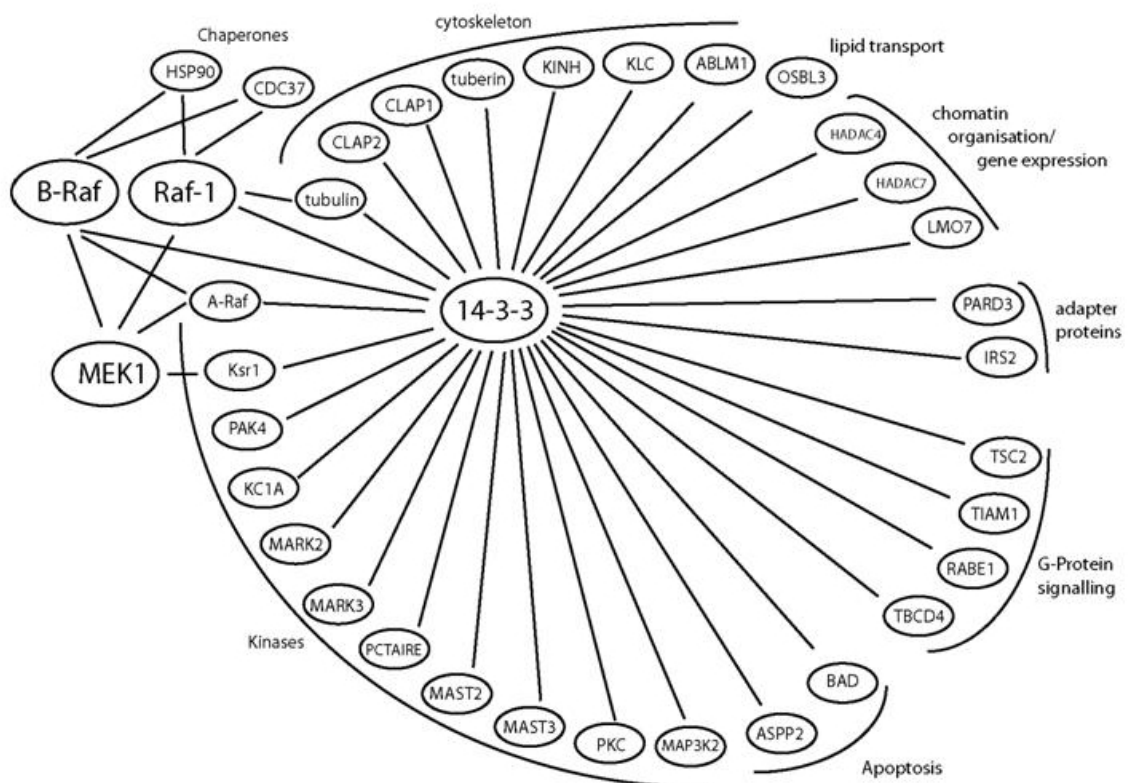

Figure 18: LC-MALDI analysis of MAPK complexes. The same samples as used for the gel-based analysis were subjected to in-solution tryptic digestion and analyzed by LC-MALDI. The proteins identified with at least two unique peptides were arranged in a network image. A full list of identifications is given in supplemental table 1.

2. Protein complex analysis of lebercilin

Lebercilin was found to be mutated in patients affected by Leber congenital amaurosis[73]. The normal function lebercilin exerts as well as the mechanism by which mutated lebercilin causes the disease is unknown. The lebercilin protein consists of four coiled coiled domains. Except this domain structure, nothing else was known about the structure or the function of lebercilin at the time this study was done. Therefore, the protein complex of wild type lebercilin was investigated followed by a comparative analysis of the complexes of wild type and mutated lebercilin. The first step was the analysis of the lebercilin protein complex by SF-TAP followed by identification of the protein complex components by LC-MALDI. In a second step, the lebercilin complex was analyzed by a quick one step purification combined with quantitative MS.

2.1 Complex analysis by SF–TAP

N- and C-terminally SF-TAP tagged lebercilin was used for the analysis of the lebercilin protein complex by SF-TAP. Approximately $2x10^8$ cells expressing N–terminally tagged and $4x10^8$ cells expressing C-terminally tagged lebercilin and $4x10^8$ cells transfected with the SF-TAP alone were harvested and protein extracts were generated. The lebercilin protein complex was purified using SF-TAP and the final eluates were analyzed by LC-MALDI. Therefore, the SF-TAP eluates were precipitated by the methanol-chloroform method to remove interfering substances and to concentrate the samples. The resulting pellets were subjected to in-solution tryptic digestion. The peptides were analyzed by LC-MALDI. Two biological replicates were produced and the purified complexes were analyzed for N- and C-terminally tagged lebercilin. A complete list of the identified peptides and the identification details are shown in supplemental table 2. The list of proteins, identified with at least two significant peptides in at least one biological replicate is shown in Table 14. Only dynein light chain was identified by one peptide. However, since dynein light chain is a protein of only 10 kDa, the one identified peptide resulted in approximately 20% sequence coverage. Only proteins that were not identified in the SF-TAP control are shown.

Uniprot	Protein	Identified in x of 2 analysis	
		N-SF-TAP	C-SF-TAP
Q86VQ0	Lebercilin	2	2
	Adaptor proteins		
P62258	14-3-3 protein epsilon	2	2
P31946	14-3-3 protein beta/alpha	2	2
P61981	14-3-3 protein gamma	2	2
P63104	14-3-3 protein zeta/delta	2	2
P27348	14-3-3 protein theta	2	2
Q04917	14-3-3 protein eta	1	1
	Cytoskeleton associated proteins		
P06748	Nucleophosmin		1
P19338	Nucleolin	1	2
P63167	Dynein light chain 1	1	1
Q96FJ2	Dynein light chain 2, cytoplasmatic	1	
	Cellular signaling proteins		
P67870	Casein kinase II subunit beta	2	
P68400	Casein kinase II subunit alpha	2	1
P19784	Casein kinase II subunit alpha'	1	1
Q01105	Phosphatase 2A inhibitor I2PP2A	2	
	Chaperones and co-chaperones		
P08107	Heat shock 70 kDa protein 1	2	2
P11142	Heat shock cognate 71 kDa protein	2	2
P11021	78 kDa glucose-regulated protein	2	2
P34931	Heat shock 70 kDa protein 1L	1	
P38646	Stress-70 protein	2	2
P34932	Heat shock 70 kDa protein 4	1	
	Miscellaneous		
Q96S59	Ran-binding protein 9	1	
P55209	Nucleosome assembly protein 1-like 1	2	
Q14241	Transcription elongation factor B polypeptide 3	2	
Q93008	Ubiquitin thioesterase FAF-X	1	
Q07021	Glycoprotein gC1qBP		1

Table 14: SF-TAP analysis of lebercilin. N- and C-terminally SF-TAP tagged lebercilin was expressed in HEK293 cells. The complexes were purified by the SF-TAP method. The eluates were concentrated using methanol-chloroform precipitation and subjected to in-solution tryptic digestion. The tryptic peptides were analyzed by LC-MALDI. For each tag (N- and C-terminal) two independent experiments were performed. Only proteins identified in at least on experiment with at least two peptides were considered with one exception. Dynein light chain 2 was only identified with one peptide which resulted in high sequence coverage (20%) due to the low molecular weight of only 21 kDa. The identified proteins were manually grouped according to their function. This table is reprinted with the friendly permission of the Nature Publishing Goup[73].

Many of the 24 identified proteins have been previously described to associate with microtubules, the centrosome or with the connecting cilium in photoreceptors. Among these proteins are the 14-3-3 isoforms, which bind and regulate mainly phosphorylated proteins, belonging to the cellular signaling machinery or are involved in organization of the cytoskeleton[35,36,162,184]. Additionally, 14-3-3 proteins interact with centrosomal proteins[185,186]. Nucleophosmin mainly localizes to the nucleus but it has been shown to localize to the centrosome as well[187]. In addition, it was described to bind to RPGR, a protein mutated in retinal dystrophies[111]. Like nucleophosmin, nucleolin localizes mainly to the nucleus but was described localize to photoreceptors and to bind casein kinase 2 in these cells[188-191]. Casein kinase 2 was also identified in the lebercilin protein complex. It localizes to the centrosome[192] and modulates the localization of several proteins to the cilia[193,194]. An important association was found with dynein light chain, a component of the dynein motor protein, driving the retrograde transport along cilia and microtubules. These interactors suggest that lebercilin localizes to the connecting cilium in photoreceptors. The identification of dynein in the protein complex suggests that lebercilin could exert a function in minus-end directed microtubular transport. Taken together, the proteins identified as complex components and the disease phenotype in the human eye only suggest a function of lebercilin in the ciliary transport at the connecting cilium of photoreceptors.

2.2 Quantitative protein complex analysis

The protein complex analysis of lebercilin by the SF-TAP method revealed 24 specific complex components. Most of these proteins are associated with microtubular structures like the centrosome and cilia. They suggested a role of lebercilin in the minus-end directed microtubular transport but there is no clear functional prove for the specific function of lebercilin. To further pinpoint the function of lebercilin, a second strategy for identifying complex components and comparing the complexes was applied. This strategy involved a quick one-step purification via the Strep-tag II which is included in the SF-TAP tag, and quantification by SILAC and MS.

In contrast to SF-TAP analysis this approach is advantageous for the identification of transient and lowly expressed complex components. But it

cannot be excluded that complex components are missed because they bind non-specifically to the resin or the tag as well as to the protein of interest. Figure 19 shows a scheme for the strategy applied. In the first step, the specific protein complex components are identified by quantitative comparison of the proteins in the control eluate to the proteins in the lebercilin eluate. By this comparison, specific complex components can be detected due to their enrichment in the lebercilin eluate. The second step of this workflow involves the comparison of complexes formed by the WT protein and by the LCA-associated P493TfsX1 mutant. The same principle as described above is used again but now, instead of the SF-TAP as control, the mutant protein is used for the comparison. The comparison of the normal versus mutant variant can reveal changes in the interaction pattern linked to the mutation that point to the mutant's dysfunction.
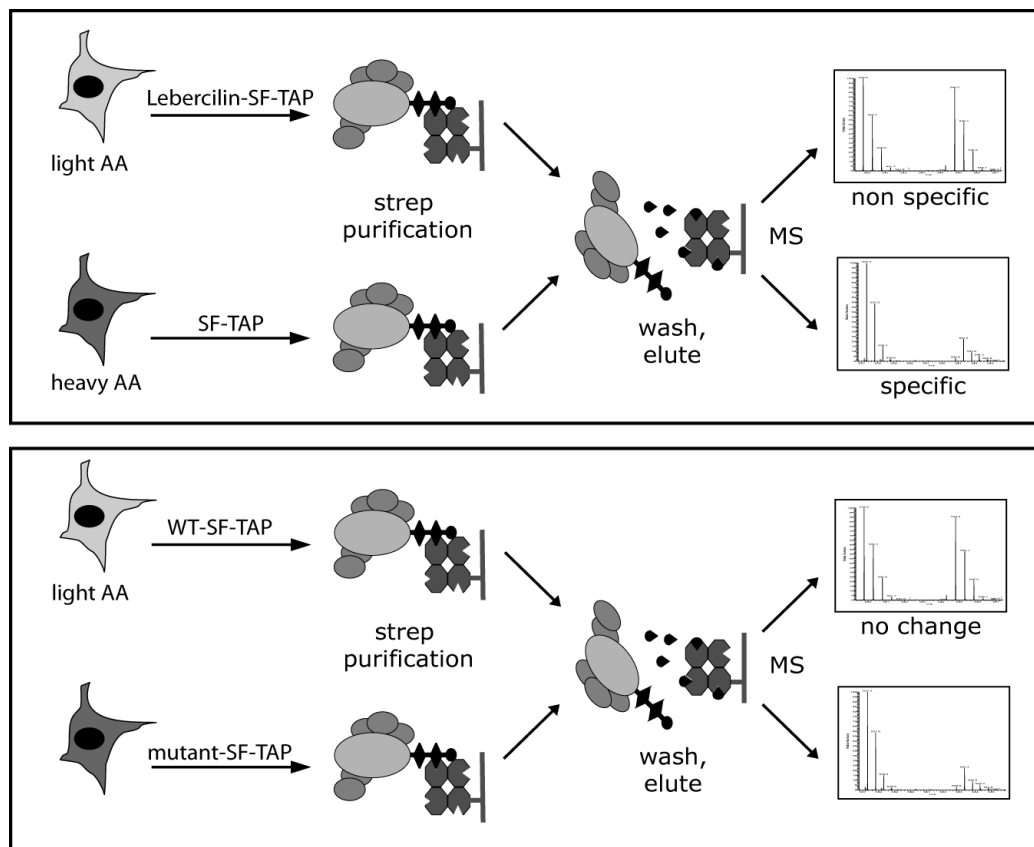
Figure 19: Quantitative complex analysis. The figure shows a scheme for the identification of specific complex components as well as for the comparison of different protein complexes. The SF-TAP tagged protein and the tag alone are expressed in SILAC-labeled cells and the complexes enriched by Strep purification before the samples are combined and eluted and quantified by MS. Non-specifically bound proteins are characterized by a 1:1 ratio, specifically bound proteins are enriched in the sample compared to the control (upper panel). For the comparison of complexes, the same principle is applied with one modification: instead of the SF-TAP as control, the mutant form of the protein is used (lower panel).

2.2.1 Detection of specific complex components

To detect specific components of the lebercilin protein complex, HEK293T cells were grown in heavy and light SILAC medium for 14 days before they were transfected with SF-TAP tagged lebercilin or the SF-TAP vector alone. 48 hours post transfection, protein extracts were generated from 2×10^8 cells for each sample and the protein concentration was determined using the Bradford protein assay. Equal protein amounts of the corresponding samples were subjected to the Strep purification separately. Before washing and elution, the Strep-Tactin beads with the bound lebercilin protein complexes were combined. The eluates were concentrated to a final volume of approximately 20 µl using 10 kDa cut-off centrifugal units. The concentrated eluates were separated by SDS-PAGE on 10% pre-cast gels. The gels were then stained with Coomassie until the first bands were visible and then de-stained. Each lane was excised and separated into 8 segments to pre-fractionate the sample. The segments were subjected to in-gel digestion.

The peptides of each fraction were analyzed by LC-MS/MS on an OrbitrapXL by a 140 minute gradient. The data were analyzed and quantified using MaxQuant[180,195]. For all further calculations, the normalized protein ratios were used. Each experiment consisted of a forward and a reverse labeling approach. This was done to exclude any possible effects due to the labeling. The whole experiment was repeated three times to get highly confident data. Proteins which deviated with a p-value between 0.01 and 0.001 from the standard distribution were further inspected whereas the proteins with p<0.001 were immediately accepted as specific complex components. The complete list of all quantified and identified proteins is given in supplemental table 3. Proteins identified as specific complex components, are shown in Table 15. The plot shown in Figure 20 clearly shows that a number of proteins are significantly enriched in the lebercilin purification when compared to the control sample. There is a clear-cut border between specifically enriched and non-specifically bound proteins.
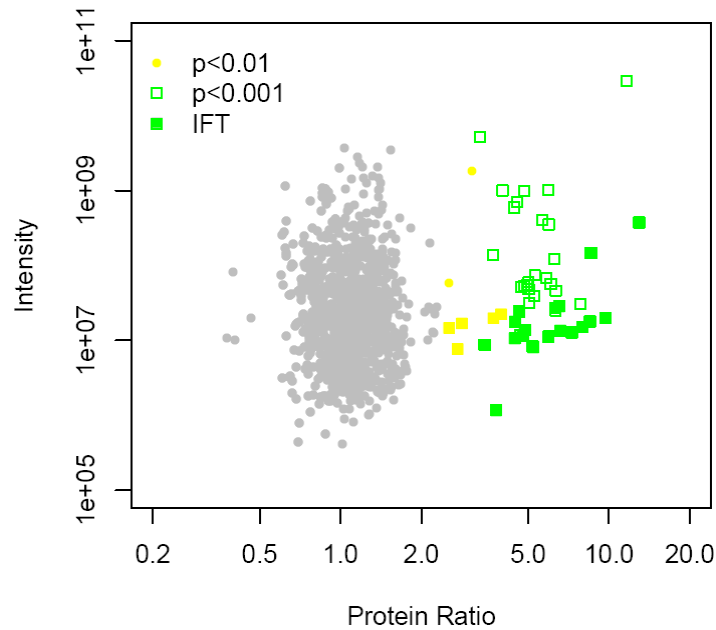
Figure 20: Distribution of the quantified proteins for lebercilin. Lebercilin and the SF-TAP alone were expressed in SILAC-labeled HEK293 cells. The lebercilin complex was enriched by Strep purification and the combined eluates were pre-fractionated by SDS-PAGE. The tryptic peptides were analyzed on an OrbitrapXL and quantified by MaxQuant. Three independent experiments were performed. For each experiment forward and reverse labeling was used. Only proteins that were quantified with at least three quantification events and identified by at least two unique peptides in at least two experiments were considered. The ratios and the corresponding intensities were plotted by an R script (supplemental file 1). Proteins which were highly significant enriched (p<0.001) are shown as green squares. Significantly enriched proteins (p<0.01) are shown as yellow squares. Filled squares are proteins belonging to the IFT machinery.

