

**Dissertation**  
**zur Erlangung des Doktorgrades**  
**der Fakultät für Chemie und Pharmazie**  
**der Ludwig–Maximilians–Universität München**

**Structural investigation of two supramolecular  
complexes of the eukaryotic cell: the proteasome and  
the mitochondrial TOM complex**

**Patrick Schreiner**

**aus**

**Höchstädt an der Donau**

**2008**

## Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Michael Groll und Herrn Prof. Walter Neupert betreut und von Herrn Prof. Patrick Cramer vor der Fakultät für Chemie und Pharmazie vertreten.

## Ehrenwörtliche Versicherung

Diese Dissertation wurde selbstständig, ohne unerlaubte Hilfe erarbeitet.

München, am 17.07.2008

*Patrick Schreiner*

---

Dissertation eingereicht am 17.07.2008

1. Gutachter: Prof. Dr. Dr. Walter Neupert

2. Gutachter: Prof. Dr. Patrick Cramer

Mündliche Prüfung am 15.10.2008

# Table of contents

<b>Preface</b>	<b>1</b>
<b>1 Structural investigation of Rpn13, the multifunctional adaptor protein of the 26S proteasome</b>	<b>2</b>
1.1 Abstract	2
1.2 Introduction	3
1.2.1 Protein degradation	3
1.2.2 Ubiquitination	5
1.2.3 26S proteasome	6
1.2.3.1 20S proteasome	7
1.2.3.2 19S regulatory subunit	9
1.2.4 Proteasome-associated proteins	12
1.2.5 The novel base component Rpn13	14
1.2.6 Goals of this study	16
1.3 Results	17
1.3.1 Cloning, expression and purification of Rpn13	17
1.3.2 Investigation of the domain architecture of Rpn13	19
1.3.2.1 Primary sequence analysis of Rpn13	19
1.3.2.2 Limited proteolysis of yeast Rpn13	21
1.3.3 Crystal structure of the Pru domain of Rpn13	23
1.3.3.1 Crystallization and structure determination of the N-terminal Pru domain of Rpn13	23
1.3.3.2 Rpn13 Pru adopts Pleckstrin Homology (PH) architecture	27
1.3.3.3 Rpn13 Pru is conserved in eukaryotes	29

1.3.4	Ubiquitin docking at the proteasome via the novel PH domain interaction of Rpn13 Pru	30
1.3.4.1	Investigation of the Pru interaction with ubiquitin by chemical shift perturbation analysis	30
1.3.4.2	Novel Ub binding mechanism of Rpn13 Pru	33
1.3.4.3	Mutational analysis of Pru domain binding to ubiquitin	36
1.3.4.4	Rpn13 binds proteasome subunit Rpn2 independently of Ub	39
1.3.4.5	K48-linked tetraUb binds hRpn13 Pru	41
1.3.5	Investigation of the interaction of Rpn13 Pru with PIPs	43
1.3.6	Interaction of Rpn13 with the ubiquitin C-terminal hydrolase (UCH) 37	45
1.4	Discussion	47
1.4.1	The N-terminal PH domain of Rpn13	47
1.4.2	Comparison of Rpn13 from mouse and yeast	48
1.4.3	Rpn13 loops bind Ub	49
1.4.4	The first proteasomal Pleckstrin Homology domain	52
1.4.5	Rpn13 and its neighbours in the proteasome	53
1.4.6	Rpn13 recruits Uch37 to the proteasome	54
1.5	Materials and Methods	56
1.5.1	Materials	56
1.5.1.1	Oligonucleotides	56
1.5.1.2	Vectors	57
1.5.1.2.1	pRSET-GST-PP	57
1.5.1.2.2	pGEX-6P-1 and pGEX-4T-1	57
1.5.1.2.3	Bicistronic vector system	58
1.5.1.3	Bacteria	59
1.5.1.4	Enzymes	60
1.5.1.5	Equipment and chemicals	60
1.5.1.6	Media and buffers	62

1.5.1.7	DNA isolation and preparation kits	64
1.5.1.8	Columns	65
1.5.1.9	Bioinformatics	65
1.5.2	Methods	66
1.5.2.1	Primer design	66
1.5.2.2	Methods in molecular biology	66
1.5.2.2.1	Polymerase chain reaction (PCR)	66
1.5.2.2.2	Digestion of DNA with restriction endonucleases	67
1.5.2.2.3	Ligation	68
1.5.2.2.4	DNA agarose gel electrophoresis	69
1.5.2.2.5	Competent <i>E. coli</i> cells	69
1.5.2.2.5.1	Chemical competent <i>E. coli</i> cells	69
1.5.2.2.5.2	Electro-competent <i>E. coli</i> cells	70
1.5.2.2.6	Transformation	70
1.5.2.2.7	Isolation of plasmid DNA from <i>E. coli</i>	70
1.5.2.2.8	Mutagenesis	71
1.5.2.2.9	DNA sequencing	71
1.5.2.3	Methods in protein biochemistry	72
1.5.2.3.1	Gene expression	72
1.5.2.3.2	Cell lysis	72
1.5.2.3.3	Purification of GST-fused proteins	72
1.5.2.3.4	Anion exchange chromatography (MonoQ)	73
1.5.2.3.5	Size exclusion chromatography	73
1.5.2.3.6	Expression and purification of selenomethionine labelled protein	74
1.5.2.3.7	Transfer of proteins to nitrocellulose membrane (Western-blot)	75
1.5.2.4	Detection and quantitative determination	75
1.5.2.4.1	SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)	75
1.5.2.4.2	Staining SDS-PA gels with Coomassie brilliant blue	76
1.5.2.4.3	Immunodecoration	76
1.5.2.4.4	Determination of the protein concentration by Bradford assay	77
1.5.2.4.5	Determination of the protein concentration by UV spectroscopy	77
1.5.2.4.6	Concentration of proteins	77

1.5.2.4.7	Protein precipitation	78
1.5.2.5	Limited proteolysis	78
1.5.2.6	Mass spectrometry	78
1.5.2.7	Crystallization	78
1.5.2.8	Structure determination	79
1.5.2.9	Protein–lipid overlay assay	80
1.5.2.10	NMR spectroscopy	80
1.5.2.10.1	$^{15}\text{N}$ , $^{13}\text{C}$ and $^2\text{H}$ labelling of proteins	80
1.5.2.10.2	NMR spectra	81
1.5.2.10.3	Chemical shift perturbation analysis	82
1.5.2.10.4	Docking protocol	83
1.5.2.11	Modelling of mRpn13 Pru:diUb complex	84
1.5.2.12	Biochemical pull–down assays	84
1.5.2.13	Rpn2 binding assays	85
<b>2</b>	<b>Crystallographic studies of the TOM core complex</b>	<b>86</b>
2.1	Abstract	86
2.2	Introduction	87
2.2.1	Mitochondria	87
2.2.2	Protein import into mitochondria	89
2.2.2.1	Presequence import pathway	90
2.2.2.2	Carrier pathway	91
2.2.2.3	Intermembrane space import and assembly pathway	92
2.2.2.4	Outer membrane sorting and assembly pathway	92
2.2.3	The TOM complex	93
2.2.4	Goals of this study	95
2.3	Results	96
2.3.1	Isolation and purification of TOM core complex	96

2.3.2	Generation of murine monoclonal antibodies recognizing native epitopes of TOM core complex	97
2.3.3	Characterization of selected monoclonal antibodies	99
2.3.3.1	Immunological subtyping of murine antibodies	99
2.3.3.2	Western Blot analysis for selection of conformation-specific antibodies	100
2.3.3.3	Pull-down assay with hybridoma supernatants	100
2.3.3.4	Purification of murine antibodies from hybridoma cell culture	102
2.3.3.5	Binding studies via analytical gel filtration	103
2.3.4	The TOM core in complex with monoclonal antibody fragments	104
2.3.4.1	Expression and purification of Fv fragments	104
2.3.4.2	Copurification of Fv fragments with TOM core complex	105
2.3.5	Crystallographic analysis of TOM core complex	107
2.3.5.1	Crystallization of TOM core complex	107
2.3.5.2	Optimization of the diffraction quality of the crystals	108
2.3.5.2.1	Experiments to improve crystallization	108
2.3.5.2.2	Manipulation of the crystals	109
2.3.5.2.3	Heavy metal atom soaks and cocrystallization	110
2.3.5.3	Crystallization of TOM core:P1C10 antibody complex	111
2.3.5.4	Detergent exchange during purification	111
2.3.5.5	Crystallographic data on TOM core complex	113
2.3.6	Investigation of the pore-forming component Tom40	115
2.3.6.1	Secondary structure prediction of Tom40	115
2.3.6.2	Sequence conservation of Tom40	117
2.3.6.3	Recombinant expression and refolding of Tom40	117
2.4	Discussion	119
2.4.1	Membrane protein crystallization	119
2.4.1.1	Detergents in crystallization of TOM core complex	119
2.4.1.2	Lipid requirements of TOM core complex	121
2.4.1.3	Additive approach in TOM core complex crystallization	122

2.4.2	Antibody-fragment mediated crystallization of TOM core complex	123
2.4.3	Expression of $\beta$ -barrel protein Tom40	126
2.5	Materials and Methods	128
2.5.1	Materials	128
2.5.1.1	Oligonucleotides	128
2.5.1.2	Vectors	130
2.5.1.2.1	pRSET-PL0 and pRSET-PL1	130
2.5.1.2.2	pET-Duet™-1	130
2.5.1.2.3	pCR2.1-TOPO and pDrive	130
2.5.1.2.4	pASK68	131
2.5.1.3	Bacteria	131
2.5.1.4	Enzymes	132
2.5.1.5	Equipment and chemicals	132
2.5.1.6	Media and Buffers	134
2.5.1.7	DNA/RNA Isolation and Preparation Kits	138
2.5.1.8	Columns	139
2.5.1.9	Bioinformatics	139
2.5.2	Methods	140
2.5.2.1	Primer design	140
2.5.2.2	Methods in molecular biology	140
2.5.2.2.1	Polymerase chain reaction (PCR)	140
2.5.2.2.2	RNA isolation from hybridoma cell culture and Reverse Transcription	141
2.5.2.2.3	Digestion of DNA with restriction endonucleases	142
2.5.2.2.4	Ligation	143
2.5.2.2.5	DNA agarose gel electrophoresis	143
2.5.2.2.6	Competent <i>E. coli</i> cells	144
2.5.2.2.6.1	Chemical competent <i>E. coli</i> cells	144
2.5.2.2.6.2	Electro-competent <i>E. coli</i> cells	144
2.5.2.2.7	Transformation	144
2.5.2.2.8	Isolation of plasmid DNA from <i>E. coli</i>	145
2.5.2.2.9	Mutagenesis	146