The proteins detected with p<0.01 were further analyzed by POINeT, which compares lists of proteins with protein interaction databases and the results visualized by Cytoscape. Complexes described in literature that were not annotated by POINeT as well as the lebercilin complex were annotated manually. The lines in the network do not necessarily mean direct interaction but rather co-occurrence in the same protein complex. It is not possible to distinguish between direct and indirect interactions by this approach and the focus of this study was laid on the analysis of protein complexes.

The largest group of proteins identified by the quantitative approach which was not detected by the SF-TAP approach was the group of intraflagellar transport (IFT) proteins and putative IFT proteins. The IFT proteins are part of a strongly conserved protein complex necessary for protein transport along cilia and flagella[28,106,196]. The IFT complex consists of three sub-complexes: complex A, B and the BBS complex[105] and links cargo to be transported along cilia to the motor proteins dynein or kinesin. These

complexes were mainly analyzed and described in *Chlamydomonas reinhardtii* or *C. elegans* but were at least partially shown to be conserved in mammalian cells, including ciliated kidney cells, including HEK293 cells and mammalian photoreceptors[30,31,106,197]. In addition to the well described components of these complexes, several putative IFT proteins exist which could not yet be assigned to a IFT sub-complex[28]. In addition to the described homologies, protein sequence comparison with *Chlamydomonas reinhardtii* proteins by protein blast analysis revealed 37% identity of WDR19 with IFT144, 37% identity of TTC21B with IFT139. These two proteins are classified as IFT complex A proteins[28]. 37% identity of HSPB11 was found with IFT25 and C11ORF60 is 94% identical to mouse IFT46, both IFT complex B proteins. All of the known IFT complex B proteins were identified by the Strep-SILAC method and additionally four of the putative IFT proteins and three proteins of IFT complex A (Figure 21, Table 15). At least one of the putative IFT proteins (RABL5) was described to be involved in the IFT[114], the others were inferred by homology to known *Chlamydomonas* or *C. elegans* IFTs and many of them were shown to be involved in the IFT machinery in these organisms[28,107-109]. Additionally to the IFT proteins, both dynein light chain isoforms and six 14-3-3 isoforms as well as two casein kinase 2 subunits and the uncharacterized C20orf11 protein were identified as specific lebercilin complex components. The kinesin-like protein KIF7 is a homolog of the drosophila protein costal-2 which mediates hedgehog signalling[198] and *C.elegans* kinesin 2, which is a microtubule-associated motor protein. Yippee-like 5 belongs to a protein family, highly conserved among eukaryotes and localizes to the centrosome[199]. The identification of ribosomal proteins is probably an artifact due to the over expression of lebercilin. The other proteins identified as specific lebercilin complex components are only sparsely or even not characterized at all.

Taken together these data clearly show that lebercilin is part of the IFT machinery. Even though the data were generated in the HEK293T cell line, it is highly likely that the results can be transferred to the situation in human photoreceptors and that lebercilin is part of the IFT machinery at the connecting cilium in photoreceptors because the IFT is such a high conserved mechanism. A finding supporting this hypothesis is the localization of lebercilin to microtubular and ciliary structures as shown by electron- and confocal microscopy (Figure 5, Figure 6, Figure 7).

Gene Name	Protein	Uniprot	Normalized Ratio	Significance (B)
Lebercilin				
LCA5	Lebercilin	Q86VQ0	11.64	0.000
Microtubule associated motor proteins				
DYNLL1	Dynein light chain 1, cytoplasmic	P63167	12.96	0.000
DYNLL2	Dynein light chain 2, cytoplasmic	Q96FJ2	9.70	0.000
KIF7	Kinesin-like protein KIF7	Q2M1P5	3.80	0.000
Intraflagellar transport complex B				
IFT20	Intraflagellar transport protein 20 homolog	Q8IY31	7.97	0.000
HSPB11	Intraflagellar transport protein 25 homolog	Q9Y547	3.44	0.000
C11orf60	Intraflagellar transport protein 46 homolog	A8K0F6	6.59	0.000
IFT52	Intraflagellar transport protein 52 homolog	Q9Y366	7.27	0.000
TRAF3IP1	Intraflagellar transport protein 54 homolog	A8MTK4	5.23	0.000
IFT57	Intraflagellar transport protein 57 homolog	Q9NWB7	6.53	0.000
IFT74	Intraflagellar transport protein 74 homolog	Q96LB3	6.32	0.000
IFT80	Intraflagellar transport protein 80 homolog	Q9P2H3	4.46	0.001
IFT81	Intraflagellar transport protein 81 homolog	Q8WYA0	4.62	0.001
IFT88	Intraflagellar transport protein 88 homolog	Q13099	4.78	0.000
IFT172	Intraflagellar transport protein 172 homolog	Q9UG01	8.57	0.000
RABL4	Putative GTP-binding protein RAY-like	Q9BW83	3.96	0.002
Intraflagellar transport complex A				
TTC21B	Intraflagellar transport protein 139 homolog	Q7Z4L5	2.83	0.010
IFT140	Intraflagellar transport protein 140 homolog	Q96RY7	3.71	0.003
WDR19	Intraflagellar transport protein 144 homolog	Q8NEZ3	2.54	0.009
Putative intraflagellar transport proteins				
CLUAP1	Clusterin-associated protein 1	Q96AJ1	8.49	0.000
RABL5	Rab-like protein 5	Q9H7X7	4.86	0.000
TTC26	Tetratricopeptide repeat protein 26	A0AVF1	4.47	0.001
TTC30A	Tetratricopeptide repeat protein 30A	Q86WT1	5.95	0.000
Centrosomal proteins				
CEP170	Centrosomal protein of 170 kDa	Q5SW79	6.34	0.000
YPEL5	Protein yippee-like 5	P62699	5.17	0.000
Cellular signaling proteins				
CSNK2A2	Casein kinase 2, alpha	P19784	5.65	0.000
CSNK2B	Casein kinase 2, beta	Q5SRQ6	5.97	0.000
WDR26	WD repeat-containing protein 26	Q9H7D7	4.84	0.000

Gene Name	Protein	Uniprot	Normalized Ratio	Significance (B)
	Adapter proteins			
YWHAE	14-3-3 protein epsilon	P62258	3.31	0.000
YWHAG	14-3-3 protein gamma	P61981	4.84	0.000
YWHAH	14-3-3 protein eta	Q04917	4.43	0.000
YWHAQ	14-3-3 protein theta	P27348	4.02	0.000
YWHAZ	14-3-3 protein zeta/delta	P63104	3.08	0.001
	Varia			
KIAA0564	KIAA0564	A3KMH1	6.25	0.000
RPLP0	60S acidic ribosomal protein P0	P05388	3.70	0.000
RPL12	60S ribosomal protein L12	P30050	4.72	0.000
RPL13	60S ribosomal protein L13	P26373	5.30	0.000
RPL13A	60S ribosomal protein L13a	P40429	5.02	0.000
RPL21	60S ribosomal protein L21	P46778	6.33	0.000
RPL4	60S ribosomal protein L4	P36578	5.86	0.000
RPL8	60S ribosomal protein L8	P62917	5.26	0.000
ATAD3A	ATPase family AAA domain-containing protein 3A	Q9NVI7	2.74	0.001
RPL7	ENSP00000381447	A8MY04	5.02	0.000
RP11	FLJ53585	B4DEN1	5.05	0.000
RPL6	FLJ92129	B2R4K7	4.96	0.000
PTCD3	Pentatricopeptide repeat-containing protein 3	Q96EY7	2.54	0.006
PGAM5	Phosphoglycerate mutase family member 5	Q96HS1	7.81	0.000
USP9X	Probable ubiquitin carboxyl-terminal hydrolase FAF-X	Q93008	4.54	0.000
C20orf11	Protein C20orf11	Q9NWU2	6.07	0.000
RMND5A	Protein RMD5 homolog A	Q9H871	4.67	0.001

Table 15: Specific complex components of the lebercilin complex. Lebercilin and the SF-TAP alone were expressed in SILAC labeled HEK293 cells. The lebercilin complex was enriched by Strep purification and the combined eluates were pre-fractionated by SDS-PAGE. The tryptic peptides were analyzed on an OrbitrapXL and quantified by MaxQuant. Three independent experiments were performed. For each experiment forward and reverse labeling was used. Only proteins that were quantified with at least three quantification events and identified by at least two unique peptides in at least two experiments were considered. The colors correspond to the colors in Figure 20.

2.2.2   Characterization of the mammalian IFT complex B

The IFT complex B which was found associated with lebercilin was described to exist in various organisms, tissues and cell types[28,106,110]. Several of the putative IFT proteins were described to be part of the IFT machinery before[107,108,114]. Even though IFT was described in kidney as

well, it is not clear whether the very same complex exists in HEK293T cells and especially if the complex inventory described so far is comprehensive.

Protein name	Uniprot	IFT88-SF 1	IFT88-SF 2	IFT88-SF 3
Intraflagellar transport complex B				
Intraflagellar transport protein 25	Q9Y547		100% (3)	100% (2)
Intraflagellar transport protein 46	Q9NQC8	100% (2)	100% (5)	100% (2)
Intraflagellar transport protein 52	Q9Y366	100% (15)	100% (18)	100% (9)
Intraflagellar transport protein 54	A8MTK4		100% (2)	100% (2)
Intraflagellar transport protein 57	Q9NWB7	100% (10)	100% (19)	100% (9)
Intraflagellar transport protein 74	Q96LB3	100% (5)	100% (9)	100% (5)
Intraflagellar transport protein 80	Q9P2H3	100% (4)	100% (14)	100% (3)
Intraflagellar transport protein 81	Q8WYA0	100% (3)	100% (4)	
Intraflagellar transport protein 88	Q13099	100% (34)	100% (41)	100% (38)
Intraflagellar transport protein 172	Q9UG01	100% (34)	100% (41)	100% (38)
RABL4	Q9BW83	100% (2)	100% (3)	100% (2)
Putative IFTs				
Tetratricopeptide repeat protein 26	A0AVF1	100% (8)	100% (12)	100% (10)
Tetratricopeptide repeat protein 30A	Q86WT1	100% (2)	100% (9)	100% (3)
Tetratricopeptide repeat protein 30B	Q84P2	89% (1)	100% (3)	
Clusterin-associated protein 1	Q96AJ1	100% (3)	100% (2)	94% (1)
Rab-like protein 5	Q9H7X7	100% (4)	100% (5)	100% (3)
Varia				
Transmembrane protein 43	Q9BTV4	100% (8)	100% (6)	100% (7)
Sodium/potassium-transporting ATPase	P05023	100% (8)	100% (5)	100% (6)
ADP/ATP translocase 1	P12235	89% (1)	100% (2)	100% (2)
ADP/ATP translocase 2	P05141	100% (2)	100% (2)	
Tubulin beta-6 chain	Q9BUF5		100% (2)	100% (2)

Table 16: IFT88-SF-TAP analysis. SF-TAP tagged IFT88 was expressed in HEK293T cells and the complexes purified by SF-TAP. The eluates were concentrated by methanol-chloroform precipitation, subjected to in-solution tryptic digestion and analyzed on an OrbitrapXL. The data were analyzed by Sequest and Scaffold. Three independent experiments were performed and only proteins that were identified in at least one experiment with more than one peptide and in a second experiment with more than 90% probability were considered.

To analyze the IFT complex B in HEK293T cells, N-terminally SF-TAP tagged IFT88 was expressed and purified from HEK293T cells and analyzed by LC-MS/MS on the OrbitrapXL. As negative control, HEK293T cells expressing the SF-TAP tag alone were used. The OrbitrapXL data were analyzed by Sequest database search followed by Scaffold analysis. The experiment was repeated three times and only proteins, identified at least

once with at least 2 peptides (minimum peptide probability: 80%) and in a second experiment with at least one highly significantly identified peptide (minimum peptide probability: 95%) were considered. The minimum protein probability was set to 90%.

Besides the known IFT complex B except IFT20, all putative IFT proteins that were found to be complex components of the lebercilin protein complex, could be identified in the IFT88 SF-TAP analysis (Table 16). These putative IFTs were: TTC26, TTC30A/B, CLUAP1 and RABL5. Since no IFT complex A proteins were identified (Table 16), this strongly suggests that these proteins, previously described as putative IFT proteins[28], are actually part of the IFT complex B or at least closely associated with this complex. This is the first time that these proteins could be assigned to one of the sub-complexes of the IFT and it is the first near comprehensive description of the IFT complex B in a mammalian cell line.

A graphical representation of the lebercilin protein complex shown in Figure 21 was created based on the data of the identification of specific complex components by the Strep-SILAC workflow, the IFT88-SF-TAP data as well as literature data. Additionally, the results of the POINeT analysis were included. The compilation of all data reveals that lebercilin associates with IFT proteins and among these mainly with IFT complex B proteins. IFT complex A proteins were identified as well but with lower significance. This could hint towards a weaker association of the complex A proteins with the lebercilin protein complex which could be an indirect interaction that is mediated by either the IFT complex B or by one of the motor proteins. Because these sub-complexes are closely associated and the role of each sub-complex is not entirely clear yet, it is not possible to describe a exact structure of the IFT complexes and the nature of the association of these sub-complexes with lebercilin.
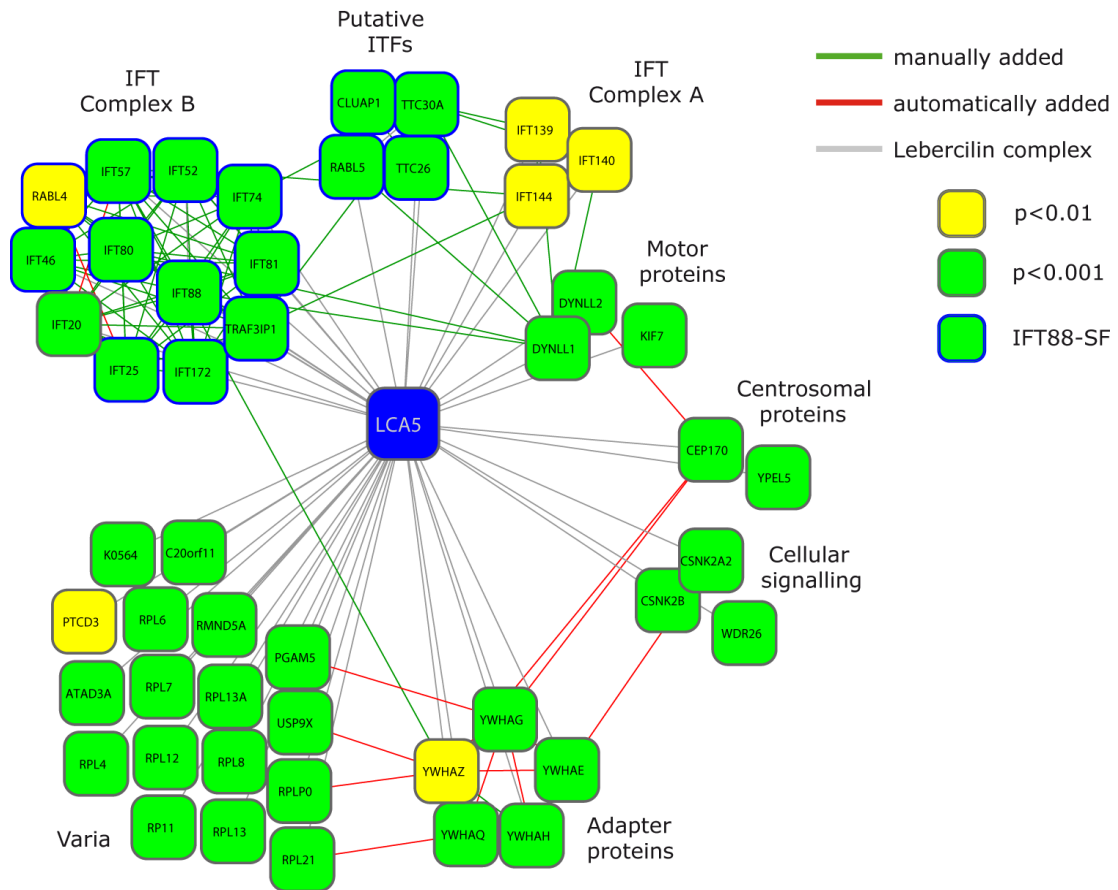
Figure 21: The lebercilin complex. The figure shows a graphical representation of the lebercilin protein complex. The colors correspond to the colors shown in Figure 20 and Table 15. Proteins shown in the table were subjected to POINeT analysis (red lines) and literature known complexes were annotated manually (green lines). The grey lines indicate that the proteins were identified as specific complex components of the lebercilin protein complex. This does not necessarily mean that these proteins interact directly. The proteins which are not in one complex but are involved in the same biological process are grouped and the process is annotated. Proteins that were identified in the IFT88-SF-TAP analysis are marked by blue surroundings.

2.2.3 Complex comparison of WT and mutated lebercilin

The detection of specific protein complex components of lebercilin strongly suggests that it is part of the IFT machinery. Additionally, the IFT88-SF-TAP analysis revealed that the putative IFTs are part of the IFT complex B. However, it is not clear by which mechanism the mutated forms of lebercilin leads to Leber congenital amaurosis since the molecular deficits associated with the genetically recessive disease are unknown. To elucidate alterations on the level of protein interactions associated with the mutant, the protein complexes of wild type lebercilin and the P493TfsX1 variant, derived from a Moroccan patient, were compared by the Strep-SILAC method. The P493TfsX1 mutant was selected because it is the mutant which leads to the longest transcript and thereby to the longest

protein when compared to the other mutants[200]. Wild type lebercilin and the P493TsfX1 mutant were expressed in HEK293T cells which were grown in SILAC medium. 48 hours post transfection, protein extracts were generated and the protein concentration was determined by a Bradford assay. Equal amounts of protein were used for each pair of samples. The protein complexes were enriched by Strep purification and the Strep-Tactin resins were combined before washing and elution. The samples were concentrated using 10 kDa centrifugal units to a final volume of approximately 20 µl before the samples were pre-fractionated by SDS-PAGE on 10% pre-cast gels. Each lane was cut into 8 segments and each segment was subjected to in-gel digestion and analyzed by LC-MS/MS on an OrbitrapXL using a 140 minute gradient for peptide separation as in the experiment for detection of specific complex components. The resulting spectra were analyzed and quantified using MaxQuant. The normalized protein ratios of all 3 (reverse and forward labeling in each experiment) experiments were calculated by MaxQuant.
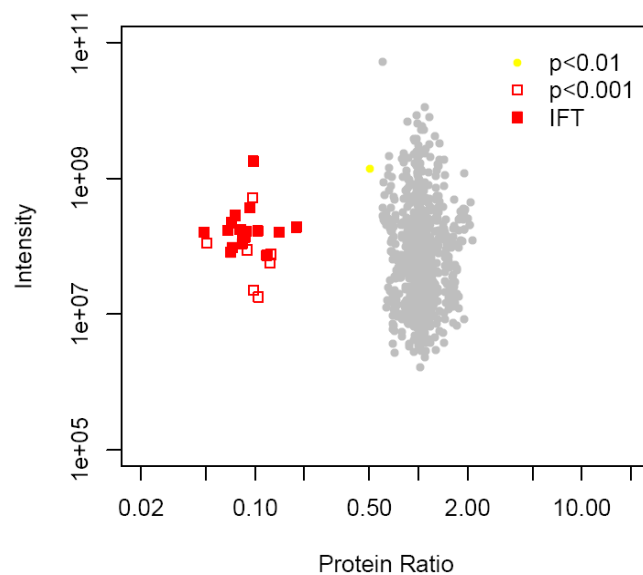
Figure 22: Distribution of the quantified proteins for lebercilin. Lebercilin and the P493T mutant were expressed in SILAC-labeled HEK293 cells. The lebercilin complex was enriched by Strep purification and the combined eluates pre-fractionated. Tryptic peptides were analyzed by LC-MS/MS and quantified by MaxQuant. Three independent experiments were performed. For each experiment forward and reverse labeling was used. Only proteins quantified with at least three quantification events and identified by at least two unique peptides in at least two experiments were considered. Proteins which were highly significant enriched (p<0.001) are shown as red squares. Significantly enriched proteins (p<0.01) are shown as yellow circles. Filled squares are proteins belonging to the IFT machinery.

The complex comparison revealed that most of the proteins identified and quantified show no significant alteration in binding (Figure 22). These proteins all cluster between rations of 0.5 and 2. However, 26 of the 51 proteins detected as specific lebercilin complex components show altered association with the P493TfsX1 mutant. These proteins all cluster between ratios of 0.15 and 0.05 and thereby bind more than 6-fold weaker to the protein complex of mutant lebercilin. Most of these proteins are part of the IFT machinery (Figure 22: filled squares).

Several components of the protein complex of WT lebercilin did not show significant alteration in association with the P493TfsX1 mutant. Under these were both dynein light chain isoforms and all 14-3-3 isoforms and most of the proteins not described to associate with microtubular structures. One Casein kinase isoform showed slightly weaker association with the protein complex of mutated lebercilin. The association of all IFT complex B proteins, IFT complex A and the putative IFT proteins associated significantly weaker with the complex of mutated lebercilin (Table 17, Figure 24). The centrosomal protein Yippee-like 5 as well as the C20orf11 protein, the KIA0564 protein, the Phosphoglycerate mutase family member 5 and the RMD5A protein showed a similar reduction in association with the complex of mutated lebercilin. Some proteins could not be identified in the comparison of WT and mutant lebercilin. However, most proteins were identified and only a part of them showed alterations in binding. Since nucleolin, nucleophosmin and HSP70 were not identified as specific lebercilin complex components, they were not considered in the complex comparison (supplemental table 3 and 4).

Gene Name	Protein	Uniprot	Normalized Ratio	Significance (B)
Lebercilin				
LCA5	Lebercilin	Q86VQ0	NA	
Microtubule associated motor proteins				
DYNLL1	Dynein light chain 1, cytoplasmic	P63167	0.87	0.234
DYNLL2	Dynein light chain 2, cytoplasmic	Q96FJ2	1.64	0.078
KIF7	Kinesin-like protein KIF7	Q2M1P5	NI	
Intraflagellar transport complex B				
IFT20	Intraflagellar transport protein 20 homolog	Q8IY31	0.09	0.000
HSPB11	Intraflagellar transport protein 25 homolog	Q9Y547	0.12	0.000
C11orf60	Intraflagellar transport protein 46 homolog	A8K0F6	0.05	0.000
IFT52	Intraflagellar transport protein 52 homolog	Q9Y366	0.07	0.000
TRAF3IP1	Intraflagellar transport protein 54 homolog	A8MTK4	0.07	0.000
IFT57	Intraflagellar transport protein 57 homolog	Q9NWB7	0.18	0.000
IFT74	Intraflagellar transport protein 74 homolog	Q96LB3	0.07	0.000
IFT80	Intraflagellar transport protein 80 homolog	Q9P2H3	0.09	0.000
IFT81	Intraflagellar transport protein 81 homolog	Q8WYA0	0.08	0.000
IFT88	Intraflagellar transport protein 88 homolog	Q13099	0.08	0.000
IFT172	Intraflagellar transport protein 172 homolog	Q9UG01	0.10	0.000
RABL4	Putative GTP-binding protein RAY-like	Q9BW83	0.09	0.000
Intraflagellar transport complex A				
TTC21B	Intraflagellar transport protein 139 homolog	Q7Z4L5	0.08	0.000
IFT140	Intraflagellar transport protein 140 homolog	Q96RY7	0.09	0.000
WDR19	Intraflagellar transport protein 144 homolog	Q8NEZ3	0.08	0.000
Putative intraflagellar transport proteins				
CLUAP1	Clusterin-associated protein 1	Q96AJ1	0.14	0.000
RABL5	Rab-like protein 5	Q9H7X7	0.05	0.000
TTC26	Tetratricopeptide repeat protein 26	A0AVF1	0.10	0.000
TTC30A/B	Tetratricopeptide repeat protein 30A/B	Q86WT1	0.07	0.000
Centrosomal proteins				
CEP170	Centrosomal protein of 170 kDa	Q5SW79	NI	
YPEL5	Protein yippee-like 5	P62699	0.10	0.000
Cellular signaling proteins				
CSNK2A2	Casein kinase 2, alpha	P19784	0.73	0.078
CSNK2B	Casein kinase 2, beta	Q5SRQ6	0.51	0.002
WDR26	WD repeat-containing protein 26	Q9H7D7	0.13	0.000

Gene Name	Protein	Uniprot	Normalized Ratio	Significance (B)
Adapter proteins				
YWHAE	14-3-3 protein epsilon	P62258	1.09	0.433
YWHAG	14-3-3 protein gamma	P61981	1.23	0.300
YWHAH	14-3-3 protein eta	Q04917	1.20	0.326
YWHAQ	14-3-3 protein theta	P27348	1.10	0.429
YWHAZ	14-3-3 protein zeta/delta	P63104	0.95	0.362
Varia				
KIAA056	KIAA0564	A3KMH1	0.10	0.000
RPLP0	60S acidic ribosomal protein P0	P05388	1.89	0.032
RPL12	60S ribosomal protein L12	P30050	1.42	0.165
RPL13	60S ribosomal protein L13	P26373	NI	
RPL13A	60S ribosomal protein L13a	P40429	1.92	0.028
RPL21	60S ribosomal protein L21	P46778	1.43	0.159
RPL4	60S ribosomal protein L4	P36578	NI	
RPL8	60S ribosomal protein L8	P62917	2.06	0.017
ATAD3A	ATPase family AAA domain-containing protein 3A	Q9NVI7	NI	
RPL7	ENSP00000381447	A8MY04	1.94	0.026
RP11	FLJ53585	B4DEN1	NI	
RPL6	FLJ92129	B2R4K7	1.82	0.041
PTCD3	Pentatricopeptide repeat-containing protein 3	Q96EY7	1.66	0.073
PGAM5	Phosphoglycerate mutase family member 5	Q96HS1	0.12	0.000
USP9X	Probable ubiquitin carboxyl-terminal hydrolase FAF-X	Q93008	1.18	0.341
C20orf11	Protein C20orf11	Q9NWU2	0.09	0.000
RMND5A	Protein RMD5 homolog A	Q9H871	0.10	0.000

Table 17: The protein complex of mutated lebercilin. Lebercilin and the P493TfsX1 mutant were expressed in SILAC labeled HEK293 cells. The lebercilin complexes were enriched by Strep purification and the combined eluates were pre-fractionated by SDS-PAGE. The tryptic peptides were analyzed on an OrbitrapXL and quantified by MaxQuant. Three independent experiments were performed. For each experiment forward and reverse labeling was used. Only proteins that were quantified with at least three quantification events and identified by at least two unique peptides in at least two experiments were considered. The colors correspond to the colors in Figure 22.

To confirm the results obtained by the Strep-SILAC approach through an independent experimental method, lebercilin and the P493TfsX1 mutant were again expressed in HEK293T cells and the complexes were purified by Strep purification separately. The eluates were separated by SDS-PAGE and analyzed by Western blot using the anti-Flag-M2 antibody to detect lebercilin, the IFT74 antibody, the IFT88 antibody as well as the 14-3-3ε antibody. The result showed that wild type lebercilin and the P493TfsX1

mutant were expressed at the same level and equal amounts were purified. Additionally, the results show that IFT74 and IFT88 specifically bind to the lebercilin complex and that they do not bind to the Strep-Tactin resin or the tag itself. In addition, the analysis confirmed that association of IFT74 and IFT88 with the complex of mutated lebercilin is lost. There is no change in association for the 14-3-3 isoforms (Figure 23: right panel) which is also in-line with the results obtained by MS. To exclude differences in association due to alterations in expression of the proteins, the protein lysates used for the Strep purification were analyzed by Western blot using the same antibodies as used for the analysis of the eluates. No change was visible for lebercilin or for IFT74, IFT88 and 14-3-3 isoforms (Figure 23: left panel).
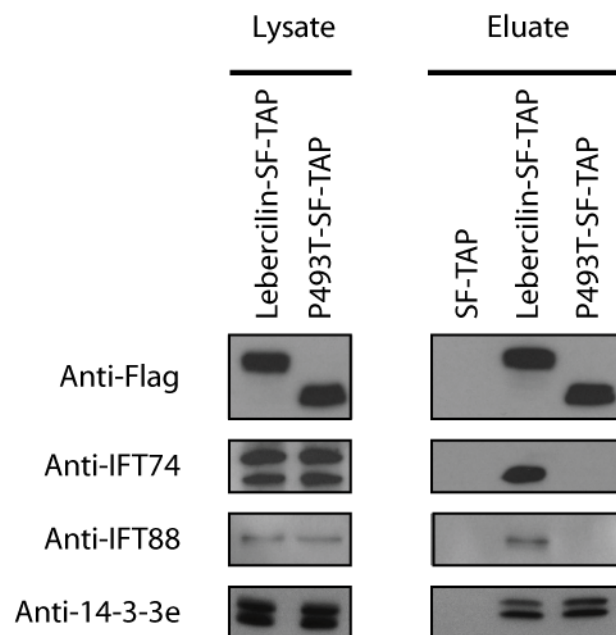
Figure 23: Validation of IFT interaction. Lebercilin and the P493TfsX1 mutant were expressed in HEK293 cells and the complexes enriched by Strep purification. The eluates were concentrated by 10 kDa cut-off spin columns, separated by SDS-PAGE analyzed by Western blot using specific antibodies to IFT74, IFT88, 14-3-3ε and Flag (lebercilin). The lysates are shown to assure equal expression of the proteins and the eluates. The SF-TAP alone was used as negative control.

A graphical representation of the lebercilin protein network is shown in Figure 24. In this network, the data derived from the IFT88 SF-TAP analysis, the detection of specific protein complex components by the Strep-SILAC approach and the protein complex comparison are compiled. It clearly shows that a majority of the proteins loosing their association with the complex of mutated lebercilin are IFT proteins. All confirmed IFT

proteins of the IFT complexes A and B as well as the putative IFT proteins, demonstrated to be actually part of the IFT complex B show the same behavior as do several proteins that are not described as IFT proteins. Since these proteins were not identified in the SF-TAP analysis of IFT88 and all known complex B components were, it is unlikely that these are part of the IFT complex B. However, they could be yet not discovered components of the IFT complex A or other IFT sub-complexes. Further experiments need to be done to identify further components, especially of IFT complex A.
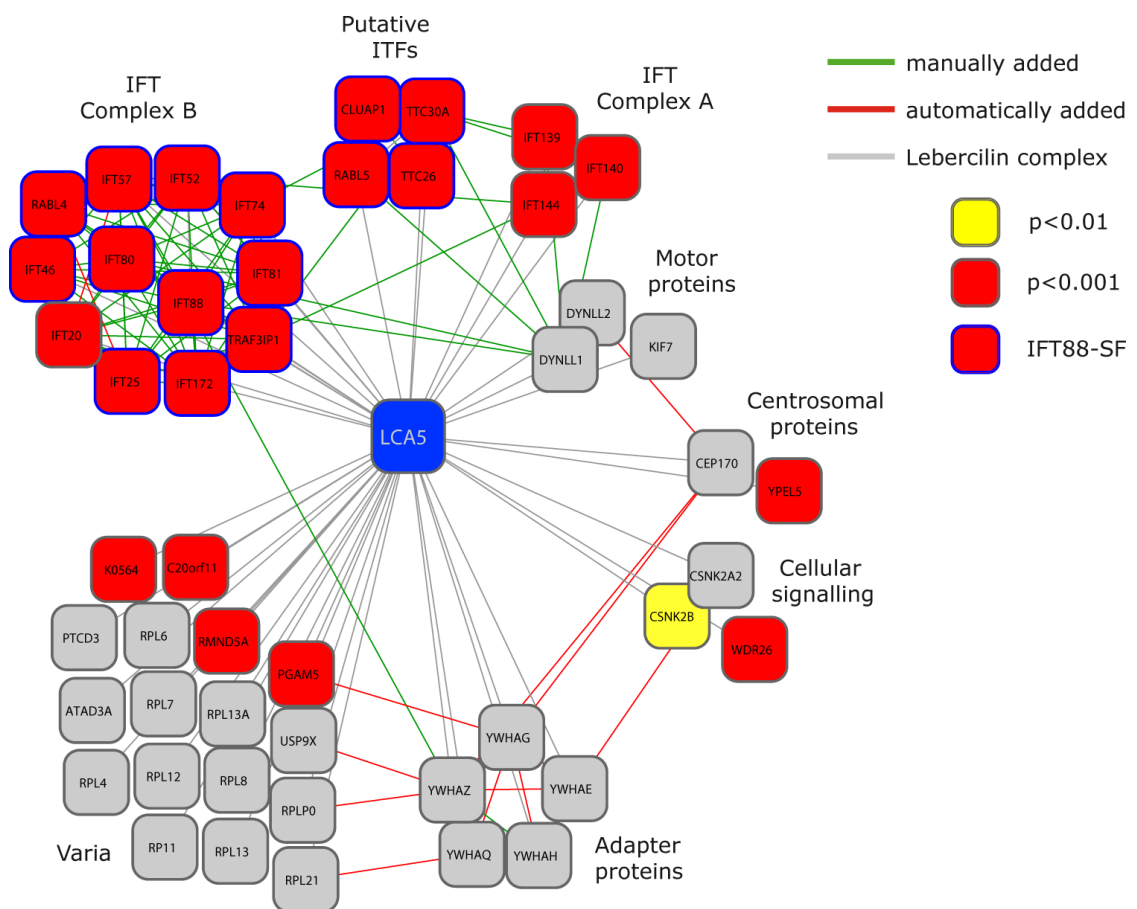
Figure 24: The complex of mutated lebercilin. The Cytoscape network is was produced as described in Figure 20. The network shows the proteins that are weaker associated with the complex of mutated lebercilin in red (p<0.001) and yellow (p<0.01). The colors correspond to the colors in Figure 22 and Table 17.