	2.5.2.2.10	DNA sequencing	146
2.5.2.3		Production of monoclonal antibodies	146
	2.5.2.3.1	Immunization of mice and cell fusion	146
	2.5.2.3.2	Hybridoma cell culture	148
	2.5.2.3.3	Single cell cloning	149
	2.5.2.3.4	Freezing and cryo-storage of hybridoma cells	149
	2.5.2.3.5	Enzyme Linked ImmunoSorbent Assays (ELISA)	149
	2.5.2.3.6	Determination of the antibody subtype	150
	2.5.2.3.7	Large-scale production of monoclonal antibodies	150
	2.5.2.3.8	Monoclonal antibody purification by protein A affinity chromatography	151
	2.5.2.3.9	Analytical gel filtration binding assay	151
	2.5.2.3.10	Biochemical pull-down assay	152
2.5.2.4		Methods in protein biochemistry	152
	2.5.2.4.1	Gene expression	152
	2.5.2.4.2	Cell lysis	153
	2.5.2.4.3	Periplasmic extraction	153
	2.5.2.4.4	Isolation of inclusion bodies	153
	2.5.2.4.5	Protein refolding	153
	2.5.2.4.6	Streptavidin affinity purification	154
	2.5.2.4.7	Purification of TOM core complex	154
	2.5.2.4.8	Co-complex formation of TOM core complex and Fv fragments	155
	2.5.2.4.9	Detergent exchange	155
2.5.2.5		Detection and quantitative determination of the protein	155
	2.5.2.5.1	SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)	155
	2.5.2.5.2	Staining SDS-PA gels with Coomassie brilliant blue	157
	2.5.2.5.3	Silver staining of SDS-PA gels	157
	2.5.2.5.4	Transfer of proteins to nitrocellulose membrane (Western-blot)	158
	2.5.2.5.5	Immunodecoration	158
	2.5.2.5.6	Determination of the protein concentration by Bradford assay	158
	2.5.2.5.7	Determination of the protein concentration by UV spectroscopy	159

	2.5.2.5.8	Concentration of proteins	159
	2.5.2.5.9	Protein precipitation	159
2.5.2.6		Crystallization	160
	2.5.2.6.1	Crystallization of TOM core complex	160
	2.5.2.6.2	Crystallization in capillary	160
	2.5.2.6.3	Crystallization with additives	160
	2.5.2.6.4	Treatment of crystals with glutaraldehyde	161
	2.5.2.6.5	Crystallization with heavy metal atoms	161
2.5.2.7		Data collection and processing	161
<b>3</b>		<b>Abbreviations</b>	<b>163</b>
<b>4</b>		<b>References</b>	<b>165</b>

## **Preface**

This study is focussed on the structural investigation of large molecular assemblies such as the 26S proteasome and the translocation machinery of the outer mitochondrial membrane. It is divided in two chapters and in both parts the structural and further functional analysis is based on X-ray crystallography.

### **Chapter 1: Structural investigation of Rpn13, the multifunctional adaptor protein of the 26S proteasome**

The results in chapter 1 reveal that the multifunctional adaptor protein Rpn13 acts as a novel ubiquitin receptor of the 26S proteasome and deliver structural and biophysical details of its interaction with ubiquitin and with other proteasomal subunits. The crystal structure of the ubiquitin binding domain of Rpn13 reveals the molecular architecture of a Pleckstrin Homology (PH) domain and the NMR structure of the complex with ubiquitin shows a novel ubiquitin-binding mode. Additional NMR studies and domain mapping by truncation analysis provide further insights in the domain architecture of Rpn13 and the interaction with its partners Rpn2 and Uch37.

### **Chapter 2: Crystallographic studies of the TOM core complex**

Chapter 2 presents the purification and crystallization of the mitochondrial protein translocase, the TOM core complex, from *Neurospora crassa*. Preliminary crystallographic data lead to the determination of space group and cell dimensions. This chapter also describes various experiments to improve the diffraction quality of the crystals and the co-crystallization of TOM core complex with specific monoclonal antibody fragments. Furthermore, expression and refolding of the main component Tom40 is raised as an alternative approach in structural investigation of the TOM complex.

# **1 Structural investigation of Rpn13, the multifunctional adaptor protein of the 26S proteasome**

## **1.1 Abstract**

Proteasomes typically recognize substrate proteins by their covalently attached ubiquitin chains. The components that perform this role have remained elusive, with Rpn10/S5a being the only established proteasomal ubiquitin receptor. This study provides structural and biophysical details of the multifunctional, proteasomal adaptor protein Rpn13 and presents the crystal structure of its ubiquitin binding domain, which forms a Pleckstrin Homology (PH) domain fold and thus was named the Pru domain (Pleckstrin-like recceptor for ubiquitin). The NMR structure of Rpn13 Pru:ubiquitin complex reveals a novel ubiquitin-binding mode in which loops rather than secondary structural elements are used to capture ubiquitin. This is supported by mutational analysis of the interacting amino acid residues of Rpn13. Moreover, NMR studies exhibit that ubiquitin and the proteasomal subunit Rpn2/S1/PSMD1 bind simultaneously through surfaces at opposite sides of the Pru domain. Rpn13 enhances the deubiquitinating activity of its binding partner Uch37, and thus jointly promotes the recognition and processing of ubiquitin chains. As a step towards the mechanism of Rpn13:Uch37 interaction, this study describes the heterologous bicistronic coexpression of the human co-complex. Domain mapping localized the binding module at the C-terminal ends of both proteins. A model structure of Rpn13 complexed to diubiquitin is provided, which affords insights into how the role of Rpn13 as an ubiquitin receptor is coupled to substrate deubiquitination by Uch37. Rpn13 represents a multifunctional adaptor protein involved in recognition and targeting of ubiquitylated substrate proteins to the proteasome and in recruitment and activation of Uch37<sup>1</sup>.

## 1.2 Introduction

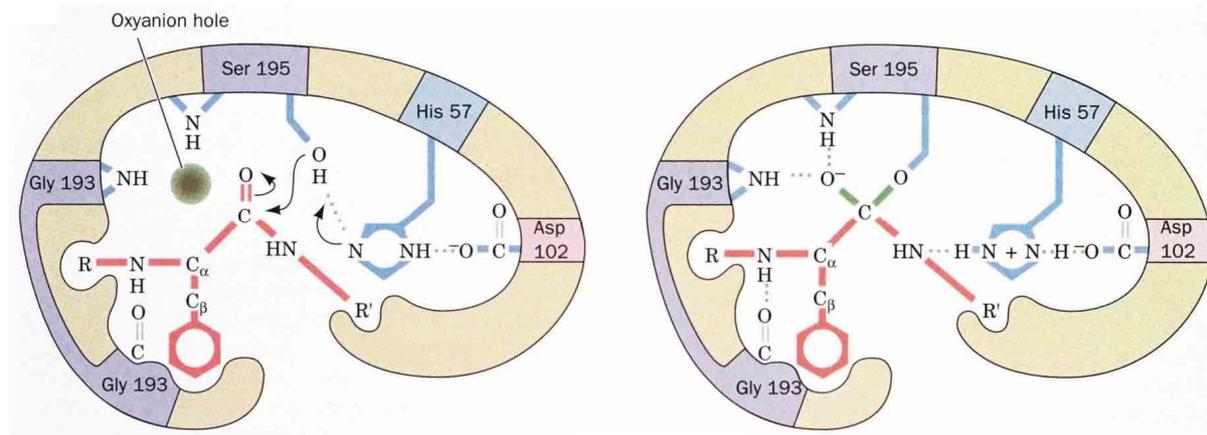
### 1.2.1 Protein degradation

Protein degradation plays a key role in regulation of various biological processes such as metabolism, cell cycle progression, antigen presentation and apoptosis<sup>2,3</sup>. Degradation is linked to the large class of enzymes called proteases and peptidases, which naturally occur in all organisms and constitute 1–5% of the gene content. Systematically, proteases belong to the group of C–N–hydrolases and are capable of catalyzing the cleavage of peptide bonds. They can either break specific peptide bonds (limited proteolysis), depending on the amino acid sequence of a protein, or break down a complete peptide to amino acids (unlimited proteolysis). Depending on the type of attack on the protein, proteases are divided into two categories, namely exo– and endopeptidases. Exopeptidases detach the terminal amino acids from N– or C–terminus of the polypeptide chain (*e.g.* aminopeptidases, carboxypeptidases), whereas endopeptidases attack internal peptide bonds (*e.g.* pepsin and trypsin). The two groups are further subdivided according to the character of their catalytic active site and reaction mechanism.

Proteases are currently classified into six groups:

- |    |                         |  |
|----|-------------------------|--|
| 1. | Serine proteases        | Ser and His residue in the active centre |
| 2. | Threonine proteases     | Thr residue in the active centre         |
| 3. | Cysteine proteases      | Cys residue in the active centre         |
| 4. | Aspartic acid proteases | Asp residue in the active centre         |
| 5. | Metalloproteases        | Metal ion in the active centre           |
| 6. | Glutamic acid proteases | Glu residue in the active centre         |

The mechanism used to cleave a peptide bond involves a catalytic amino acid residue (serine, threonine and cysteine proteases) or a water molecule (as for aspartic acid, metallo– and glutamic acid proteases) which has to be converted to a nucleophile, so that it is able to attack the peptide carbonyl group. For example, the mechanistic class of serine proteases was originally distinguished by the presence of the Asp-His-Ser “charge relay” system or “catalytic triad”<sup>4</sup>. The catalytic triad spans the active site cleft, with serine on one side and aspartic acid and histidine on the other (Figure 1.1). The backbone NHs of glycine and serine built a pocket of positive charge, the oxyanion hole, which activates the carbonyl of the scissile peptide bond and stabilizes the negatively charged oxyanion of the tetrahedral intermediate (Figure 1.1).