These results clearly demonstrate that there are pronounced differences in complex composition between the complex of wild type and mutated lebercilin. The most prominent alteration is the clear reduction of association with the IFT complex B as well as the putative IFT proteins.

This suggests that the loss of association with the IFT complex is the mechanism by which mutation of lebercilin causes the disease.

The application of the Strep-SILAC workflow and the results obtained by this approach demonstrate that it is a valuable tool for the identification of components of protein complexes and also for the comparison of protein complexes.

H Discussion

1. SF-TAP, a new method for protein complex analysis

In the first part of this study, the optimization and evaluation of the new tandem affinity purification strategy, the SF-TAP, is described. First of all, the method was optimized to achieve highly enriched protein complexes in minimal time with minimal loss of efficiency. The goal of this approach was to increase the possibility to co-purify complex components, characterized by quick dissociation rates or lower affinity. Therefore the incubation times were reduced, resulting in a complete workflow for purification of protein complexes completed within two and a half hours. This is substantially faster than the original TAP method and compared to other, newly developed methods as well[34,39-41,55-58,201]. Basically all of these TAP tags are reliant on proteolytic cleavage after the first purification step. The cleavage process is dependent on high quality proteases, which need time for the process of cleavage. Additionally, proteases are less efficient at 4°C and for that reason the temperature is elevated in many protocols to achieve efficient cleavage. The elevated temperature increases the dissociation and association rates of potential complex components, thereby leading to less efficient enrichment of those proteins or even to their complete loss. The major advantages of the SF-TAP tag is that the protein complexes can be eluted in both steps by competition, there is no need for proteolytic cleavage. This improvement saves time and simplifies the procedure because all steps can be done in one single buffer system. Additionally, in the original TAP (TAPo) procedure, the binding of calmodulin to the CBP domain is dependent on calcium which is chelated by EDTA for elution. This change in calcium concentration could interfere with calcium dependent binding of proteins to be isolated and is therefore not desired. Especially for the MAPK pathway, alterations in calcium concentrations might strongly alter the outcome of an experiment since a cross-talk between the calcium dependent PKA pathway, for example, exists and especially B-Raf can be phosphorylated and regulated by PKA and PKC[202-205].

Despite less time consuming purification using the SF-TAP method, the efficiency is equal or better than the efficiency of the TAPo method. The efficiency was compared for B-Raf and 14-3-3ε. For 14-3-3ε, the expression was equal for both tags and the purification efficiency as well.

For B-Raf, the comparison is difficult since the expression of SF-TAP tagged B-Raf was significantly higher when compared to the TAPo tagged version. Higher amounts of DNA for transfection led to comparable levels of B-Raf and to comparable purification efficiencies that made quantification possible. The efficiency of the SF-TAP approach was still at least comparable. The expression levels of SF-TAP tagged proteins might be higher because it only adds approximately 5 kDa to the protein and the original tag approximately 25 kDa. This large tag might interfere with the function of the protein since it might alter the protein structure, mask signal peptides or functional domains of the protein. Nevertheless, it can not be excluded that a smaller tag leads to similar folding problems. The small size of the tag additionally increases the risk that the tag is folded into the protein structure. This would lead to impairment or even prevention of binding of the tag to the affinity capture resin. So far however, none such impairment has been observed with at least 25 different proteins being tagged and purified.

In this study, the efficiency of the SF-TAP tag was only compared to the original tag, composed of a protein A tag, the TEV cleavage site and the calmodulin binding peptide domain and not to other variations of this tag like the one described by Burckstummer and colleagues[39]. The authors claimed that the efficiency of a tag, in which the protein A was replaced by protein G and the calmodulin binding peptide by a streptavidin-binding peptide, was up to 10-fold higher when compared to the original tag. However, this might be highly dependent on the bait protein since the efficiency of the SF-TAP tag was between 40% and 50% for the purification of B-Raf, 14-3-3ε and MEK1 from cell lines, stably expressing the bait proteins. A 10-fold higher efficiency is in this case not achievable and with efficiencies up to 50%, there is not much room for improvement.

One more aspect that should be considered when comparing the different technologies is the complexity of the protocol, which is strongly simplified for the SF-TAP procedure. The procedure is done in one buffer system without additional washes to change the buffer for TEV cleavage or calmodulin binding. Additionally the avoidance of proteolytic cleavage saves time, especially hands on time. This straightforward system greatly improves the throughput and in combination with the versatile GATEWAY cloning system, it makes the technology suitable for high throughput studies as well.

110

So far it is not proven that the SF-TAP is superior to the original TAP tag[162]. In contrast to the TAPo which was extensively tested[34] and used for the large scale analysis of protein complexes in yeast[33], the SF-TAP tag approach was used for the analysis of few proteins only. Even though it was possible to efficiently enrich the protein of interest and to analyze the protein complexes, the application for large scale protein complex mapping could finally show its usefulness for the analysis of protein complexes.

Even though the second step of every TAP approach is most likely the crucial one for the loss of transient complex components, the overall time for the purification procedure influences the probability for identifying these components. In the second step, the eluted proteins are strongly diluted because most of the proteins present in the lysate and a high proportion of individual complex components may be lost during the first purification step. This shifts the equilibrium from constant association and dissociation towards dissociation. If the dissociation and association rates are high and the affinity of a potential binder not high enough, it gets lost during this process. As mentioned earlier, this is a major disadvantage of the SF-TAP method but of other TAP strategies as the original one as well. Here lies one advantage of single-step purifications which can result in co-purification of transient or less stable associated proteins. As shown in several examples, including the protein complex analysis of lebercilin described in this work, single-step purifications have dramatic advantages in detection of those interactions. The efforts for the identification of complex components might be higher if the eluate fractions containing the protein complexes are less pure. This however, can be compensated by the identification of highly interesting interactions, as shown for example for lebercilin. Several alternative strategies and approaches, involving single-step purifications were described and successfully used for the analysis of protein complexes as well.

One important and successfully applied strategy for the analysis of protein complexes by single-step purification utilizes the Strep-tag or the Strep-tag II. In combination with the Strep-Tactin it is in many cases possible to use this system for the simple, efficient purification and analysis of protein complexes without the need for a second purification step as in TAP approaches[53,54]. The power of this method was demonstrated in various publications, showing its value for the analysis of protein complexes. The Strep system was for example applied to identify potential substrates of

the periplasmic chaperone Skp in *Escherichia coli*. Therefore, the protein complexes formed by WT Skp and the ones formed by a mutant with a more hydrophobic interior where analyzed by high resolution 2D-gel electrophoresis in combination with mass spectrometric protein identification. The analysis resulted in the identification of more than 30 interacting proteins, including known as well as several new potential substrate proteins for this chaperone[206]. The Strep system was not only used for the analysis of protein complexes in bacteria, it was also applied for the analysis of protein complexes in mammalian cells. Juntilla and co-workers demonstrated, that both, the Strep- tag II and a Strep-tag II double tag, for which the two tags are separated by a linker region, are suitable for the identification of complex components of the PP2A subunit PR65 from a human fibrosarcoma cell line, stably expressing the tagged proteins[49]. The double Strep-tag II was however reported to have a higher binding affinity towards Strep-Tactin compared to a single Strep-tag II which can be helpful, especially when aiming to purify low abundant target proteins. This double Strep-tag II is also known as One-STrEP-tag (IBA). Utilizing this tag, Groth and co-workers could identify a protein complex of the Chaperone Asf1 in HeLa cells and could thereby learn that this protein functions as histone acceptor and donor during DNA replication and that it is linked via histones H3-H4 to putative replicative helicases[207]. The One-STrEP-tag could also be used to identify novel interaction partners of the IκB-kinase-associated protein, especially Filamin A. Further characterization of this interaction led to the finding that these proteins are both involved in cytoskeleton remodeling, possibly cooperatively[208].

To further increase the affinity of the tag towards Strep-Tactin, Schaffitzel and Ban even used a triple repeat of the Strep-tag II. This increase in affinity was necessary to purify ribosome nascent chains with sufficient efficiency. In combination with arrest sequences to stabilize the nascent chain in complex with ribosomes, affinity purification using this tag led to the possibility of purifying stable ribosome nascent chains and analyzing the complexes by 3D cryo-electron microscopy[209].

A modification of this method, similar to the SF-TAP tag, involves a combination of a double Strep-tag II and a Flag tag (Two-TAP, IBA). This tag was successfully applied in combination with SDS-PAGE and mass spectrometry to identify new protein complex components and thereby new function in cytokinesis[210] and HIV budding[211] of proteins known to be

involved in vesicular sorting. Even though, only the Strep-tag was used for the purification, the use of the Flag-tag enabled a sensitive and specific detection of the proteins by anti-Flag antibodies as an alternative to detection by the Strep tag using Strep antibodies or Strep-Tactin for example.

That the achievable purity of Strep-purified protein complexes can be really high was also demonstrated by Ostermeier and colleagues. Using eluates of Strep purified F_v-fragments in complex with cytochrome C oxidases they could produce clean crystals suitable for high resolution x-ray crystallographic studies[212].

A highly interesting approach is the combination of the Strep system with reversible, chemical cross-linking using formaldehyde. This approach called SPINE (Strep-protein interaction experiment)[213] is of special interest because a quick single-step Strep purification can be combined with stringent washing and thereby enabling high purity. Due to the combination with *in vivo* cross-linking, this approach also enables the detection of transient and weakly associated protein complex components. However, it could also lead to the identification of artifacts due to the *in vivo* cross-linking. The authors decided to use the Strep system especially because of the high purity of the eluted protein complexes, when compared to the His-tag system[213]. But also compared to other tag-based affinity purification systems, the Strep system is highly efficient and a high purity can be achieved[51].

These reports demonstrate that there are several methods, mainly utilizing the Strep-based affinity purification system to achieve eluates with protein complexes of high purity. In many cases, a single Strep-tag II can be sufficient to achieve efficient and pure complexes. It should be additionally noted that the fusion of only a Strep-tag to the protein of interest instead of a combination of tags can be favorable because the addition of long linear stretches or large tags can lead alterations in the biochemical properties of the protein of interest or can even lead to aggregation or increased non-specific binding[206]. More than one tag can be added, however, to produce higher affinity Strep tags, or addition of a alternative tag can be favorable if the purity of the single step purifications is not high enough and a second, independent purification step is necessary. In these cases, the SF-TAP can be a good alternative.

A question which can not easily be answered by the SF-TAP approach but neither with any other affinity tag-based purification approach is how the protein complexes assemble *in vivo.* This would only be possible if knock-in animals would be produced from which the protein complexes could be purified. Without undertaking this laborious task, if good antibodies exist, antibody-based affinity purifications or pull-downs can be used for the same purpose with less effort and similar chances for success.

2.   Analysis of the MAPK protein network by SF-TAP

The first prove of principle experiment for the SF-TAP system was the application to the analysis of the MAPK network. In this application I could identify many of the network components known from literature by the gel based approach already. The LC-MALDI approach resulted in even more identifications of complex components, especially for 14-3-3ε. For B-Raf, C-Raf and MEK1 the main effectors MEK1 and MEK2, known interacting adapter proteins as well as the HSP90 and CDC37 chaperones were identified. The two chaperones where shown before to bind to B-Raf, especially for highly active B-Raf mutants. It was shown, that these mutants can only stay active with the support of HSP90[214,215]. Since it is known that 14-3-3 proteins bind a variety of proteins[214,215], it is not surprising that this bait led to the highest number of identifications. 14-3-3 proteins are known to bind proteins in a phosphorylation-dependent manner to two conserved motifs[216,217]. Up to date more than 300 potential interaction partners were described[218]. For most of the proteins identified here it has been previously shown that they are interaction partners of 14-3-3 proteins and for many of them that the interaction has physiological impact[35,36,214,215,219-221]. This demonstrates that the SF-TAP method is not only efficient and straightforward but that the protein complexes that were identified are real complexes and not artifacts due to the mild over expression or tagging of the protein.

However, it should also be mentioned that only the core complexes, the most stably associated proteins, were identified. Ras for example was not identified as Raf-1 or B-Raf complex and Erk1/2 not as MEK1 protein complex component[10,183]. As discussed before, this again demonstrates that only stable protein complex components can be detected with TAP approaches.

3. Analysis of lebercilin complexes

3.1 Analysis of the lebercilin protein complex by SF-TAP

By application to the MAPK network, the SF-TAP technology was proven to be suited for the analysis of real protein complexes. Therefore it was applied to the protein complex analysis of a, at that time, uncharacterized protein lebercilin, found to be mutated in Leber congenital amaurosis. The resulting protein complex consists of many proteins, previously shown to be associated with microtubules, centrosomes or cilia. An important association was found with the motor protein dynein, strengthening the link to microtubular dynamics and suggesting involvement in minus end-directed microtubular transport. Nucleophosmin has been shown to localize to centrosomes[187] where it is involved in the regulation of centrosome duplication by interaction with ROCKII[222]. It is also a substrate of one of the key kinases in centrosome function, the polo-like kinase 1 (Plk1)[223] and interestingly, was found to bind to RPGR, involved in X-linked retinal dystrophies[111]. Similar to nucleophosmin, nucleolin is a mainly nuclear multifunctional protein found to be involved in nucleo-cytoplasmic shuttling[224] and to bind extra nuclear to casein kinase 2 in photoreceptors[188,189,225]. Casein kinase 2 is also part of the lebercilin inter-actome. It was described to regulate microtubular dynamics[226], associate with the centrosome[192] and regulate localization of several proteins to cilia via phosphorylation[193,194]. Finally, 14-3-3 scaffold proteins that bind and regulate function of many phosphorylated proteins[185] have been previously described to associate with centrosomal protein complexes[184,186]. Since the SF-TAP analysis from HEK293 cells might not be the ideal model system for photoreceptors, the protein complex needed to be validated[73]. This validation was done by Monika Beer in our department by co-immunoprecipitation from porcine retinae using a lebercilin specific antibody. The association of 14-3-3 protein, nucleolin, nucleophosmin, HSP70 could be confirmed. Dynein could not be tested due to the unavailability of an antibody. However, the two dynactin subunits p50-dynamitin and p150[glued] could be detected (Figure 25). These two proteins are directly bound to dynein and are involved in linking it to the cargo to be transported. The confirmation by co-immunoprecipitation shows that the results obtained by SF-TAP analysis in HEK293 cells reflect protein interactions that are also occurring in the retina where endogenous

lebercilin is expressed.

Taken together, immunohistochemical studies (Figure 6, Figure 7) immunoprecipitation studies (Figure 25) as well as results from protein complex analysis studies presented here suggest that lebercilin is a true ciliary protein. Additionally, the SF-TAP analysis showed by identification of dynein that lebercilin is possibly involved in the retrograde ciliary transport. When considering the photoreceptor context and the disease phenotype, lebercilin possibly plays an important role in the ciliary transport of outer segments components from the inner to the outer segment or rather the other way around which is important for recycling of components.
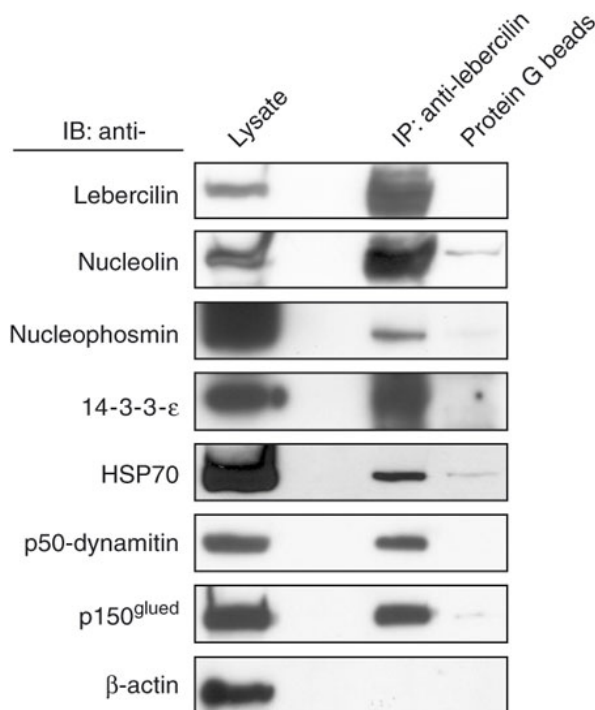
Figure 25: Co-immunoprecipitation of lebercilin from porcine retinae. Lebercilin was affinity purified using a lebercilin specific antibody from porcine retinae. The eluates were separated by SDS-PAGE and analyzed by Western blotting against several proteins identified by lebercilin SF-TAP analysis. By this method, nucleolin, nucleophosmin, 14-3-3ε and HSP70 could be confirmed by this method. Actin was used as negative control. Dynein is not shown but the two subunits of the dynactin complex which link dynein to the cargo. This figure was reprinted with the friendly permission of the Nature Publishing Group[73].

3.2 Quantitative lebercilin complex analysis

3.2.1 Strep-SILAC compared to SF-TAP

For an in-depth analysis of the lebercilin protein complex, a one step enrichment of SF-TAP tagged lebercilin was combined with quantitative mass spectrometry. SF-TAP analysis in combination with immuno-precipitation from porcine retina showed that HEK293 cells are a suitable model for identifying components of the lebercilin interactome. In comparison to the SF-TAP approach, the Strep purification alone results in

enrichment rather than a purification of the protein complex. Therefore, an additional, truly quantitative approach appeared to be necessary for the identification of further specific complex components. This is obvious when looking at the number of non-specific proteins identified in the approach. Altogether, several hundreds of proteins were identified in the single-step Strep purifications from which only 74 were significantly enriched in comparison to the control (Figure 20, supplemental table 3). The specific drawback here is that due to the threshold selected, specific binders may not be detected because they bind to the resin equally or similarly strong as to lebercilin or its complex. Nucleolin and nucleophosmin for example which were both detected in the SF-TAP approach could not be detected as specific complex component with the threshold selected. Nucleolin, for example, was enriched but the p-value was slightly above the significance threshold selected. Nucleophosmin was only weakly enriched and was shown to bind strongly to the resin (supplemental table 3). These examples demonstrate the drawback of the single-step strategy. It is much more impure when compared to the SF-TAP approach and thereby complex components with high affinity to the used affinity resin can not be detected. Nevertheless, the value of the Strep-SILAC approach is immense because the number of specific complex components was significantly higher than compared to the SF-TAP approach and at least one sub-complex (IFT complex) and a second motor protein (KIF7) could be identified which were not detected by the conventional SF-TAP approach. The advantage of the method is that given the quick single-step purification leads to impurity on one hand but enables the detection of transient and weak complex components on the other hand. This is possible especially because the second purification step is not necessary in the SILAC approach where strong dilution of binders is happening.

Most importantly, the SILAC approach revealed a new set of complex components. By combining single-step purification in combination with quantitative mass spectrometry, the association of many IFT proteins and a kinesin isoform with the lebercilin protein complex could be detected. All known proteins of IFT complex B, four putative IFT proteins and three IFT complex A proteins were identified. Two centrosomal proteins, the CEP170 protein and the protein yippee-like 5 (YPEL5), two casein kinase isoforms, one additional cellular signaling associated protein, WDR26 and the 14-3-3 complex could be identified as specific components of the lebercilin protein

complex. Binding of ribosomal proteins is most likely due to the transient over expression of the bait protein lebercilin. This analysis shows that the established method for the detection of specific complex components can give valuable insights into a protein complex and thereby into a function of a protein. Even though, HEK293 cells might not be the perfect model for protein complexes in photoreceptors but as most mammalian cells and many cell lines, HEK293 cells form cilia. Additionally, the IFT complexes are highly conserved and the proteins are expressed in these cells. For these two reasons HEK293 cells are a suitable model system to investigate cilia associated mechanisms. Although there is a risk to miss specific complex components due to the background binding to the affinity resin or to photoreceptor specific expression of certain proteins, it is still possible to detect even low abundant protein complex components and proteins that are missed by a conventional TAP approach due to low stability, affinity or low abundance.

3.2.2   The lebercilin protein complex

The association of IFT proteins to the lebercilin complex demonstrates that lebercilin has a function in the IFT machinery. Most of the IFT proteins identified to be associated with the lebercilin protein complex are part of the IFT complex B. However, some of the proteins characterized as putative IFT proteins, have been shown to be important for the IFT but the sub-complex they are involved in was not clear[107,108,114,227]. The SF-TAP analysis of IFT88 shows that all of these putative IFTs are actually part of the IFT complex B in HEK293 cells. Additionally, it has not been clarified so far whether the IFT complexes, mainly described in *Caenorhabditis elegans* and *Chlamydomonas rheinhardtii*[28,105] are conserved in mammals, although a great degree of homology to human proteins does suggest a similar functional role. Parts of the complexes were also characterized in mammalian photoreceptors and kidney cells but this study for the first time presents a comprehensive analysis of the IFT complex B in mammalian cells.

The identification of the IFT proteins as part of the lebercilin complex is especially interesting because IFT is the mechanism by which cargo is transported from the inner to the outer segments in photoreceptors and the other way around[28,102]. The transport is very specific and only certain proteins are selected as cargo. The mechanism by which the correct cargo

is selected is only rudimentarily investigated yet and is still largely unknown. But not only IFT proteins were identified to interact with lebercilin. In addition to dynein light chain which was identified by the SF-TAP and the Strep-SILAC approach, KIF7, a homologue of Kinesin-2 heavy chain of *Chlamydomonas* and a homologue of *Drosophila* Costal-2, a repressor of hedgehog signaling was identified[228,229]. Surprisingly knock out of KIF7 in mice does not lead to any phenotype which implies that KIF7 is a motor protein and that the involvement in Hedgehog signaling is rather an indirect effect[228,229]. The identification of KIF7 as complex component of lebercilin opens the question whether lebercilin is involved only in retrograde, dynein-driven or the anterograde, kinesin driven IFT as well. It is also possible that the interaction of lebercilin with IFT complex A proteins is an indirect one, mediated by the IFT complex B. The same could be true for KIF7 since kinesins are transported as cargo in the retrograde transport. The exact structure of the complexes was not elucidated yet. Therefore it is not known how the motor proteins are linked to the IFT complexes when they are transported as cargo. It is possible that KIF7 is bound to the IFT complex B during retrograde dynein driven transport and that lebercilin is only involved in the retrograde transport.

3.2.3 The complex of the LCA associated lebercilin variant

The second step of the quantitative protein complex analysis of lebercilin was the comparison of WT and a mutant variant lebercilin. The hypothesis was that if one can conclude from the protein complex to the protein function, it should also be possible to associate eventual losses or gains in the binding pattern of a mutated protein to a loss or gain of its function. Because several mutations in lebercilin cause Leber congenital amaurosis and all of these mutations either lead to a truncated form of the protein or, in case of the deletion of the promoter region, to the absence of the protein, a loss of function has to be expected for all of the mutants. To identify alterations that possibly affect all mutant variants, I chose a mutant variant encoding for the smallest truncation and thereby to the longest form of all lebercilin mutants[200] for the complex comparison with the wild type complex.

The complex comparison clearly demonstrated that not all components loose their association with mutated lebercilin but only a certain sub-complex does not bind to the mutant variant of lebercilin. Surprisingly, all

members of the IFT complex B as well as the identified IFT complex A proteins loose their association with the complex of mutated lebercilin. Many of the other proteins do not show a significant alteration in binding. The two dynein light chain isoforms for example show no significant alteration in binding to the protein complex of mutated lebercilin as did all 14-3-3 isoforms. Unfortunately, KIF7 could not be identified in this comparison. It was only present in one of the three experiments. This is not enough to conclusively state whether its binding is altered or lost. The proteins, showing the same behavior as the IFT proteins, namely association with the lebercilin complex and loss of association with the complex of mutated lebercilin, could be associated with IFT complexes as well. These proteins were not identified in the IFT88-SF-TAP analysis but they could be part of the IFT complex A which is not further characterized, yet.

3.2.4 A model for the LCA disease mechanism

Not only the structure of the IFT complex is conserved but the function is as well, at least in zebra fish. Knock out of IFT proteins (IFT88, IFT52, IFT57) was shown to cause severe phenotypes, especially in the eye, very similar to the phenotype of Leber congenital amaurosis in humans[30,31,102,197]. When considering these findings, the results obtained in this study and the localization of lebercilin as well as the IFT proteins IFT20, IFT52, IFT57 and IFT88 that mainly localize to the basal body and the axoneme of the cilium[30,31,73], it is possible that lebercilin has a function as scaffold, linking the motor proteins to the IFT proteins. The exact role of lebercilin in the IFT is not clear yet, also because the specific functions the IFT sub-complexes exert are still not completely understood.

Taken together, the knowledge gained about lebercilin and its associated protein so far, suggest a role for lebercilin in the assembly or the stabilization of the IFT complexes. Mutation of lebercilin and thereby loss of the stabilizing effect could lead to impairment of the IFT complex stability and to decreased transport rates (Figure 26). Lebercilin is not only expressed in photoreceptors but also in several other tissues. The disease phenotype, however, is restricted to retinal photoreceptors. This suggests that lebercilin function can be compensated in all other tissues. The reason therefore could be that the connecting cilium is the only connection between the inner and outer segment of photoreceptors and all the

components, especially rhodopsin, have to be transported via the connecting cilium.
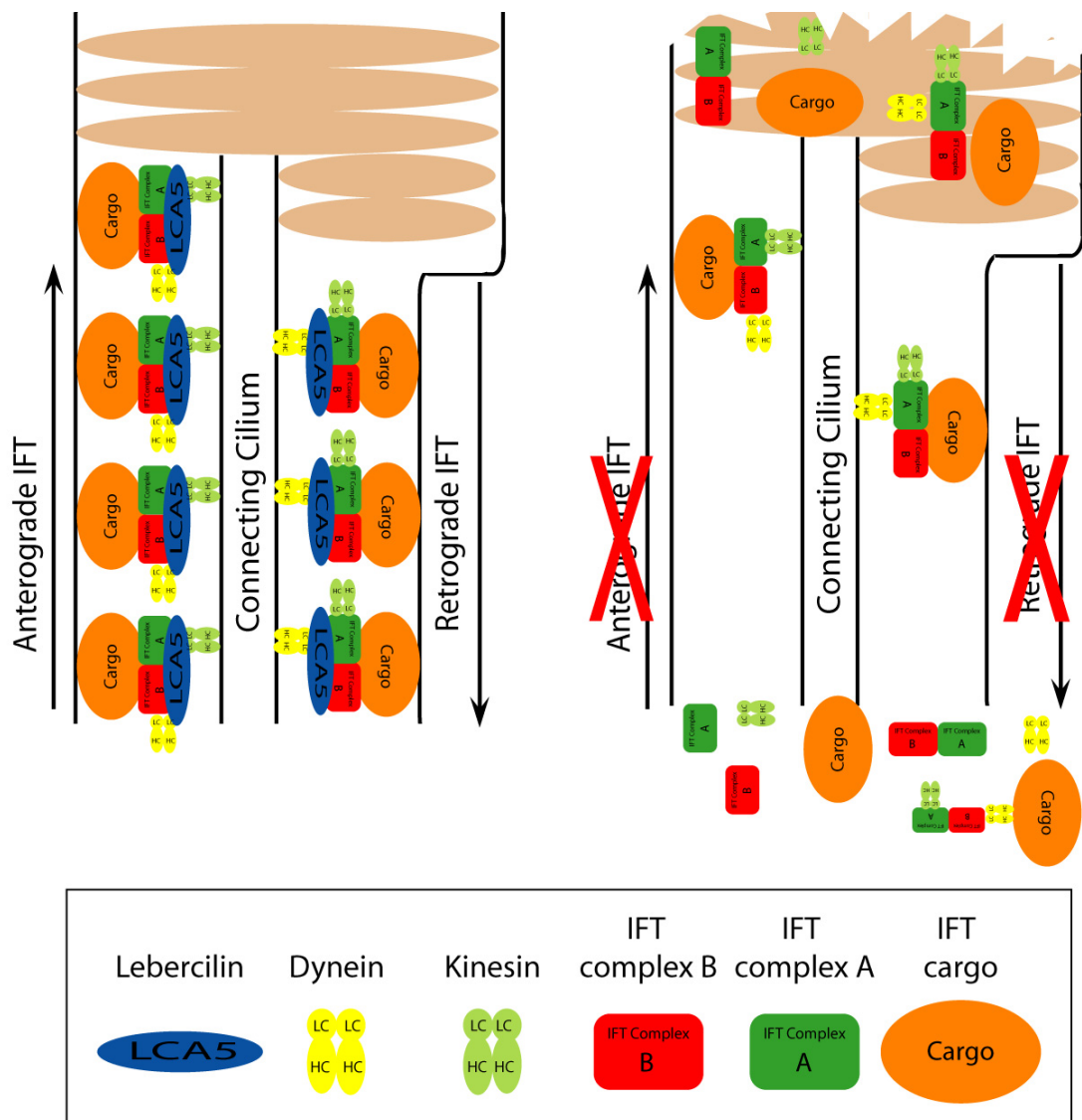
Figure 26: A model for lebercilin involvement in IFT. The left panel shows a model for IFT in normal photoreceptors. Lebercilin might act as a scaffold which helps assembling or stabilizing the IFT complexes. The right panel shows a model for the disease mechanism. Malfunction or absence of lebercilin of lebercilin leads to impairment of the assembly or stabilization of the IFT complexes. Impairment of the IFT leads to less transport of components of the outer segments and thereby finally to the degradation of the outer segments and blindness.

In contrast to other ciliated cells, the rate of transport is substantially higher in photoreceptors because there is a constant need for recycling and supplies of components due to disc shedding[124]. Therefore, the IFT in photoreceptors is subjected to a very high load. As such, dysfunction of a single component may not be compensated. This is also in-line with the disease phenotype, characterized by degeneration of the outer segments

of the photoreceptors after birth when rhodopsin is expressed and light reception initiated[138,143].

4. Perspectives

This dissertation presents the development, establishment and application of methods for protein complex analysis, including quantitative protein complex analysis. A new approach for tandem affinity purification, the SF-TAP method was applied to the analysis of the MAPK network and to the protein complex of lebercilin. The quantitative method based on single-step purification in combination with stable isotope labeling by amino acids in cell culture (SILAC) was applied to define the composition of the Lebercilin complex at a high confidence level and identify mutation specific alterations in protein interactions of a lebercilin mutant, associated with the hereditary eye disease LCA. SF-TAP tag based purifications in combination with SILAC prompted to be highly suitable tools for the analysis of protein complexes, protein complex comparison and protein network analysis. The analysis of the lebercilin complex and the characterization of the IFT complex B around the lebercilin complex component IFT88 showed that these methods can be applied for the comprehensive analysis of protein networks, by choosing different components of the complex as bait.

Future studies will focus on an in-depth analysis of the ciliary protein network associated with lebercilin, as already started with the analysis of the IFT complex B by SF-TAP analysis of IFT88. Candidate IFT proteins and so far uncharacterized proteins will therefore be analyzed by SF-TAP and/or Strep-SILAC. The sub-structure of the network will then be deciphered by combining affinity based purification of the complexes with methods to separate sub-complexes like blue-native gel electrophoresis or isopycnic gradients. The detailed architecture of the lebercilin associated complexes might give further insights into the IFT machineries functional modules associated with directed transport in retinal photoreceptors and eventually other ciliated cells.