**Figure 1.1:** (a) Schematic view of the catalytic mechanism of serine proteases. The peptide and the catalytic residues are depicted in red and blue, respectively<sup>5</sup>.

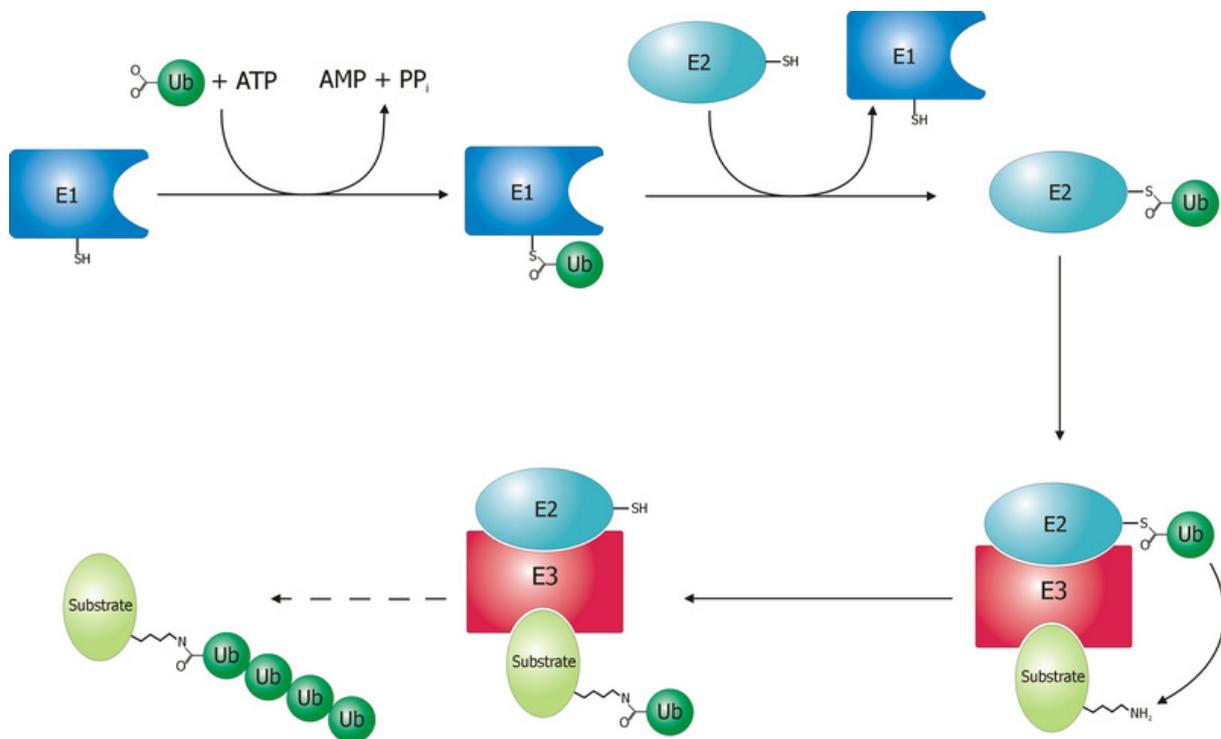
Proteinases are involved in a multitude of physiological reactions from simple digestion of food proteins to highly-regulated cascades<sup>2,3</sup>. Proteolysis can be a destructive change, abolishing a protein's function or digesting to its principal components; it can be an activation of a function, or a signal in a signaling pathway. Acid proteases secreted into the stomach (such as pepsin) and serine proteases present in the duodenum (trypsin and chymotrypsin) enable digestion of proteins ingested by food. Blood serum contains proteinases (thrombin, plasmin, *etc.*) playing an important role in blood-clotting, as well as lysis of the clots, and in maintenance of the immune system. Degradation by proteases determines the lifetime of proteins which accomplish important physiological roles like hormones, antibodies, enzymes or transcription factors, thereby providing a fast regulatory relay system in the cell. The process is highly selective: some proteins are degraded within minutes, while others are practically stable.

Protein hydrolysis represents an irreversible process and therefore substrate selection for degradation has to be strictly controlled. This is achieved by different mechanisms such as high substrate specificity (*e.g.* signal peptidases), inhibitory regulation (*e.g.*  $\alpha_1$ -antitrypsin and elastase), activation by limited proteolysis of zymogens (*e.g.* blood clotting system), co-localization (*e.g.* in membranes), binding to cofactors (*e.g.* metalloproteases) and the confinement to specialized compartments (*e.g.* proteasome)<sup>6</sup>.

The vast majority of cellular proteins is degraded by the cytosolic ubiquitin (Ub)-proteasome pathway. In this pathway, substrate proteins which are destined for degradation are marked by covalent attachment of ubiquitin in a process termed ubiquitination or ubiquitylation.

## 1.2.2 Ubiquitination

Proteins targeted for proteasomal degradation are tagged by the small protein ubiquitin which consists of 76 amino acids and which is highly conserved among eukaryotic species. Ubiquitin molecules are then further added on to previously-conjugated ubiquitin molecules to form a polyubiquitin chain mediating the recognition and degradation of substrate proteins by the 26S proteasome<sup>2</sup>. If the chain is longer than 3 ubiquitin molecules, the tagged protein is rapidly degraded into small peptides. Ubiquitin moieties are cleaved off the protein by deubiquitinating enzymes and are recycled for further use. The process of ubiquitination involves the ubiquitin-activating enzyme E1, ubiquitin-conjugating enzyme E2, and ubiquitin ligase E3 and consists of a several steps (Figure 1.2).



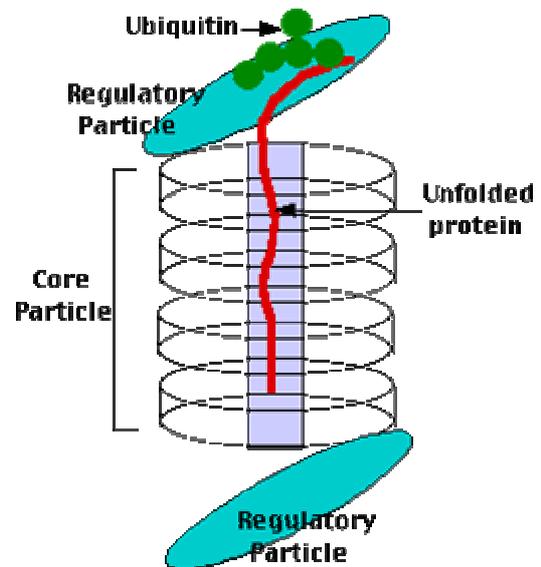
**Figure 1.2:** Proteins are targeted for degradation by the proteasome by covalent modification of a lysine residue that requires the coordinated reactions of three enzymes. In the first step, an ubiquitin-activating enzyme (known as E1) hydrolyzes ATP and adenylates an ubiquitin molecule. This is then transferred to E1's active-site cysteine residue in concert with the adenylation of a second ubiquitin. This adenylated ubiquitin is then transferred to a cysteine of a second enzyme, the ubiquitin-conjugating enzyme (E2). Lastly, a member of a highly-diverse class of enzymes known as ubiquitin ligases (E3) recognizes the specific protein to be ubiquitinated and catalyzes the transfer of ubiquitin from E2 to this target protein.

It starts with the activation of ubiquitin in a two-step reaction by an E1 ubiquitin-activating enzyme. This initial process requires ATP as an energy source and involves production of an ubiquitin-adenylate intermediate. The second step transfers ubiquitin to the E1 active site cysteine residue upon release of AMP which results in a thioester linkage between the C-terminal carboxyl group of ubiquitin and the E1 cysteine sulfhydryl group. Afterwards, the ubiquitin is transferred from E1 to the active site cysteine of a ubiquitin-conjugating enzyme E2 via a trans(thio)esterification reaction. The final step of the ubiquitylation cascade is mediated by one of the hundreds of E3 ubiquitin-protein ligases. The E3, which may be a multi-protein complex, is generally responsible for targeting ubiquitination to specific substrate proteins and therefore acts as the substrate recognition module of the system (Figure 1.2). E3 ligases are capable of interaction with both E2 and substrate. The number of expressed E1, E2, and E3 proteins depends on the organism and cell type, but there are many different E3 enzymes present in eukaryotes, indicating that there is a huge number of targets for the 26S proteasome.

### 1.2.3 26S proteasome

The 26S proteasome is a huge macromolecular machine of approximately 2.5 MDa and represents a compartmentalized, ATP-dependent protease composed of more than 30 subunits that recognizes and degrades polyubiquitinated substrates. It is responsible for most of the nonlysosomal protein degradation in both the nucleus and cytosol. Its function is critical to a wide variety of cellular processes such as protein quality control, antigen processing, signal transduction, cell-cycle control, cell differentiation and apoptosis<sup>7,8</sup>. In higher eukaryotes, the subcellular location of the proteasome is mainly cytoplasmic and nuclear. In yeast cells, however, the nuclear envelope-endoplasmic reticulum network and the nuclear periphery seem to be predominant sites of 26S proteasome action (ERAD = Endoplasmic Reticulum Associated Protein Degradation)<sup>9</sup>. Inhibition of this enzyme allowed identification of numerous protein substrates and revealed the importance of the proteasome-ubiquitin pathway. Consequently, a functionally impaired ubiquitin-proteasome system has been implicated in several human diseases<sup>10</sup>. The 26S proteasome contains a proteolytically active 20S core particle (CP) capped at one or both ends by 19S regulatory particles (RP)<sup>8</sup>, which are responsible for substrate recognition, unfolding and translocation (Figure 1.3).

**Figure 1.3:** Schematic representation of the 26S proteasome consisting of 20S core and the 19S regulatory particle. Substrate proteins are marked with polyubiquitin, unfolded and translocated into the core.

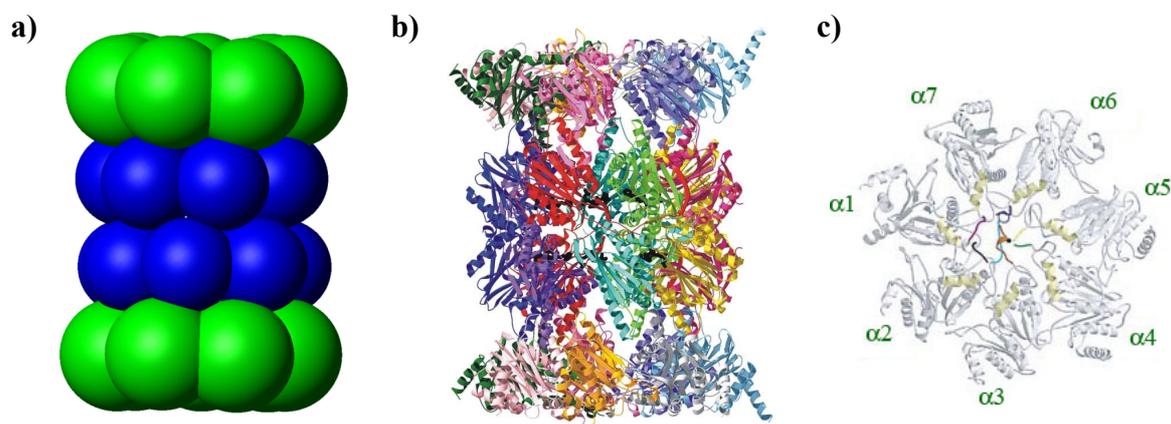


Core particles are ubiquitous among all three kingdoms of life. However, in contrast to eukaryotes, archaea contain a much more simply organized core particle, whereas eubacteria possess a proteasome-like protease termed HslV<sup>3</sup>. Despite all the differences, a common feature of the CPs and proteasome-related systems is the generation of peptide fragments with an average length distribution of about 2–29 amino acids residues. As a rule, these peptides have to be further degraded by other proteases to single amino acids, which can be utilized for *de novo* synthesis<sup>3</sup>. It is currently believed that proteasome degradation products in eukaryotes are further digested by prolyl oligopeptidase, thimet oligopeptidase and tripeptidyl peptidase II<sup>3,11,12</sup>.

### 1.2.3.1 20S proteasome

The most stable subcomplex of the 26S proteasome is the proteolytic 20S core particle (CP). It is a large threonine protease with a molecular mass of about 700 kDa. All CPs consist of four stacked heptameric ring structures following a cylindrical architecture (Figure 1.4)<sup>13-15</sup>. The heptameric rings themselves are composed of two different types of subunits, termed  $\alpha$  and  $\beta$ . As anticipated from sequence comparison, the noncatalytic  $\alpha$  subunit and the catalytic  $\beta$  subunit show similar folding, a four-layer structure in which two sheets of  $\beta$  strands are sandwiched between two layers of  $\alpha$  helices<sup>14,16</sup>. The two outer rings of the barrel shaped core are built of seven  $\alpha$  subunits, which serve as docking domains for the regulatory particles and form an exterior gate blocking unregulated access to the interior cavity<sup>17,18</sup>. As evident from the crystal structure of yeast CP, the N-termini of the  $\alpha$  subunits protrude into the gate region of the cylinder in the absence of activating complexes, thus occluding substrate entry

(Figure 1.4c)<sup>14</sup>. Docking of an activator complex to the outer rim of the CP cylinder induces opening of the gate by a structural rearrangement of these N-termini<sup>18,19</sup>. The  $\beta$  subunits form the inner two heptameric rings, which contain the proteolytically active centres of the proteasome that face into the hollow interior of the CP. The interior chamber of the 20S proteasome is at most 53 Å wide, though the entrance can be as narrow as 13 Å, suggesting that substrate proteins must be at least partially unfolded to enter<sup>20</sup>.



**Figure 1.4:** (a) Schematic diagram of the proteasome 20S core particle viewed from one side. The  $\alpha$  subunits that make up the outer two rings are shown in green, and the  $\beta$  subunits that constitute the inner two rings are shown in blue. (b) Structure of yeast 20S proteasome viewed from the side. All subunits are depicted in different colours. (c) Top view of yeast 20S proteasome showing the  $\alpha$  subunits. The N-termini block the substrate entry.

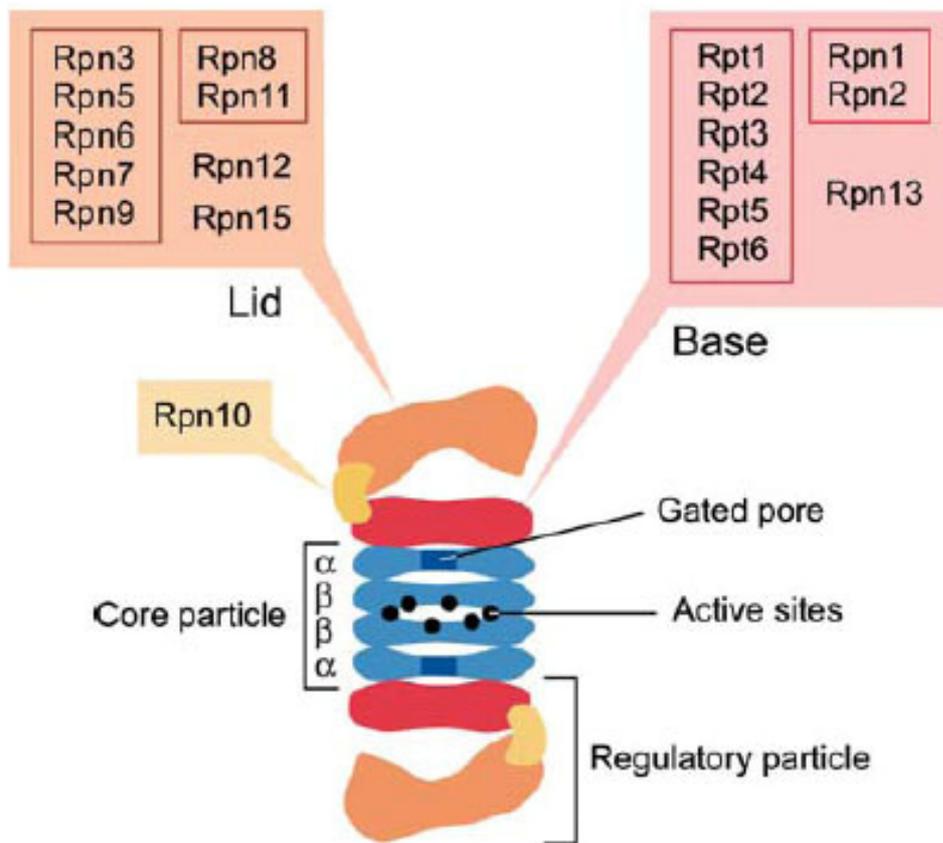
The number and the diversity of  $\alpha$  and  $\beta$  subunits depend on the organism. The number of distinct and specialized subunits increases from prokaryotes to eukaryotes. In archaea such as *Thermoplasma acidophilum*, all  $\alpha$  and  $\beta$  subunits are identical, whereas proteasomes from eukaryotic organisms such as yeast contain seven distinct but related types of each subunit. All 14 different eukaryotic proteasomal subunits contain characteristic insertion segments and termini, which represent well-defined contact sites between related subunits and lead to their unique locations at special positions within the particle. In eukaryotes, only the  $\beta 1$ ,  $\beta 2$ , and  $\beta 5$  subunits are proteolytically active and follow a common peptide bond cleavage mechanism. The architecture and the polarity of the substrate binding channel is unique for each of the sites, thereby generating a caspase-like cleavage site for  $\beta 1$ , a trypsin-like cleavage site for  $\beta 2$  and a chymotryptic-like activity for  $\beta 5$ <sup>14,21,22</sup>. Proteasomal active subunits belong to the family of Ntn-hydrolases, which use their N-terminal residue Thr1 as the catalytic nucleophile. All Ntn-hydrolases require a processing step that results in the exposure of the N-terminal amino group acting as a nucleophile. Therefore, 20S proteasomes have to follow a defined maturation pathway based on intrasubunit autolysis leading to the functionally active

Ntn–protease complex<sup>3</sup>. In eukaryotes, the 20S core particle can be capped by regulatory complexes such as 11S (PA28) and 19S (PA700), which regulate substrate access and stimulate peptide hydrolysis. The 19S regulatory complex forms the most prominent and important regulator providing the link for proteasome–mediated proteolysis with the ubiquitin pathway of protein degradation<sup>2,8,23,24</sup>.

### 1.2.3.2 19S regulatory subunit

The ATP–independent 20S proteasome is fairly well characterized, and its maturation, catalytic mechanism, broad specificity, regulation and interactions have been described in detail. So far, there is less available data on the ATP–dependent 19S regulatory particle (RP), which functions as a gate opener in addition to its important roles in recognition of ubiquitinated substrate, deconjugation of Ub chains, and unfolding of substrates prior to their translocation into the CP<sup>8,25</sup>. The regulatory particle, consisting of at least 19 subunits (Figure 1.5), can occupy one or both ends of the cylindrical core particle. Although a large number of subunits has been studied genetically, the functions of many subunits are still unknown<sup>8</sup>. Based on biochemical characteristics, the RP can be further subdivided in two assemblies, the “base” and the “lid”, which are located proximal and distal to the CP, respectively (Figure 1.5)<sup>26</sup>. The base is involved in unfolding of substrate proteins and their subsequent translocation into the central proteolytic chamber of the CP, whereas the lid–complex is expected to function as the recognition scaffold for the marked polyubiquitinated substrates. The 19S subunits are designated by the mammalian “S” (subunit)<sup>27</sup> or yeast “Rp” (regulatory particle)<sup>28</sup> nomenclatures.