Lebercilin was described in this study to be part of the intraflagellar transport machinery. However, the exact function within this machinery is not entirely clear yet. Besides IFT proteins, especially of complex B, lebercilin interacts with two motor proteins, dynein light chain and KIF7. This leaves open on whether lebercilin is involved in the anterograde or in the retrograde transport along cilia. Here, lebercilin could possibly function as a scaffold protein, necessary for assembling IFT sub-complexes, linking

them to specific cargo and/or to motor proteins. To elucidate this, functional assays and co-localization studies will be undertaken. In addition a knock out mouse generated by Ronald Roepman and co-workers will be analyzed with respect to the interplay between lebercilin and IFT. In cellular model systems, like HEK293 cells and ARPE or MDCK cells, knock down of various components of the IFT complex, including lebercilin and motor proteins, followed by the analysis of their impact on the protein complex assembly and the localization of the other components is planned to gain mechanistic insights into complex assembly and dynamics.

I References

1. Papaioannou, V.E., McBurney, M.W., Gardner, R.L. & Evans, M.J. Fate of teratocarcinoma cells injected into early mouse embryos. *Nature* **258**, 70-73 (1975).

2. Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. & Kucherlapati, R.S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230-4 (1985).

3. Thomas, K.R., Folger, K.R. & Capecchi, M.R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419-28 (1986).

4. Elbashir, S.M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-8 (2001).

5. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-11 (1998).

6. Lee, R., Feinbaum, R. & Ambros, V. A short history of a short RNA. *Cell* **116**, S89-92, 1 p following S96 (2004).

7. Lee, R.C., Feinbaum, R.L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-54 (1993).

8. Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **92**, 291-4 (1998).

9. Alberts, B.M. The DNA enzymology of protein machines. *Cold Spring Harb Symp Quant Biol* **49**, 1-12 (1984).

10. Kolch, W. Coordinating ERK/MAPK signalling through scaffolds and inhibitors. *Nat Rev Mol Cell Biol* **6**, 827-37 (2005).

11. O'Neill, E. & Kolch, W. Conferring specificity on the ubiquitous Raf/MEK signalling pathway. *Br J Cancer* **90**, 283-8 (2004).

12. O'Neill, E., Rushworth, L., Baccarini, M. & Kolch, W. Role of the kinase MST2 in suppression of apoptosis by the proto-oncogene product Raf-1. *Science* **306**, 2267-70 (2004).

13. Wu, X., Carr, H.S., Dan, I., Ruvolo, P.P. & Frost, J.A. p21 activated kinase 5 activates Raf-1 and targets it to mitochondria. *J Cell Biochem* **105**, 167-75 (2008).

14. Rocks, O., Peyker, A. & Bastiaens, P.I. Spatio-temporal segregation of Ras signals: one ship, three anchors, many harbors. *Curr Opin Cell Biol* **18**, 351-7 (2006).

15. Sasagawa, S., Ozaki, Y., Fujita, K. & Kuroda, S. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nat Cell Biol* **7**, 365-73 (2005).

16. von Kriegsheim, A., Pitt, A., Grindlay, G.J., Kolch, W. & Dhillon, A.S. Regulation of the Raf-MEK-ERK pathway by protein phosphatase 5. *Nat Cell Biol* **8**, 1011-6 (2006).

17. Niethammer, P. et al. Discrete states of a protein interaction network govern interphase and mitotic microtubule dynamics. *PLoS Biol* **5**, e29 (2007).

18. Santos, S.D., Verveer, P.J. & Bastiaens, P.I. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat Cell Biol* **9**, 324-30 (2007).

19. Deng, X., Gao, F. & May, W.S. Protein phosphatase 2A inactivates Bcl2's antiapoptotic function by dephosphorylation and up-regulation of Bcl2-p53 binding. *Blood* (2008).

20. Guillaud, L., Wong, R. & Hirokawa, N. Disruption of KIF17-Mint1 interaction by CaMKII-dependent phosphorylation: a molecular model of kinesin-cargo release. *Nat Cell Biol* **10**, 19-29 (2008).

21. Ke, B., Oh, E. & Thurmond, D.C. Doc2beta is a novel Munc18c-interacting partner and positive effector of syntaxin 4-mediated exocytosis. *J Biol Chem* **282**, 21786-97 (2007).

22. Miller, S.E., Collins, B.M., McCoy, A.J., Robinson, M.S. & Owen, D.J. A SNARE-adaptor interaction is a new mode of cargo recognition in clathrin-coated vesicles. *Nature* **450**, 570-4 (2007).

23. Yeung, K. et al. Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature* **401**, 173-7 (1999).

24. Nguyen, A. et al. Kinase suppressor of Ras (KSR) is a scaffold which facilitates mitogen-activated protein kinase activation in vivo. *Mol Cell Biol* **22**, 3035-45 (2002).

25. Therrien, M., Michaud, N.R., Rubin, G.M. & Morrison, D.K. KSR modulates signal propagation within the MAPK cascade. *Genes Dev* **10**, 2684-95 (1996).

26. Ishibe, S., Joly, D., Liu, Z.X. & Cantley, L.G. Paxillin serves as an ERK-regulated scaffold for coordinating FAK and Rac activation in epithelial morphogenesis. *Mol Cell* **16**, 257-67 (2004).

27. Ishibe, S., Joly, D., Zhu, X. & Cantley, L.G. Phosphorylation-dependent paxillin-ERK association mediates hepatocyte growth factor-stimulated epithelial morphogenesis. *Mol Cell* **12**, 1275-85 (2003).

28. Blacque, O.E., Cevik, S. & Kaplan, O.I. Intraflagellar transport: from molecular characterisation to mechanism. *Front Biosci* **13**, 2633-52 (2008).

29. Scholey, J.M. Intraflagellar transport. *Annu Rev Cell Dev Biol* **19**, 423-43 (2003).

30. Pazour, G.J. et al. The intraflagellar transport protein, IFT88, is essential for vertebrate photoreceptor assembly and maintenance. *J Cell Biol* **157**, 103-13 (2002).

31. Tsujikawa, M. & Malicki, J. Intraflagellar transport genes are essential for differentiation and survival of vertebrate sensory neurons. *Neuron* **42**, 703-16 (2004).

32. Fromont-Racine, M., Rain, J.C. & Legrain, P. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* **16**, 277-82 (1997).

33. Gavin, A.C. et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-7 (2002).

34. Rigaut, G. et al. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-2 (1999).

35. Angrand, P.O. et al. Transgenic mouse proteomics identifies new 14-3-3-associated proteins involved in cytoskeletal rearrangements and cell signaling. *Mol Cell Proteomics* **5**, 2211-27 (2006).

36. Benzinger, A., Muster, N., Koch, H.B., Yates, J.R., 3rd & Hermeking, H. Targeted proteomic analysis of 14-3-3 sigma, a p53 effector commonly silenced in cancer. *Mol Cell Proteomics* **4**, 785-95 (2005).

37. Bouwmeester, T. et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol* **6**, 97-105 (2004).

38. Blagoev, B. et al. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotechnol* **21**, 315-8 (2003).

39. Burckstummer, T. et al. An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat Methods* **3**, 1013-9 (2006).

40. Drakas, R., Prisco, M. & Baserga, R. A modified tandem affinity purification tag technique for the purification of protein complexes in mammalian cells. *Proteomics* **5**, 132-7 (2005).

41. Knuesel, M. et al. Identification of novel protein-protein interactions using a versatile mammalian tandem affinity purification expression system. *Mol Cell Proteomics* **2**, 1225-33 (2003).

42. Letteboer, S.J. & Roepman, R. Versatile screening for binary protein-protein interactions by yeast two-hybrid mating. *Methods Mol Biol* **484**, 145-59 (2008).

43. Simonis, N. et al. Empirically controlled mapping of the Caenorhabditis elegans protein-protein interactome network. *Nat Methods* **6**, 47-54 (2009).

44. Yu, H. et al. High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104-10 (2008).

45. Gorden, N.T. et al. CC2D2A is mutated in Joubert syndrome and interacts with the ciliopathy-associated basal body protein CEP290. *Am J Hum Genet* **83**, 559-71 (2008).

46. van Wijk, E. et al. Usher syndrome and Leber congenital amaurosis are molecularly linked via a novel isoform of the centrosomal ninein-like protein. *Hum Mol Genet* **18**, 51-64 (2009).

47. Suter, B., Kittanakom, S. & Stagljar, I. Two-hybrid technologies in proteomics research. *Curr Opin Biotechnol* **19**, 316-23 (2008).

48. Brizzard, B.L., Chubet, R.G. & Vizard, D.L. Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution. *Biotechniques* **16**, 730-5 (1994).

49. Junttila, M.R., Saarinen, S., Schmidt, T., Kast, J. & Westermarck, J. Single-step Strep-tag purification for the isolation and identification of protein complexes from mammalian cells. *Proteomics* **5**, 1199-203 (2005).

50. Kolodziej, P.A. & Young, R.A. Epitope tagging and protein surveillance. *Methods Enzymol* **194**, 508-19 (1991).

51. Lichty, J.J., Malecki, J.L., Agnew, H.D., Michelson-Horowitz, D.J. & Tan, S. Comparison of affinity tags for protein purification. *Protein Expr Purif* **41**, 98-105 (2005).

52. Schmidt, T.G., Koepke, J., Frank, R. & Skerra, A. Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *J Mol Biol* **255**, 753-66 (1996).

53. Schmidt, T.G. & Skerra, A. One-step affinity purification of bacterially produced proteins by means of the "Strep tag" and immobilized recombinant core streptavidin. *J Chromatogr A* **676**, 337-45 (1994).

54. Schmidt, T.G. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* **2**, 1528-35 (2007).

55. Rohila, J.S., Chen, M., Cerny, R. & Fromm, M.E. Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J* **38**, 172-81 (2004).

56. Rubio, V. et al. An alternative tandem affinity purification strategy applied to Arabidopsis protein complex isolation. *Plant J* **41**, 767-78 (2005).

57. Schimanski, B., Nguyen, T.N. & Gunzl, A. Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell* **4**, 1942-50 (2005).

58. Yang, P., Sampson, H.M. & Krause, H.M. A modified tandem affinity purification strategy identifies cofactors of the Drosophila nuclear receptor dHNF4. *Proteomics* **6**, 927-35 (2006).

59. Follit, J.A., Xu, F., Keady, B.T. & Pazour, G.J. Characterization of mouse IFT complex B. *Cell Motil Cytoskeleton* (2009).

60. Vermeulen, M., Hubner, N.C. & Mann, M. High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr Opin Biotechnol* **19**, 331-7 (2008).

61. Selbach, M. & Mann, M. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat Methods* **3**, 981-3 (2006).

62. Ong, S.E. et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-86 (2002).

63. Gygi, S.P. et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994-9 (1999).

64. Schmidt, A., Kellermann, J. & Lottspeich, F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**, 4-15 (2005).

65. Ross, P.L. et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **3**, 1154-69 (2004).

66. Wang, W. et al. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* **75**, 4818-26 (2003).

67. Trinkle-Mulcahy, L. et al. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol* **183**, 223-39 (2008).

68. Vermeulen, M. et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58-69 (2007).

69. DeRouen, M.C. & Oro, A.E. The primary cilium: a small yet mighty organelle. *J Invest Dermatol* **129**, 264-5 (2009).

70. Evans, J.E. et al. Functional modulation of IFT kinesins extends the sensory repertoire of ciliated neurons in Caenorhabditis elegans. *J Cell Biol* **172**, 663-9 (2006).

71. Greenwood, M. On Retractile Cilia in the Intestine of Lumbricus Terrestris. *J Physiol* **13**, 239-59 (1892).

72. Fliegauf, M., Benzing, T. & Omran, H. When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol* **8**, 880-93 (2007).

73. den Hollander, A.I. et al. Mutations in LCA5, encoding the ciliary protein lebercilin, cause Leber congenital amaurosis. *Nat Genet* **39**, 889-95 (2007).

74. Roepman, R. et al. Interaction of nephrocystin-4 and RPGRIP1 is disrupted by nephronophthisis or Leber congenital amaurosis-associated mutations. *Proc Natl Acad Sci U S A* **102**, 18520-5 (2005).

75. Arts, H.H. et al. Mutations in the gene encoding the basal body protein RPGRIP1L, a nephrocystin-4 interactor, cause Joubert syndrome. *Nat Genet* **39**, 882-8 (2007).

76. Ansley, S.J. et al. Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. *Nature* **425**, 628-33 (2003).

77. Pazour, G.J., San Agustin, J.T., Follit, J.A., Rosenbaum, J.L. & Witman, G.B. Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Curr Biol* **12**, R378-80 (2002).

78. Badano, J.L., Mitsuma, N., Beales, P.L. & Katsanis, N. The ciliopathies: an emerging class of human genetic disorders. *Annu Rev Genomics Hum Genet* **7**, 125-48 (2006).

79. Toremalm, N.G., Hakansson, C.H., Mercke, U., Dahlerus, B. & Huberman, D. Mucociliary wave pattern. A comparative analysis of extracellular and intracellular activities. *Acta Otolaryngol* **79**, 436-41 (1975).

80. Hirokawa, N., Tanaka, Y., Okada, Y. & Takeda, S. Nodal flow and the generation of left-right asymmetry. *Cell* **125**, 33-45 (2006).

81. Nonaka, S. et al. Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell* **95**, 829-37 (1998).

82. Papermaster, D.S., Schneider, B.G., DeFoe, D. & Besharse, J.C. Biosynthesis and vectorial transport of opsin on vesicles in retinal rod photoreceptors. *J Histochem Cytochem* **34**, 5-16 (1986).

83. Wolfrum, U. & Schmitt, A. Rhodopsin transport in the membrane of the connecting cilium of mammalian photoreceptor cells. *Cell Motil Cytoskeleton* **46**, 95-107 (2000).

84. Quarmby, L.M. & Parker, J.D. Cilia and the cell cycle? *J Cell Biol* **169**, 707-10 (2005).

85. Santos, N. & Reiter, J.F. Building it up and taking it down: the regulation of vertebrate ciliogenesis. *Dev Dyn* **237**, 1972-81 (2008).

86. Plotnikova, O.V., Golemis, E.A. & Pugacheva, E.N. Cell cycle-dependent ciliogenesis and cancer. *Cancer Res* **68**, 2058-61 (2008).

87. Spektor, A., Tsang, W.Y., Khoo, D. & Dynlacht, B.D. Cep97 and CP110 suppress a cilia assembly program. *Cell* **130**, 678-90 (2007).

88. Dawe, H.R., Farr, H. & Gull, K. Centriole/basal body morphogenesis and migration during ciliogenesis in animal cells. *J Cell Sci* **120**, 7-15 (2007).

89. Lemullois, M., Boisvieux-Ulrich, E., Laine, M.C., Chailley, B. & Sandoz, D. Development and functions of the cytoskeleton during ciliogenesis in metazoa. *Biol Cell* **63**, 195-208 (1988).

90. Fowkes, M.E. & Mitchell, D.R. The role of preassembled cytoplasmic complexes in assembly of flagellar dynein subunits. *Mol Biol Cell* **9**, 2337-47 (1998).

91. Avidor-Reiss, T. et al. Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. *Cell* **117**, 527-39 (2004).

92. Rosenbaum, J.L. & Witman, G.B. Intraflagellar transport. *Nat Rev Mol Cell Biol* **3**, 813-25 (2002).

93. Pugacheva, E.N. & Golemis, E.A. The focal adhesion scaffolding protein HEF1 regulates activation of the Aurora-A and Nek2 kinases at the centrosome. *Nat Cell Biol* **7**, 937-46 (2005).

94. Pugacheva, E.N., Jablonski, S.A., Hartman, T.R., Henske, E.P. & Golemis, E.A. HEF1-dependent Aurora A activation induces disassembly of the primary cilium. *Cell* **129**, 1351-63 (2007).

95. Salathe, M. Regulation of mammalian ciliary beating. *Annu Rev Physiol* **69**, 401-22 (2007).

96. Mesland, D.A., Hoffman, J.L., Caligor, E. & Goodenough, U.W. Flagellar tip activation stimulated by membrane adhesions in Chlamydomonas gametes. *J Cell Biol* **84**, 599-617 (1980).

97. Perkins, L.A., Hedgecock, E.M., Thomson, J.N. & Culotti, J.G. Mutant sensory cilia in the nematode Caenorhabditis elegans. *Dev Biol* **117**, 456-87 (1986).

98. Snow, J.J. et al. Two anterograde intraflagellar transport motors cooperate to build sensory cilia on C. elegans neurons. *Nat Cell Biol* **6**, 1109-13 (2004).