The base complex contains six related, essential ATPases (Rpt1/S7, Rpt2/S4, Rpt3/S6, Rpt4/S10b, Rpt5/S6, Rpt6/S8), which are responsible for the ATP requirement of proteasome activity, the two largest subunits Rpn1/S2 and Rpn2/S1 and the recently identified subunit Rpn13/Adrm1<sup>29–32</sup>. ATPase complexes that regulate protein degradation in eukaryotes, bacteria and archaea are all members of the AAA<sup>+</sup> (ATPases associated with various cellular activities) ATPase superfamily<sup>33</sup>. Specific interactions between the 19S Rpts (for regulatory particle triple–A protein) supported by cross–linking studies<sup>34</sup> and the analogy to their bacterial counterparts suggest formation of a six–membered ring, which presumably associates directly with the  $\alpha$  subunits of the core particle<sup>6,35</sup>.



**Figure 1.5:** Schematic representation of the yeast 26S proteasome. The CP is shown in blue, with active sites indicated in black. Docking of the RP (shown in red, orange, and yellow) allows opening of the gate and creates an open channel into the central cavity. The regulatory subunit is composed of two entities: the base (red) and the lid (orange), connected by Rpn10 (yellow). Grouping of subunits is based on structural motifs or homology. Lid: PCI motifs (Rpn3, Rpn5, Rpn6, Rpn7, and Rpn9), Jab/MPN motifs (Rpn8 and Rpn11), with the remaining subunits ungrouped. Base: AAA proteins (all Rpts), UBL receptors bearing LRR-like repeats (Rpn1 and Rpn2), and Rpn13 remains ungrouped. Rpn10 lies outside of the traditional lid/base definition

Since structural studies indicate that the gate region of the 20S core is blocked by the N-termini of the  $\alpha$  subunits, substrate entry into the proteolytic chamber is permitted<sup>14,18</sup>. It has been proposed that the ATPases open a channel to the interior of the core and are additionally responsible for substrate unfolding, which is required for efficient degradation by the proteasome<sup>36</sup>. The proteasomal ATPases possess a chaperone-like activity and are able to change the tertiary structure of a bound protein under ATP hydrolysis, eventually resulting in unfolding<sup>37,38</sup>. The two large subunit of the base, Rpn1 and Rpn2 (regulatory particle non-ATPase), are related in sequence and show similarity to the leucine-rich repeat (LRR) domain, a common site for protein-protein interaction<sup>9,39</sup>. They are proposed to form  $\alpha$ -solenoid structures that contact the ATPases from the side opposite to the CP<sup>40</sup>. Functional

studies provide evidence that both subunits are able to bind to the UBL (ubiquitin-like) domains of Rad23 and Dsk2 through their LRR-like domain and mediate recognition of these proteins by proteasome<sup>41,42</sup>. The yeast Rad23 and Dsk2 belong to a family of proteins that contain a UBL and an ubiquitin-chain binding UBA (ubiquitin-associated) domain. Given that these UBL-UBA proteins can bind both ubiquitin and the proteasome, they might serve as adaptors that deliver ubiquitylated cargo to proteasomes<sup>43-45</sup>.

Besides unfolding of substrate proteins, a second major function of the RP is to provide binding sites for substrates conjugated to ubiquitin chains. So far, two components of the RP have been implicated in binding of ubiquitin conjugates, Rpn10/S5a<sup>46,47</sup> and Rpt5<sup>48</sup>. Rpn10, which lies outside of the traditional lid/base definition, connects the base and the lid and recognizes ubiquitin-marked proteins via its UIM (ubiquitin-interacting motif) element. The base alone is competent to activate the peptide and casein-hydrolyzing activity of the proteasome CP to levels almost equivalent to those of wild-type proteasomes. However, degradation of ubiquitin-protein conjugates requires the intact RP.

The lid as the second subassembly of the RP is composed of nine subunits (Rpn3/S3, Rpn5, Rpn6/S9, Rpn7/S10a, Rpn8/S12, Rpn9/S11, Rpn11/S13, Rpn12/S14 and Rpn15/Sem1/Dss1, Figure 1.5) and although seven of them are essential for viability in yeast, indicating a critical function during proteasomal substrate degradation, there is little structural and functional information available on their mechanistic roles. The lid subunits share sequence motifs (the PINT/PCI and the MPN motif) with components of the COP9/signalosome and the eukaryotic initiation factor 3 (eIF3) complexes, suggesting that these functionally diverse particles have a common evolutionary ancestry<sup>49-51</sup>. The PINT/PCI domain is up to 200 residues in length and is predicted to form an  $\alpha$ -helical structure<sup>52,53</sup>. These domains have been found at the C-termini of Rpn3, Rpn5, Rpn6, Rpn7, and Rpn9. The MPN domain, located in the N-termini of Rpn8 and Rpn11, spans approximately 140 residues and is predicted to assume  $\alpha/\beta$  structure<sup>52</sup>. Remarkably, all of the proteasome subunits that possess these motifs are found in the lid of the RP. Rpn12 and Rpn15 are the two components of the lid, which lack these structural motifs. Rpn11 functions as a metallo-isopeptidase that cleaves polyubiquitin chains from substrates en bloc<sup>54,55</sup>. It belongs to the family of deubiquitinating enzymes (DUBs), which are suggested to assist in the removal of the polyubiquitin moiety from protein substrates prior to their translocation into the catalytic chamber. The smallest, non-essential proteasomal subunit, Rpn15, was initially identified as a suppressor of the growth arrest induced by overexpression of the UBL-UBA protein Dsk2<sup>31</sup>. Purified proteasomes lacking Rpn15 exhibit stability defects and its human ortholog, Dss1, has been implicated in split-hand/split-foot disease<sup>31,56</sup>. All of the known RP components are tightly associated with the particle, with the two exceptions of Rpn4/Son1/Ufd5 and Rpn14/PAAF1<sup>57,58</sup> and all are

conserved in eukaryotes. The Rpt subunits are 66–67% identical between yeast and human, whereas the non-ATPase subunits show lower amount of sequence identity, typically in the range of 33–47%<sup>39</sup>.

In addition to the proteins that form the proteasome holocomplex, there is increasing variety of associating proteins, which not only influence the overall stability of the complex, but assist in different steps of proteasomal degradation. The increasing number of these associated proteins suggests that proteasomes are larger and more diverse in composition than previously assumed.

### 1.2.4 Proteasome-associated proteins

Proteasomal function is not only affected by its intrinsic components, but also by a wide variety of associating proteins. A group of ubiquitin-binding proteins assists in delivery of substrates to the proteasome, whereas proteasome-associated ubiquitin ligases and deubiquitinating enzymes may alter dynamics of ubiquitin chains already associated with the proteasome. The recognition of polyubiquitinated substrates by the 26S proteasome is essential for protein degradation in the cell, but the underlying mechanisms are not fully understood. Subunit Rpn10 clearly contributes to the capacity of the proteasome to bind ubiquitinated substrates<sup>32,46,47</sup> through binding to ubiquitin chains via its UIM element. However, deletion of the RPN10 gene only produces mild cellular defects<sup>59</sup>, suggesting alternative recognition pathways or existence of other ubiquitin receptors. Two such alternatives may be substrate recognition by Rpt5<sup>48</sup> and the UBL-UBA shuttle system. UBL-UBA proteins, such as Rad23, Dsk2 and Ddi1 in *S. cerevisiae*, which are ubiquitin receptors involved in proteasomal targeting, generally have an amino-terminal region that contains one UBL domain that is recognized by the proteasome<sup>41,43,47,60-62</sup>, and one or more UBA domains, which bind to ubiquitin<sup>63</sup>. Presumably, these proteins transport ubiquitinated substrates to the proteasome and in some cases initially recognize the ubiquitin conjugates while they are still bound to the E3 ligase that catalyzed their synthesis<sup>64-66</sup>. The proteasomal receptors for the UBL-UBA proteins are the base subunits Rpn1 and Rpn2<sup>41,42,67</sup>.

The existence of deubiquitinating activity within the proteasome had been anticipated well before its discovery, based on a likely requirement to recycle ubiquitin. At least three deubiquitinating enzymes are thought to contribute to deubiquitination by the proteasome: Rpn11, Ubp6 and Uch37<sup>54,55,68-70</sup>. Rpn11 couples the deubiquitination and degradation of proteasome substrates and point mutations in its metal-coordinating site impair substrate degradation<sup>54,55</sup>. The second proteasomal deubiquitinating enzyme, the cysteine protease

Ubp6, associates with the base and contains an N-terminal UBL domain which is necessary for the interaction with Rpn1<sup>68,71</sup>. Association with the proteasome markedly stimulates Ubp6's catalytic activity<sup>68</sup>, indicating an intimate functional relationship between Ubp6 and the proteasome. Moreover, Ubp6 seems to prevent translocation of ubiquitin into the core particle of the proteasome, and is therefore crucial for the maintenance of normal ubiquitin levels within the cell<sup>68,72,73</sup>. The third deubiquitinating enzyme (DUB) within the regulatory particle of proteasome and the only ubiquitin carboxy-terminal hydrolase (UCH)-family protease that is associated with mammalian proteasomes, is Uch37. In contrast to Rpn11, Uch37 disassembles polyubiquitin specifically from the distal end of the chain, a property that may be used to clear poorly ubiquitinated or unproductively bound substrates from the proteasome<sup>54,55</sup>. Therefore, Uch37 was proposed to function as an "editing" protease<sup>74,75</sup>. The base subunit Rpn13 recruits Uch37 to the proteasome and is able to stimulate its enzymatic activity<sup>76-78</sup>. Uch37 appears to be an abundant component of proteasomes from mammalian cells, *Drosophila*, and fission yeast, but is absent from budding yeast<sup>75,76,79,80</sup>.

A third class of proteasome-associated proteins are the ubiquitin ligases (also called E3 ubiquitin ligases) which covalently attach ubiquitin to a lysine residue on a target protein. The ubiquitin ligase is typically involved in polyubiquitination, which marks proteins for degradation by the proteasome. It operates in conjunction with an E1 ubiquitin-activating enzyme and an E2 ubiquitin-conjugating enzyme (see Section 1.2.2). Traditionally, the ubiquitination of substrates and their delivery to the proteasome have been viewed as independent steps in proteasomal protein degradation. The direct association of a variety of E2 ubiquitin-conjugating enzymes and E3 ubiquitin ligases with proteasomes suggests that, for some substrates, these two steps may be coupled. The most abundant of the ubiquitin ligases is the HECT domain-containing E3 ubiquitin ligase Hul5 (KIAA10 in mammals), which binds proteasomes through its N-terminus<sup>81</sup>. A second HECT E3 ubiquitin ligase is Ufd4, which also mediates proteasome binding through an N-terminal domain. Ufd4 binds to 26S proteasomes apparently via the base subunits Rpt4 and Rpt6<sup>82,83</sup>. In addition to Hul5 and Ufd4, a large number of other ligases from distinct families as well as E2 enzymes are associated with the proteasome. Recently, it has been shown that at least four E2 enzymes, Ubc1, Ubc2, Ubc4 and Ubc5, are able to interact with the proteasome in a non-competitive manner<sup>84</sup>.

## 1.2.5 The novel base component Rpn13

Several recent studies reported the identification of novel, non-essential subunits of the 19S regulatory particle, among them the yeast subunit Daq1/Rpn13<sup>31,71,85</sup>. Rpn13 from *S. cerevisiae* is an acidic 17.9 kDa protein that was identified by mass spectrometry in affinity-purified proteasomes and deletion of RPN13 results in stabilization of the N-end-rule model substrate Ub<sup>V76</sup>-Val-e<sup>AK</sup>-βGal<sup>71</sup>. Significantly, yeast Rpn13 was found to be present in highly purified proteasomes in apparently stoichiometric amounts and identified as an additional subunit of the base subcomplex<sup>33</sup>. Database search analysis suggested that Adrm1/Arml is a mammalian ortholog of the yeast Rpn13 and indeed, several publications in 2006 could support this assumption<sup>76-78,86</sup>. Adhesion regulating molecule 1 (Adrm1) was originally described as an interferon-γ-inducible, heavily glycosylated membrane protein of molecular mass 110 kDa<sup>87</sup> with a proposed role in regulation of cell adhesion, but recent studies identified Adrm1 as a 42 kDa non-glycosylated base subunit of the 19S regulatory particle<sup>78-80,88</sup>. Like other proteasome subunits, human Rpn13 is localized in both the cytosol and the nucleus and cosediments with proteasomes upon glycerol-density centrifugation<sup>79</sup>. Cellular localization by differential centrifugation could also show that a small quantity of the protein is associated with cell membranes, also similar to subcellular localisation of the 26S proteasome<sup>86,88,89</sup>. The N-terminal region of Rpn13 binds to the base subunit Rpn2/S1<sup>76,78,90,91</sup>, whereas a C-terminal region interacts with the base-associated deubiquitinating proteins, Uch37/UCHL5<sup>78-80</sup>. Uch37 catalyzes deubiquitination from the distal end of Ub chains and is proposed to function as an editing isopeptidase that rescues poorly ubiquitinated substrates from being degraded<sup>74</sup>. The functional relationship of Rpn13 with Uch37 extends beyond simply tethering it to the proteasome as Uch37 exhibits very little catalytic activity in its absence. As well as providing the determinant for recruitment to the proteasomes, the unique C-terminal tail of Uch37 serves as an inhibitory domain that keeps the uncomplexed enzyme inactive. Following association of the C-terminal domain of Rpn13/Adrm1 with the C-terminal tail of Uch37, autoinhibition is relieved, leading to increased activity<sup>76</sup>.

However, yeast Rpn13, which contains 156 amino acids, is much shorter than the 407-amino-acid human homologue. The N-terminal one-third of the human protein is 41% similar to yeast Rpn13, whereas the C-terminal two-thirds of human Rpn13, which are absent in budding yeast, has substantial similarity to its homologues from other eukaryotes, including the fission yeast *S. pombe*. Interestingly, Uch37 orthologues have been identified in *S. pombe* (Uch2) and metazoan species, but not in budding yeast, suggesting that the C-terminal domain evolved to recruit Uch37 to the proteasome<sup>78,80</sup>.

Deletion of yeast Rpn13 leads to the stabilization of some model substrates of the ubiquitin–proteasome pathway<sup>71</sup>. In frogs, the homolog of Rpn13, Xoom, is required for embryonic development<sup>92</sup>, which presumably reflects a critical role for protein degradation. Knockdown of human Rpn13 (hRpn13) by small interfering RNA (siRNA) in HEK293T cells reduces overall degradation of short–lived proteins and causes accumulation of the rapidly degraded model substrate Ub–R–GFP. Similar decreases in proteolysis are observed upon overexpression of the C–terminal half of hRpn13<sup>79</sup>. Furthermore, the knockdown markedly reduces total cellular Uch37, the amount of proteasome–bound Uch37 and decreases deubiquitinating activities of 26S proteasomes<sup>76,78</sup>. Since the amount of Uch37 mRNA is unaltered in knockdown cells, the reduction of cellular Uch37 levels is probably a posttranscriptional effect. Interestingly, expression of the C–terminal half of hRpn13 in HeLa cells appears to induce cell death<sup>77</sup> and hRpn13 is upregulated in several carcinoma cell lines and multiple types of solid tumors<sup>87,93,94</sup>. This property may involve interactions of hRpn13 with Uch37 and proteasomes, although it may have other functions that are yet to be discovered. Double knockout of the two non–essential components Rpn10 and Rpn13 in yeast shows a synthetic growth defect, indicating overlapping function of the two proteins<sup>95</sup> and recent data suggests that Rpn13p and the proteasome–associated Rpn14p are involved in the stable formation of 26S proteasome and the recognition of the ubiquitinated Gcn4p in proteasomal degradation<sup>96</sup>. Gcn4 is a typical eukaryotic transcriptional activator, which is degraded by proteasome<sup>97,98</sup> and its stability seems to be increased in the deletion strain of Rpn13.

## 1.2.6 Goals of this study

The ubiquitin-dependent proteasomal degradation represents a complex pathway and is based on the interplay of a variety of different proteasomal subunits and associated factors. Although a large number of subunits have been studied genetically, the composition and function of the 19S regulatory complex is hardly understood. The main goal of this study was the structure determination of the domain architecture of base subunit Rpn13 and the structure based functional analysis. This should be achieved by the establishment of a suitable expression system to obtain soluble recombinant Rpn13 for crystallization, the determination of the unknown phases by SeMet derivatization with subsequent structure solution and refinement. Furthermore, it was aimed to investigate interactions between Rpn13 and its binding partners Rpn2 and Uch37 as well as to reveal the underlying mechanisms of binding to monoUb (ubiquitin) and Ub-chains. In this study, pull-down assays and NMR spectroscopy should shed light on the interaction of the N-terminal domain of Rpn13 with ubiquitin and mutational analysis of interacting amino acid residues of Rpn13 should further support structural data. Additionally, NMR shift perturbation analysis should map the putative Rpn2 binding site on Rpn13. For the investigation of the interaction between the C-terminal domains of mammalian Rpn13 and Uch37, the two proteins had to be coexpressed and purified. Thereby, the design of various truncated variants of Rpn13 and Uch37, which are based on secondary structure prediction, could lead to the localization of their interaction module. A part of the results of this study was published in Schreiner et al. 2008<sup>1</sup>.

## 1.3 Results

### 1.3.1 Cloning, expression and purification of Rpn13

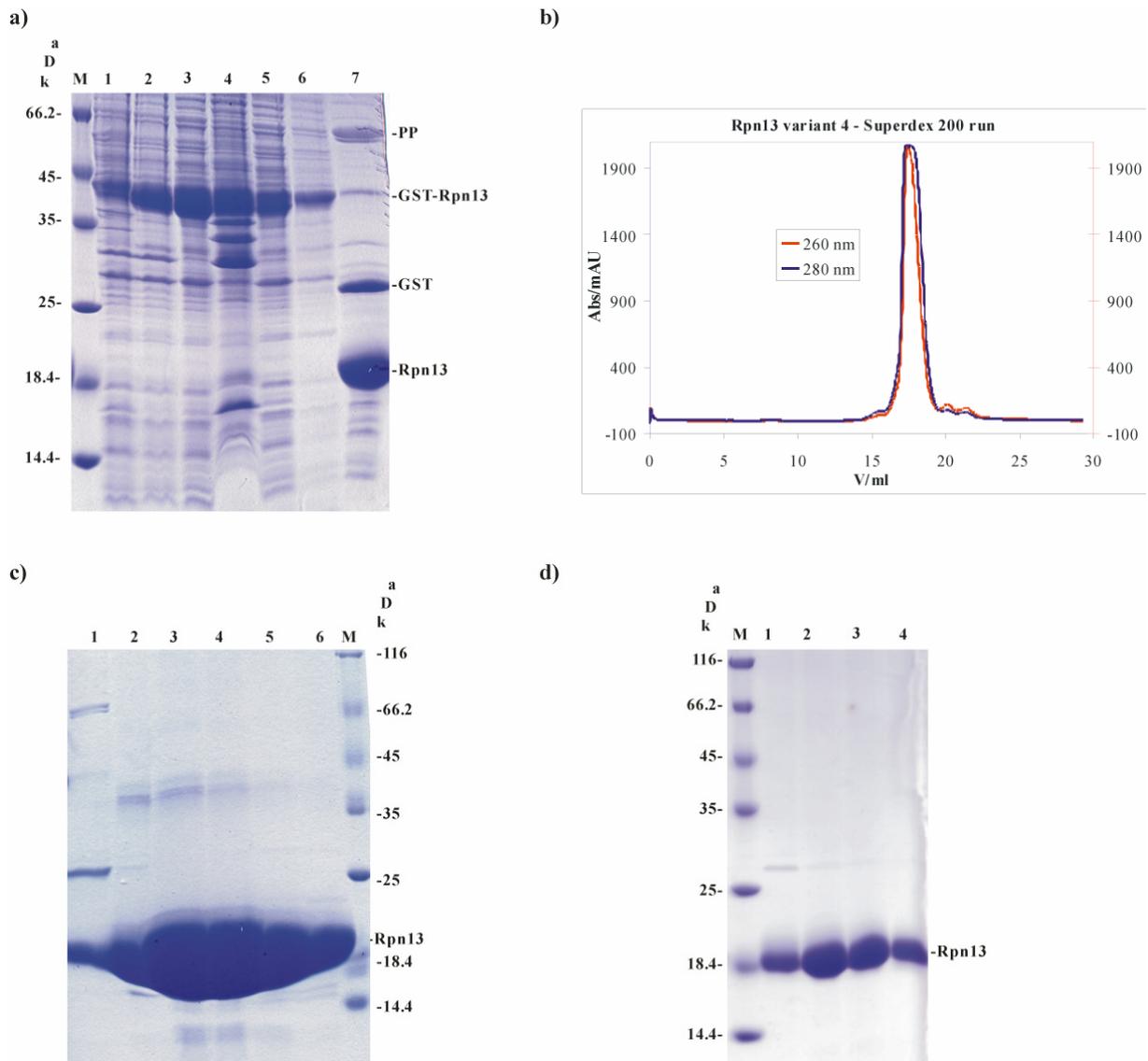
Mammalian Rpn13 could be divided in three main parts: the N-terminal domain (called Pru domain for Pleckstrin-like rceptor for ubiquitin), a linker region and the C-terminal domain. In yeast, Rpn13 consists only of the N-terminal domain. To gain more insight into the domain architecture of Rpn13, different variants of the protein from *Saccharomyces cerevisiae*, *Mus Musculus* and *Homo sapiens* were cloned, expressed and purified (variants are listed in Table 1.1) preceding further investigations like crystallization, binding assays, NMR spectroscopy and structure determination of the N-terminal Pru domain of murine Rpn13.