99. Feistel, K. & Blum, M. Three types of cilia including a novel 9+4 axoneme on the notochordal plate of the rabbit embryo. *Dev Dyn* **235**, 3348-58 (2006).

100. O'Toole, E.T., Giddings, T.H., McIntosh, J.R. & Dutcher, S.K. Three-dimensional organization of basal bodies from wild-type and delta-tubulin deletion strains of Chlamydomonas reinhardtii. *Mol Biol Cell* **14**, 2999-3012 (2003).

101. Dutcher, S.K. Elucidation of basal body and centriole functions in Chlamydomonas reinhardtii. *Traffic* **4**, 443-51 (2003).

102. Insinna, C. & Besharse, J.C. Intraflagellar transport and the sensory outer segment of vertebrate photoreceptors. *Dev Dyn* **237**, 1982-92 (2008).

103. Deane, J.A., Cole, D.G., Seeley, E.S., Diener, D.R. & Rosenbaum, J.L. Localization of intraflagellar transport protein IFT52 identifies basal body transitional fibers as the docking site for IFT particles. *Curr Biol* **11**, 1586-90 (2001).

104. Hao, L. & Scholey, J.M. Intraflagellar transport at a glance. *J Cell Sci* **122**, 889-92 (2009).

105. Ou, G. et al. Sensory ciliogenesis in Caenorhabditis elegans: assignment of IFT components into distinct modules based on transport and phenotypic profiles. *Mol Biol Cell* **18**, 1554-69 (2007).

106. Baker, S.A., Freeman, K., Luby-Phelps, K., Pazour, G.J. & Besharse, J.C. IFT20 links kinesin II with a mammalian intraflagellar transport complex that is conserved in motile flagella and sensory cilia. *J Biol Chem* **278**, 34211-8 (2003).

107. Efimenko, E. et al. Caenorhabditis elegans DYF-2, an orthologue of human WDR19, is a component of the intraflagellar transport machinery in sensory cilia. *Mol Biol Cell* **17**, 4801-11 (2006).

108. Tran, P.V. et al. THM1 negatively modulates mouse sonic hedgehog signal transduction and affects retrograde intraflagellar transport in cilia. *Nat Genet* **40**, 403-10 (2008).

109. Kunitomo, H. & Iino, Y. Caenorhabditis elegans DYF-11, an orthologue of mammalian Traf3ip1/MIP-T3, is required for sensory cilia formation. *Genes Cells* **13**, 13-25 (2008).

110. Follit, J.A., Tuft, R.A., Fogarty, K.E. & Pazour, G.J. The intraflagellar transport protein IFT20 is associated with the Golgi complex and is required for cilia assembly. *Mol Biol Cell* **17**, 3781-92 (2006).

111. Shu, X. et al. RPGR ORF15 isoform co-localizes with RPGRIP1 at centrioles and basal bodies and interacts with nucleophosmin. *Hum Mol Genet* **14**, 1183-97 (2005).

112. Khanna, H. et al. RPGR-ORF15, which is mutated in retinitis pigmentosa, associates with SMC1, SMC3, and microtubule transport proteins. *J Biol Chem* **280**, 33580-7 (2005).

113. Hong, D.H. et al. RPGR isoforms in photoreceptor connecting cilia and the transitional zone of motile cilia. *Invest Ophthalmol Vis Sci* **44**, 2413-21 (2003).

114. Adhiambo, C., Blisnick, T., Toutirais, G., Delannoy, E. & Bastin, P. A novel function for the atypical small G protein Rab-like 5 in the assembly of the trypanosome flagellum. *J Cell Sci* **122**, 834-41 (2009).

115. McGrath, J., Somlo, S., Makova, S., Tian, X. & Brueckner, M. Two populations of node monocilia initiate left-right asymmetry in the mouse. *Cell* **114**, 61-73 (2003).

116. Tanaka, Y., Okada, Y. & Hirokawa, N. FGF-induced vesicular release of Sonic hedgehog and retinoic acid in leftward nodal flow is critical for left-right determination. *Nature* **435**, 172-7 (2005).

117. Chemes, H.E. et al. Ultrastructural pathology of the sperm flagellum: association between flagellar pathology and fertility prognosis in severely asthenozoospermic men. *Hum Reprod* **13**, 2521-6 (1998).

118. Gallo, J.M., Escalier, D., Schrevel, J. & David, G. Differential distribution of tubulin epitopes in human spermatozoa. *Eur J Cell Biol* **40**, 111-6 (1986).

119. Brunner, S. et al. Overexpression of RPGR leads to male infertility in mice due to defects in flagellar assembly. *Biol Reprod* **79**, 608-17 (2008).

120. Escalier, D. Knockout mouse models of sperm flagellum anomalies. *Hum Reprod Update* **12**, 449-61 (2006).

121. Rodieck, R.W. *The vertebrate retina; principles of structure and function*, (Freeman, San Francisco, 1973).

122. Reiners, J. & Wolfrum, U. Molecular analysis of the supramolecular usher protein complex in the retina. Harmonin as the key protein of the Usher syndrome. *Adv Exp Med Biol* **572**, 349-53 (2006).

123. Steinberg, R.H., Fisher, S.K. & Anderson, D.H. Disc morphogenesis in vertebrate photoreceptors. *J Comp Neurol* **190**, 501-8 (1980).

124. Anderson, D.H., Fisher, S.K. & Steinberg, R.H. Mammalian cones: disc shedding, phagocytosis, and renewal. *Invest Ophthalmol Vis Sci* **17**, 117-33 (1978).

125. Veland, I.R., Awan, A., Pedersen, L.B., Yoder, B.K. & Christensen, S.T. Primary cilia and signaling pathways in mammalian development, health and disease. *Nephron Physiol* **111**, p39-53 (2009).

126. Afzelius, B.A. A human syndrome caused by immotile cilia. *Science* **193**, 317-9 (1976).

127. Pazour, G.J. et al. Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. *J Cell Biol* **151**, 709-18 (2000).

128. Murcia, N.S. et al. The Oak Ridge Polycystic Kidney (orpk) disease gene is required for left-right axis determination. *Development* **127**, 2347-55 (2000).

129. Yoder, B.K., Hou, X. & Guay-Woodford, L.M. The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *J Am Soc Nephrol* **13**, 2508-16 (2002).

130. Liu, A., Wang, B. & Niswander, L.A. Mouse intraflagellar transport proteins regulate both the activator and repressor functions of Gli transcription factors. *Development* **132**, 3103-11 (2005).

131. Haycraft, C.J. et al. Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genet* **1**, e53 (2005).

132. Torban, E., Kor, C. & Gros, P. Van Gogh-like2 (Strabismus) and its role in planar cell polarity and convergent extension in vertebrates. *Trends Genet* **20**, 570-7 (2004).

133. Liu, X. et al. Usherin is required for maintenance of retinal photoreceptors and normal development of cochlear hair cells. *Proc Natl Acad Sci U S A* **104**, 4413-8 (2007).

134. Yan, D. et al. An isoform of GTPase regulator DOCK4 localizes to the stereocilia in the inner ear and binds to harmonin (USH1C). *J Mol Biol* **357**, 755-64 (2006).

135. Adato, A. et al. Usherin, the defective protein in Usher syndrome type IIA, is likely to be a component of interstereocilia ankle links in the inner ear sensory cells. *Hum Mol Genet* **14**, 3921-32 (2005).

136. Tosi, G.M., de Santi, M.M., Pradal, U., Braggion, C. & Luzi, P. Clinicopathologic reports, case reports, and small case series: usher syndrome type 1 associated with primary ciliary aplasia. *Arch Ophthalmol* **121**, 407-8 (2003).

137. Leber, T. Über Retinitis pigmentosa und angeborene Amaurose. *Archiv für Ophthalmologie* (1869).

138. den Hollander, A.I., Roepman, R., Koenekoop, R.K. & Cremers, F.P. Leber congenital amaurosis: genes, proteins and disease mechanisms. *Prog Retin Eye Res* **27**, 391-419 (2008).

139. McEwen, D.P. et al. Hypomorphic CEP290/NPHP6 mutations result in anosmia caused by the selective loss of G proteins in cilia of olfactory sensory neurons. *Proc Natl Acad Sci U S A* **104**, 15917-22 (2007).

140. Koenekoop, R.K. RPGRIP1 is mutated in Leber congenital amaurosis: a mini-review. *Ophthalmic Genet* **26**, 175-9 (2005).

141. Zhao, Y. et al. The retinitis pigmentosa GTPase regulator (RPGR)-interacting protein: subserving RPGR function and participating in disk morphogenesis. *Proc Natl Acad Sci U S A* **100**, 3965-70 (2003).

142. Dharmaraj, S. et al. A novel locus for Leber congenital amaurosis maps to chromosome 6q. *Am J Hum Genet* **66**, 319-26 (2000).

143. Mohamed, M.D. et al. Progression of phenotype in Leber's congenital amaurosis with a mutation at the LCA5 locus. *Br J Ophthalmol* **87**, 473-5 (2003).

144. Gherman, A., Davis, E.E. & Katsanis, N. The ciliary proteome database: an integrated community resource for the genetic and functional dissection of cilia. *Nat Genet* **38**, 961-2 (2006).

145. Ostrowski, L.E. et al. A proteomic analysis of human cilia: identification of novel components. *Mol Cell Proteomics* **1**, 451-65 (2002).

146. Reed, S.E., Staley, E.M., Mayginnes, J.P., Pintel, D.J. & Tullis, G.E. Transfection of mammalian cells using linear polyethylenimine is a simple and effective means of producing recombinant adeno-associated virus vectors. *J Virol Methods* **138**, 85-98 (2006).

147. Baker, A. et al. Polyethylenimine (PEI) is a simple, inexpensive and effective reagent for condensing and linking plasmid DNA to adenovirus for gene delivery. *Gene Ther* **4**, 773-82 (1997).

148. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54 (1976).

149. Smith, P.K. et al. Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**, 76-85 (1985).

150. Peterson, G.L. Determination of total protein. *Methods Enzymol* **91**, 95-119 (1983).

151. Splittgerber, A.G. & Sohl, J. Nonlinearity in protein assays by the Coomassie blue dye-binding method. *Anal Biochem* **179**, 198-201 (1989).

152. Wessel, D. & Flugge, U.I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* **138**, 141-3 (1984).

153. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5 (1970).

154. Candiano, G. et al. Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25**, 1327-33 (2004).

155. Neuhoff, V., Arold, N., Taube, D. & Ehrhardt, W. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **9**, 255-62 (1988).

156. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* **68**, 850-8 (1996).

157. Switzer, R.C., 3rd, Merril, C.R. & Shifrin, S. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal Biochem* **98**, 231-7 (1979).

158. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4 (1979).

159. Schmidt, T.G. & Skerra, A. The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment. *Protein Eng* **6**, 109-22 (1993).

160. Voss, S. & Skerra, A. Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. *Protein Eng* **10**, 975-82 (1997).

161. Hopp, T.P. et al. A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. *Nat Biotech* **6**, 1204-1210 (1988).

162. Gloeckner, C.J., Boldt, K., Schumacher, A., Roepman, R. & Ueffing, M. A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics* **7**, 4228-34 (2007).

163. Agell, N., Bachs, O., Rocamora, N. & Villalonga, P. Modulation of the Ras/Raf/MEK/ERK pathway by Ca(2+), and calmodulin. *Cell Signal* **14**, 649-54 (2002).

164. Villalonga, P. et al. Calmodulin binds to K-Ras, but not to H- or N-Ras, and modulates its downstream signaling. *Mol Cell Biol* **21**, 7345-54 (2001).

165. Gloeckner, C.J., Boldt, K., Schumacher, A. & Ueffing, M. Tandem Affinity Purification of Protein Complexes from Mammalian Cells by the Strep/FLAG (SF)-TAP Tag. *Methods Mol Biol* **564**, 359-72 (2009).

166. Perkins, D.N., Pappin, D.J., Creasy, D.M. & Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-67 (1999).

167. Hu, Q. et al. The Orbitrap: a new mass spectrometer. *J Mass Spectrom* **40**, 430-43 (2005).

168. Kebarle, P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J Mass Spectrom* **35**, 804-17 (2000).

169. Makarov, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* **72**, 1156-62 (2000).

170. Makarov, A. et al. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem* **78**, 2113-20 (2006).

171. Makarov, A., Denisov, E., Lange, O. & Horning, S. Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *J Am Soc Mass Spectrom* **17**, 977-82 (2006).

172. Olsen, J.V. et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* **4**, 2010-21 (2005).

173. Yates, J.R., 3rd, Eng, J.K., McCormack, A.L. & Schieltz, D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* **67**, 1426-36 (1995).

174. Carr, S. et al. The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Mol Cell Proteomics* **3**, 531-3 (2004).

134

175. Keller, A., Nesvizhskii, A.I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383-92 (2002).

176. Nesvizhskii, A.I. & Aebersold, R. Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS. *Drug Discov Today* **9**, 173-81 (2004).

177. Li, J., Steen, H. & Gygi, S.P. Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. *Mol Cell Proteomics* **2**, 1198-204 (2003).

178. Kruger, M. et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* **134**, 353-64 (2008).

179. Schulze, W.X. & Mann, M. A novel proteomic screen for peptide-protein interactions. *J Biol Chem* **279**, 10756-64 (2004).

180. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-72 (2008).

181. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* **389**, 1017-31 (2007).

182. Rushworth, L.K., Hindley, A.D., O'Neill, E. & Kolch, W. Regulation and role of Raf-1/B-Raf heterodimerization. *Mol Cell Biol* **26**, 2262-72 (2006).

183. Fujioka, A. et al. Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *J Biol Chem* **281**, 8917-26 (2006).

184. Chen, C.Y., Olayioye, M.A., Lindeman, G.J. & Tang, T.K. CPAP interacts with 14-3-3 in a cell cycle-dependent manner. *Biochem Biophys Res Commun* **342**, 1203-10 (2006).

185. Mhawech, P. 14-3-3 proteins--an update. *Cell Res* **15**, 228-36 (2005).

186. Pietromonaco, S.F., Seluja, G.A., Aitken, A. & Elias, L. Association of 14-3-3 proteins with centrosomes. *Blood Cells Mol Dis* **22**, 225-37 (1996).

187. Shinmura, K., Tarapore, P., Tokuyama, Y., George, K.R. & Fukasawa, K. Characterization of centrosomal association of nucleophosmin/B23 linked to Crm1 activity. *FEBS Lett* **579**, 6621-34 (2005).

188. Li, D., Dobrowolska, G. & Krebs, E.G. Identification of proteins that associate with protein kinase CK2. *Mol Cell Biochem* **191**, 223-8 (1999).

189. Li, D., Dobrowolska, G. & Krebs, E.G. The physical association of casein kinase 2 with nucleolin. *J Biol Chem* **271**, 15662-8 (1996).

190. Hollander, B.A., Liang, M.Y. & Besharse, J.C. Linkage of a nucleolin-related protein and casein kinase II with the detergent-stable photoreceptor cytoskeleton. *Cell Motil Cytoskeleton* **43**, 114-27 (1999).

191. Muresan, V., Bendala-Tufanisco, E., Hollander, B.A. & Besharse, J.C. Evidence for kinesin-related proteins associated with the axoneme of retinal photoreceptors. *Exp Eye Res* **64**, 895-903 (1997).

192. Faust, M., Gunther, J., Morgenstern, E., Montenarh, M. & Gotz, C. Specific localization of the catalytic subunits of protein kinase CK2 at the centrosomes. *Cell Mol Life Sci* **59**, 2155-64 (2002).

193. Schermer, B. et al. Phosphorylation by casein kinase 2 induces PACS-1 binding of nephrocystin and targeting to cilia. *EMBO J* **24**, 4415-24 (2005).

194. Hu, J., Bae, Y.K., Knobel, K.M. & Barr, M.M. Casein kinase II and calcineurin modulate TRPP function and ciliary localization. *Mol Biol Cell* **17**, 2200-11 (2006).

195. Cox, J. et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* **4**, 698-705 (2009).

196. Piperno, G. & Mead, K. Transport of a novel complex in the cytoplasmic matrix of Chlamydomonas flagella. *Proc Natl Acad Sci U S A* **94**, 4457-62 (1997).

197. Luby-Phelps, K., Fogerty, J., Baker, S.A., Pazour, G.J. & Besharse, J.C. Spatial distribution of intraflagellar transport proteins in vertebrate photoreceptors. *Vision Res* **48**, 413-23 (2008).