Variant	Amino acid residues	Organism	Expression vector	MM (kDa)	Comments
1	1–156	<i>S.c.</i>	pRSET–GST–PP	18.1	full length yeast protein, highly expressed
2	1–123	<i>S.c.</i>	pRSET–GST–PP	14.4	truncated yeast protein, highly expressed
3	1–407	<i>M.m.</i>	pRSET–GST–PP	42.2	full length mouse protein, highly expressed
4	1–150	<i>M.m.</i>	pRSET–GST–PP	16.8	highly expressed, crystals obtained
5	151–407	<i>M.m.</i>	pRSET–GST–PP	25.7	N-terminal truncated mouse protein, highly expressed
6	202–407	<i>H.s.</i>	pRSET–GST–PP	21.1	N-terminal truncated human protein, expressed
7	202–407	<i>H.s.</i>	Bicistronic vector	21.1	N-terminal truncated human protein, expressed
8	267–407	<i>H.s.</i>	pRSET–GST–PP	15.2	N-terminal truncated human protein, no expression

**Table 1.1:** Different expressed Rpn13 variants.

The cDNAs encoding Rpn13 variants were subcloned into *E. coli* expression vector pRSET–GST–PP using BamHI restriction sites (for details see Materials and Methods). The BamHI and EcoRI cleaved PCR product encoding Rpn13 variant 7 (Table 1.1) was inserted into a bicistronic vector. All constructs were tested by restriction analysis and sequencing. The expression conditions were optimized in small scale expression tests.

As the first purification step, the GST–fused protein variants were purified from bacterial lysate by affinity chromatography using Glutathione Sepharose. Recombinant protein was eluted from the glutathione beads by cleavage with PreScission Protease (PP) and the eluted protein was further purified using a two step size exclusion chromatography (see Materials and Methods). The purity was analyzed after each step by SDS–PAGE. The gels and the chromatogram of the purification of Rpn13 variant 4 are depicted in Figure 1.6.



**Figure 1.6:** Purification of Rpn13 variant 4. (a) Coomassie stained SDS–PAGE (16%) of affinity purification. 1=uninduced, 2=induced, 3=sup, 4=pellet, 5=flow through, 6=wash, 7=elution with PP (PreScission Protease). (b) Chromatogram of a Superdex 200 run. (c) After first run with Superdex 200 gel filtration column. 1=load, 2–6=fractions. (d) After gel filtration second run. 1–4=fractions.



```

.....130.....140.....150.....160.....170.....180
AA      CRKVNECLNNPMPGSLGASGSSGHELSALGGEGGLQSLNMGSMHSQLMQLIGPAGLGGL
PROF_sec  HHHHHHHH
Rel_sec  999877514787776445565666665566665542333467664544444566656655
SUB_sec  HHHHHHHH..LLLLLL..LLLLLLLLLLLLLLLLL.....LL..L.....LLLLLLLLL

.....190.....200.....210.....220.....230.....240
AA      GGLGALTGPGLASLLGSSGPPASSSSSSSRSQSAAVTPSSSTSSARATPAPSAPAAASAT
PROF_sec
Rel_sec  55555667665556677777777777777766666653356767666667777777666677
SUB_sec  LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL..LLLLLLLLLLLLLLLLLLLLL

.....250.....260.....270.....280.....290.....300
AA      SPSPAPSSNGTSTAASPTQPIQLSDLQSILATMNVPAGPGGSQQVDLASVLTPEIMAPI
PROF_sec
Rel_sec  77877776666777778877766531000012367777777665553212363111201
SUB_sec  LLLLLLLLLLLLLLLLLLLLLLLLLL.....LLLLLLLLLLLLL.....L.....

.....310.....320.....330.....340.....350.....360
AA      LANADVQERLLPYLPSGESLPQTADEIQNTLTSPQFQQALGMFSAALASGQLGFLMCQIG
PROF_sec  HHHHHHHH          HHHHHHHH   HHHHHHHHHHHHHH   HHHHH
Rel_sec  477356664420178766765406788764157105677878888753278631555215
SUB_sec  .LL.HHHH.....LLLLLLL..HHHHH..LL..HHHHHHHHHHH..LL..HHH..L

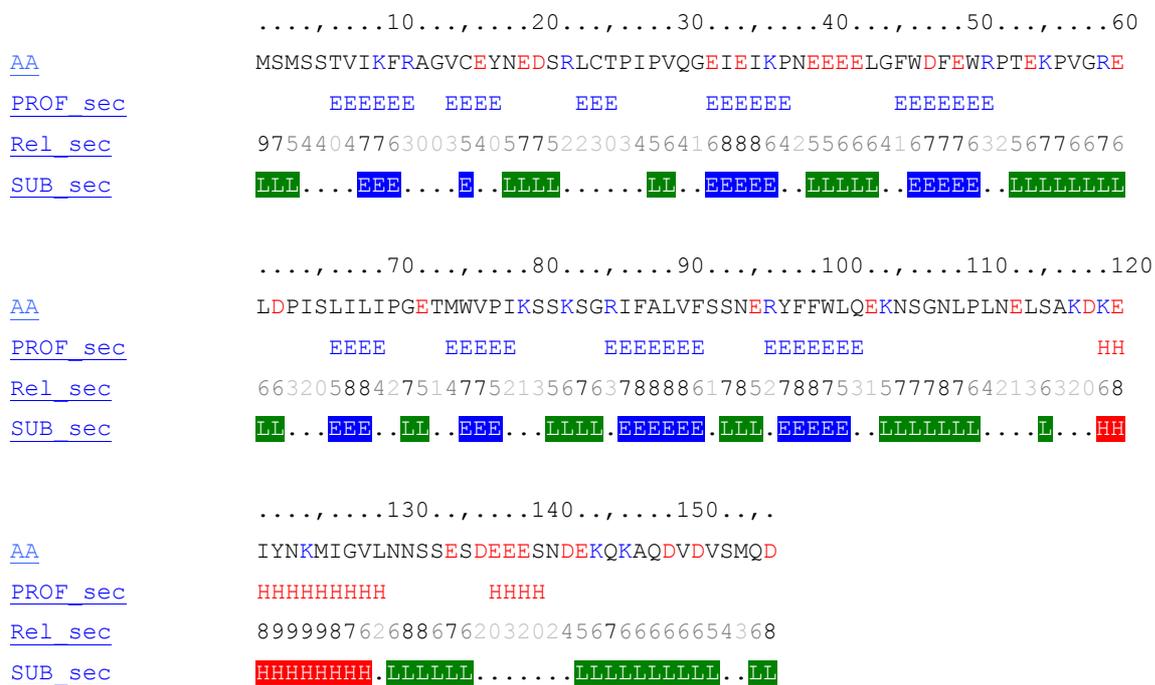
.....370.....380.....390.....400.....
AA      LPAEAVEAANKGDVEAFAKAMQNNAKSDPKEGDTKDKKDEEEDMSLD
PROF_sec  HHH          HHHHHHHHHH          HHHH
Rel_sec  66010121244636889998640256774234533354534304679
SUB_sec  LL.....L.HHHHHHHH..LLL.....L..L.L.....LL

```

**Figure 1.7:** Secondary structure prediction of *mRpn13*. *PROFsec*: predicted secondary structure in three states: helix H, sheet E, rest L. *RELsec*: reliability index for secondary structure prediction (0 = low to 9 = high). *SUBsec*: subset of the prediction, for all residues with an expected average accuracy >82%.

The secondary structure prediction of *mRpn13* reflected the domain architecture of mammalian *Rpn13* (Figure 1.7). The *PROFsec* prediction showed clearly the defined secondary structure of the conserved N-terminal Pru domain (21–130) which was determined by X-ray crystallography in this study (see section 1.3.3) and the canonical order of its structural components, the seven  $\beta$ -strands followed by the C-terminal  $\alpha$ -helix. In contrast, the C-terminal domain (301–407) exhibited solely helical structure and as expected from low complexity calculations, the serine-rich linker domain (131–300) lacks secondary structure elements. Prediction of *scRpn13s* secondary structure demonstrated the same pattern of seven

$\beta$ -strands and  $\alpha$ -helix for the N-terminal Pru domain in yeast (1–130) (Figure 1.8). In contrast to yeast Rpn13, the mammalian homolog contains a N-terminal extension of 20 amino acids, where secondary structure elements are also missing.