198. Ishii, S. Costal-2: a scaffold for kinases mediates Hedgehog signaling. *Dev Cell* **8**, 140-1 (2005).

199. Errico, A., Claudiani, P., D'Addio, M. & Rugarli, E.I. Spastin interacts with the centrosomal protein NA14, and is enriched in the spindle pole, the midbody and the distal axon. *Hum Mol Genet* **13**, 2121-32 (2004).

200. Gerber, S. et al. Mutations in LCA5 are an uncommon cause of Leber congenital amaurosis (LCA) type II. *Hum Mutat* **28**, 1245 (2007).

201. Cheeseman, I.M. & Desai, A. A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci STKE* **2005**, pl1 (2005).

202. Konig, S., Guibert, B., Morice, C., Vernier, P. & Barnier, J.V. Phosphorylation by PKA of a site unique to B-Raf kinase. *C R Acad Sci III* **324**, 673-81 (2001).

203. Gerits, N., Kostenko, S., Shiryaev, A., Johannessen, M. & Moens, U. Relations between the mitogen-activated protein kinase and the cAMP-dependent protein kinase pathways: comradeship and hostility. *Cell Signal* **20**, 1592-607 (2008).

204. Calipel, A. et al. Extracellular signal-regulated kinase-dependent proliferation is mediated through the protein kinase A/B-Raf pathway in human uveal melanoma cells. *J Biol Chem* **281**, 9238-50 (2006).

205. Lin, S. et al. High glucose stimulates synthesis of fibronectin via a novel protein kinase C, Rap1b, and B-Raf signaling pathway. *J Biol Chem* **277**, 41725-35 (2002).

206. Jarchow, S., Luck, C., Gorg, A. & Skerra, A. Identification of potential substrate proteins for the periplasmic Escherichia coli chaperone Skp. *Proteomics* **8**, 4987-94 (2008).

207. Groth, A. et al. Regulation of replication fork progression through histone supply and demand. *Science* **318**, 1928-31 (2007).

208. Johansen, L.D. et al. IKAP localizes to membrane ruffles with filamin A and regulates actin cytoskeleton organization and cell migration. *J Cell Sci* **121**, 854-64 (2008).

209. Schaffitzel, C. & Ban, N. Generation of ribosome nascent chain complexes for structural and functional studies. *J Struct Biol* **158**, 463-71 (2007).

210. Morita, E. et al. Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO J* **26**, 4215-27 (2007).

211. Morita, E. et al. Identification of human MVB12 proteins as ESCRT-I subunits that function in HIV budding. *Cell Host Microbe* **2**, 41-53 (2007).

136

212. Ostermeier, C., Harrenga, A., Ermler, U. & Michel, H. Structure at 2.7 A resolution of the Paracoccus denitrificans two-subunit cytochrome c oxidase complexed with an antibody FV fragment. *Proc Natl Acad Sci U S A* **94**, 10547-53 (1997).

213. Herzberg, C. et al. SPINE: a method for the rapid detection and analysis of protein-protein interactions in vivo. *Proteomics* **7**, 4032-5 (2007).

214. Grbovic, O.M. et al. V600E B-Raf requires the Hsp90 chaperone for stability and is degraded in response to Hsp90 inhibitors. *Proc Natl Acad Sci U S A* **103**, 57-62 (2006).

215. da Rocha Dias, S. et al. Activated B-RAF is an Hsp90 client protein that is targeted by the anticancer drug 17-allylamino-17-demethoxygeldanamycin. *Cancer Res* **65**, 10686-91 (2005).

216. Dougherty, M.K. & Morrison, D.K. Unlocking the code of 14-3-3. *J Cell Sci* **117**, 1875-84 (2004).

217. Mackintosh, C. Dynamic interactions between 14-3-3 proteins and phosphoproteins regulate diverse cellular processes. *Biochem J* **381**, 329-42 (2004).

218. Satoh, J., Nanri, Y. & Yamamura, T. Rapid identification of 14-3-3-binding proteins by protein microarray analysis. *J Neurosci Methods* **152**, 278-88 (2006).

219. Jin, J. et al. Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr Biol* **14**, 1436-50 (2004).

220. Liu, M.Y., Cai, S., Espejo, A., Bedford, M.T. & Walker, C.L. 14-3-3 interacts with the tumor suppressor tuberin at Akt phosphorylation site(s). *Cancer Res* **62**, 6475-80 (2002).

221. Pozuelo Rubio, M. et al. 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. *Biochem J* **379**, 395-408 (2004).

222. Ma, Z. et al. Interaction between ROCK II and nucleophosmin/B23 in the regulation of centrosome duplication. *Mol Cell Biol* **26**, 9016-34 (2006).

223. Zhang, H. et al. B23/nucleophosmin serine 4 phosphorylation mediates mitotic functions of polo-like kinase 1. *J Biol Chem* **279**, 35726-34 (2004).

224. Ginisty, H., Sicard, H., Roger, B. & Bouvet, P. Structure and functions of nucleolin. *J Cell Sci* **112 (Pt 6)**, 761-72 (1999).

225. Hildesheim, J., Salvador, J.M., Hollander, M.C. & Fornace, A.J., Jr. Casein kinase 2- and protein kinase A-regulated adenomatous polyposis coli and beta-catenin cellular localization is dependent on p38 MAPK. *J Biol Chem* **280**, 17221-6 (2005).

226. Lim, A.C., Tiu, S.Y., Li, Q. & Qi, R.Z. Direct regulation of microtubule dynamics by protein kinase CK2. *J Biol Chem* **279**, 4433-9 (2004).

227. Lechtreck, K.F., Luro, S., Awata, J. & Witman, G.B. HA-tagging of putative flagellar proteins in Chlamydomonas reinhardtii identifies a novel protein of intraflagellar transport complex B. *Cell Motil Cytoskeleton* (2009).

228. Katoh, Y. & Katoh, M. Hedgehog signaling pathway and gastrointestinal stem cell signaling network (review). *Int J Mol Med* **18**, 1019-23 (2006).

229. Katoh, Y. & Katoh, M. Characterization of KIF7 gene in silico. *Int J Oncol* **25**, 1881-6 (2004).

Annex

1. Figure Index

2. Table index

3. Publications and presentations

3.1 Peer Reviewed Publications

Boldt K*, Mans DA*, Won J*, van Reeuwijk J, Vogt A, Kinkl N, Letteboer SJF, Hicks WL, Hurd RE, Naggert JK, Texier Y, Pazour GJ, den Hollander AI, Koenekoop RK, Bennett J, Cremers FP, Gloeckner CJ, Nishina PM, Roepman R#, Marius Ueffing#. Disruption of the intraflagellar transport machinery in photoreceptors causes Leber congenital amaurosis. In preparation.

Meixner A, Boldt K, Van Troys M, Askenazi M, Gloeckner CJ, Bauer M, Marto JA, Ampe C, Kinkl N, Ueffing M. A QUICK screen for Lrrk2 interaction partners – leucine-rich repeat kinase 2 is involved in actin cytoskeleton dynamics. In preparation

Piccoli G, Condliffe SB, Bauer M, Giesert F, Boldt K, Meixner A, Sarioglu H, Vogt-Weisenhorn DM, Wurst W, Gloeckner JG, Matteoli M, Sala C, Ueffing M. LRRK2 controls synaptic vesicle storage and mobilization within the recycling pool. Submitted

Bengel D*, Boldt K*, Davis EE, Burtscher I, Trümbach D, Diplas B, Attié-Bitach T, Wurst W, Katsanis N, Ueffing M, Lickert H. Pitchfork regulates cilia disassembly and left-right asymmetry. In press (Developmental Cell)

Gloeckner CJ, Boldt K, von Zweydorf F, Helm S, Wiesent L, Sarioglu H, Ueffing M. Phosphopeptide analysis reveals two discrete clusters of phosphorylation in the N-terminus and the Roc domain of the Parkinson-disease associated protein kinase LRRK2. *Journal of Proteome Research.* 2010 Apr 5;9(4):1738-45.

Gloeckner CJ, Schumacher A, Boldt K, Ueffing M, The Parkinson disease-associated protein kinase LRRK2 exhibits MAPKKK activity and phosphorylates MKK3/6 and MKK4/7, in vitro. *Journal of Neurochemistry*. 2009 May;109(4):959-68.

Hauck SM, Gloeckner CJ, Harley ME, Schoeffmann S, Boldt K, Ekstrom PA, Ueffing M. Identification of paracrine neuroprotective candidate proteins by a functional assay-driven proteomics approach. *Molecular and Cellular Proteomics*. 2008 Jul; 7(7):1349-61.

Gloeckner CJ, Boldt K, Schumacher A, Roepman R, Ueffing M. A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics*. 2007; 7(23):4228-34.

den Hollander AI*, Koenekoop RK*, Mohamed MD*, Arts HH*, Boldt K, Towns KV, Sedmak T, Beer M, Nagel-Wolfrum K, McKibbin M, Dharmaraj S, Lopez I, Ivings L, Williams GA, Springell K, Woods CG, Jafri H, Rashid Y, Strom TM, van der Zwaag B, Gosens I, Kersten FF, van Wijk E, Veltman JA, Zonneveld MN, van Beersum SE, Maumenee IH, Wolfrum U, Cheetham ME, Ueffing M, Cremers FP*, Inglehearn CF*, Roepman R*. Mutations in

LCA5, encoding the ciliary protein lebercilin, cause Leber congenital amaurosis. *Nature Genetetics*. 2007; 39(7):889-95.

<u>Boldt K</u>, Rist W, Weiss SM, Weith A, Lenter MC. FPRL-1 induces modifications of migration-associated proteins in human neutrophils. *Proteomics*. 2006; 6(17):4790-9.

* these authors contributed equally

3.2 Book Chapters

<u>Boldt K</u>, van Reeuwijk J, Gloeckner CJ, Ueffing M, Roepman R. Tandem Affinity Purification of Ciliopathy-Associated Protein Complexes. *Methods in Cell Biology*. 2009; p. 143.

Gloeckner CJ, Boldt K, Ueffing M. Strep/FLAG Tandem Affinity Purification (SF-TAP) to Study Protein Interactions. *Current Protocols in Protein Science*. 2009; 19.20.1-19.20.19.

Gloeckner CJ*, <u>Boldt K*</u>, Schumacher A, and Ueffing M. Tandem affinity purification of protein complexes from mammalian cells by the Strep/FLAG (SF)-TAP tag. *Methods in Molecular Biology: Proteomics Methods and Protocols*. 2009; Vol. 564.

* these authors contributed equally

3.3 Oral presentations

Medical Research Council, Human Genetics Unit; Edinburgh, UK: Analysis of protein complexes by tandem affinity purification, September 2006.

Interaction Proteome, annual meeting; Palma de Mallorca: Analysis of signalling complexes with a new tandem affinity approach, September 2006.

Beatson Institute for Cancer Research, Glasgow, UK: Analysis of disease associated B-Raf mutations, August 2007.

Applied Biosystems TOF/TOF user meeting, Darmstadt, Germany: Analyse von Proteinkomplexen mittels SF-TAP und LC-MALDI, January 2008.

Interaction proteome, annual meeting; Palma de Mallorca, Spain: Lebercilin: the discovery of the LCA5 gene - from gene identification towards analysis of functional protein networks, September 2009.

Arbeitstagung, Mikromethoden in der Proteinchemie, Martinsried, Germany: Quantitative Analyse von Krankheitsassoziierten Protein-komplexen, Juni 2009.

Molecular Interactions Workshop, Berlin, Germany: Lebercilin: from protein complexes to disease mechanism, September 2009.

Thermo Scientific User meeting, Martinsried, Germany: Lebercilin: from protein complexes to disease mechanism, October 2009.

Nijmegen, The Netherlands: Lebercilin: from protein complexes to disease mechanism, October 2009.

3.4 Posters presentations

Interaction Proteome, Summer School, Spetses, Greece: Protein complex analysis of Lebercilin, a protein associated with Leber congenital amaurosis, September 2008.

Proteomic Forum, Berlin, Germany: Quantitative protein complex analysis of Lebercilin, a protein associated with Leber congenital amaurosis, March 2009.

4. Acknowledgements

Bei Prof. Walter Kölch möchte ich mich dafür bedanken, dass er es mir zwei längere Aufenthalte in seiner Gruppe in Glasgow ermöglichte. Außerdem war er immer zur Diskussion und zu konstruktiver Kritik bereit und gab sehr hilfreiche Anregungen.

Bei Dr. Alex von Kriegsheim möchte ich mich für seine Hilfe bei der Durchführung der Experimente in Glasgow sowie seine weitere Unterstützung und für viele anregende Diskussionen bedanken. Er ermöglichte mir die Etablierung der quantitativen Proteinkomplexanalyse.

Bei Dr. Hakan Sarioglu möchte ich mich für seine Hilfe bei den massenspektrometrischen Arbeiten bedanken. Er war stets hilfsbereit und ermöglichte mir die Aus- und Weiterbildung, sowie die Benutzung der Massenspektrometer. Außerdem hielt er mir immer den Rücken frei indem er sich um sämtlich administrative Aufgaben in der Proteomics Core Facility kümmerte.

Bei Sandra Helm und Silke Becker möchte ich mich für die großartige Unterstützung beim Betrieb und der Instandhaltung der Massenspektrometer und für die Versorgung mit Kaffee in der Core Facility bedanken.

Bei Ludwig Wiesent möchte ich mich für seine Unterstützung bei Softwareproblemen und bioinformatischen Analysen bedanken.

Bei Saskia Hanf möchte ich mich vor allem für ihre Hilfe bei sämtlichen administrativen Aufgaben bedanken. Ich konnte mit allen Problemen zu ihr kommen und sie war stets hilfsbereit.

Bei Dr. Ursula Olazabal, die leider nicht mehr in unserer Abteilung arbeitet, möchte ich mich ebenfalls für ihre Unterstützung bei administrativen Angelegenheiten bedanken.

Bei Dr. Ronald Roepman und auch bei seinen Mitarbeitern möchte ich mich für die langjährige Kooperation und für seine Unterstützung vor allem in Form von Konstrukten und Antikörpern bedanken. Außerdem war er immer ein sehr guter Diskussionspartner und lieferte gute und konstruktive Kritik.

Bei der Arbeitsgruppe von Prof. Uwe Wolfrum möchte ich mich für die Unterstützung in Form von Antikörpern bedanken.

Bei Doris Bengel möchte ich mich sowohl für nette Gespräche und Diskussionen als auch für den IFT88 Antikörper bedanken.

5. Curriculum Vitae

Name	Karsten Boldt
Address	Moritzstrasse 7
	72131 Ofterdingen
	Germany
Phone	+49 7473 50 20 643
Mobile	+49 179 455 05 08
E-mail	Karsten.Boldt@uni-tuebingen.de
Date of birth	April 19^{th,} 1979
Place of birth	Bietigheim-Bissingen
Nationality	German

Positions held

Since 01/2010	Post doctoral researcher in the Medical Proteome Center Tuebingen at the University of Tuebingen, Germany.
09/2009-12/2009	Post doctoral researcher in the department of protein science with Dr. Marius Ueffing at the Helmholtz-Zentrum München, Munich, Germany.

Education

05/2004-05/2005	Diploma thesis in the proteomics laboratory of PD Dr. Martin Lenter at the Boehringer-Ingelheim Pharma Gmbh & Co. KG, Biberach an der Riss, Germany.
	Title: "Phospho-proteomic analysis of FPRL-1 stimulated human neutrophils".
10/1999-05/2005	Studies of biology at the Eberhard Karls University Tübingen, Germany with the main subject cell biology (molecular cell biology and immunology) and the minor subjects genetics and informatics.
	Qualification: Diploma; Final grade: 1.2.
09/1989-07/1998	Christoph-Schrempf-Gymnasium in Besigheim, Germany.
	Qualification: Abitur; Final grade: 2.2.
08/1985-07/1989	Primary school in Besigheim, Germany.

Additional Work experience

06/2001-12/2003	Part time work in the laboratory of Dr. Teresa Nicolson at the Max-Planck-Institute for developmental biology in Tübingen, Germany.
09/1998-10/1999	Civil service at the Kaywald-School for mentally and physically handicapped children in Lauffen am Neckar, Germany.

Language skills

German	Mother language
English	Fluent in speaking and writing
French	Basic skills

Time abroad

07/2006-09/2006	Beatson institute for cancer research, Glasgow, UK in the proteomics and signal transduction group of professor Walter Kölch.
08/2007-09/2007	Beatson institute for cancer research, Glasgow, UK in the proteomics and signal transduction group of professor Walter Kölch.

Awards

16. Arbeitstagung: Mikromethoden in der Proteinchemie, Martinsried, Germany: Nachwuchspreis Proteinanalytik 2009.