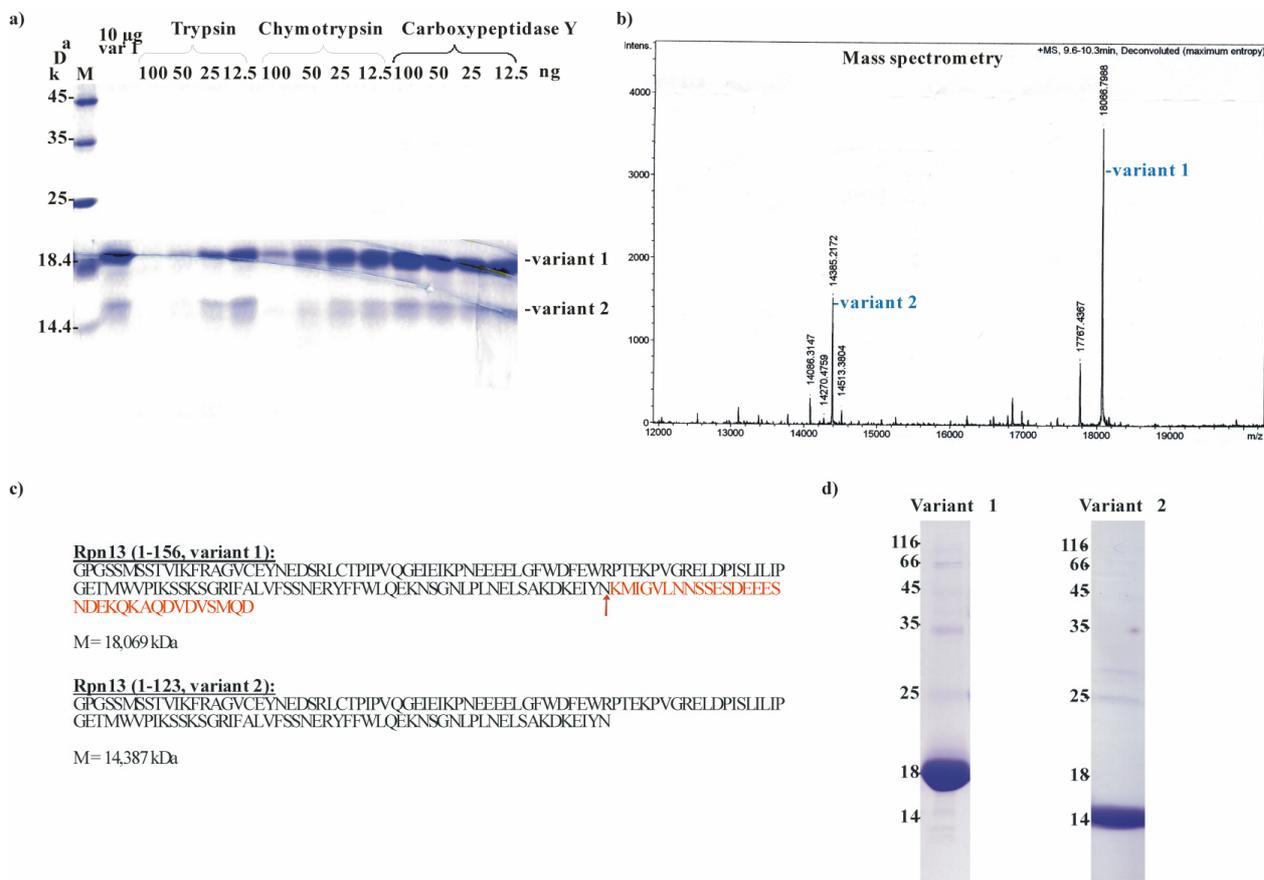
**Figure 1.8:** Secondary structure prediction of *scRpn13*. *PROFsec*: predicted secondary structure in three states: helix H, sheet E, rest L. *RELsec*: reliability index for secondary structure prediction (0 = low to 9 = high). *SUBsec*: subset of the prediction, for all residues with an expected average accuracy >82%.

### 1.3.2.2 Limited proteolysis of yeast Rpn13

In this study limited proteolysis was used as a tool for the investigation of the protein structure and domain architecture of Rpn13. In order to retrieve structural information by localization of stable domains through proteolytic digestion it is necessary to limit the enzymatic reaction for example in either concentration of the protease or incubation time.

In this experiment Rpn13 variant 1 (Table 1.1) from *S. cerevisiae* was digested with the three proteases Trypsin, Chymotrypsin and Carboxypeptidase Y. Trypsin cleaves peptide bonds C-terminal to basic amino acids, lysine and arginine, whereas Chymotrypsin prefers bulky nonpolar and aromatic side chains like phenylalanine, tyrosine and tryptophane. Carboxypeptidase Y exhibits broader amino acid specificity, and is able to release every amino acid including proline from the carboxyl-terminus of peptides, although glycine and aspartic acid are released much more slowly than other residues. The influence of the tertiary

structure is expected to be more pronounced for proteases with broader specificity such as Carboxypeptidase Y. Rpn13 variant 1 was incubated with different concentrations of each protease and the results of the digestion were monitored by SDS–PAGE analysis (Figure 1.9a).



**Figure 1.9:** Limited proteolysis of Rpn13 variant 1 from *S. cerevisiae*. **(a)** Coomassie stained SDS–PAGE (16%) of variant 1 digested with the proteases Trypsin, Chymotrypsin and Carboxypeptidase Y in different concentrations. 10  $\mu$ g of Rpn13 was used in each lane. A stable fragment at 14 kDa (variant 2) occurred **(b)** Mass spectrometrical analysis of the proteolytic digestion of variant 1. The digested fragment exhibits the size of 14385 Da in comparison to the 18066 Dalton of the unprocessed protein. **(c)** Sequence of full length protein Rpn13 variant 1 and of the shorter variant 2 (aa 1–123) found by limited proteolysis. The cleavage site and the deleted sequence are marked red. **(d)** Coomassie stained SDS–Page of purified variant 1 and 2.

SDS–PAGE analysis exhibited a processed fragment of 14 kDa, which seemed to be quite resistant to proteolytic treatment and which was also present as a degradation product of the untreated protein. The experiment was repeated using 12.5 ng of Trypsin with 10  $\mu$ g of variant 1 and the sample was analyzed by mass spectrometry as depicted in Figure 1.9b. The

determination of the exact molecular mass of 14385 Da for the fragment allowed localization of the cleavage site, which is C-terminal of asparagine 123. This result led to the design of a new variant 2 for crystallization (Figure 1.9c and d).

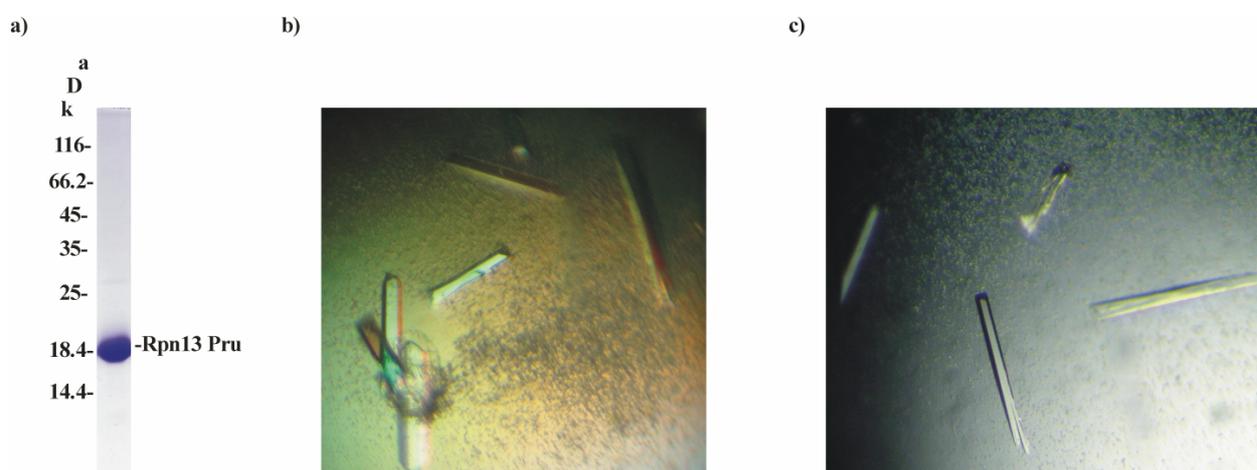
Both variants were often present as double bands arising from the fact that the four N-terminal residues GPGS, which were residual from cleavage with PreScission Protease and the BamHI restriction site, were removed by endogenous proteases during purification procedure (see less intensive peaks in Figure 1.9b). Variant 2 was cloned and purified according to Section 1.3.1 and also subjected to crystallization trials. Figure 1.9d shows SDS-PAGE of purified variant 1 and variant 2 in comparison, but none of the proteins crystallized. The crystal structure of the mammalian homologue, which was solved in this study (see Section 1.3.3), and the sequence alignment (Figure 1.13) propose that variant 2 is truncated in the middle of the C-terminal  $\alpha$ -helix of the Pru domain, which might explain problems in crystallization. Additionally, NMR data suggested low structural order of this helix as it was absent in NMR structural analysis. This low structural order could result in the observed accessibility of the  $\alpha$ -helix for protease treatment.

### **1.3.3 Crystal structure of the Pru domain of Rpn13**

#### **1.3.3.1 Crystallization and structure determination of the N-terminal Pru domain of Rpn13**

It has been recently shown that Rpn13 is a subunit of the 19S regulatory particle, being involved in contact of the 19S regulator with several other proteins, such as Uch37, Rpn2 as well as mono- and polyUb<sup>1,76-78,86,99</sup>. The major aim of this study was to crystallize Rpn13 to gain new insights into its domain architecture and to reveal its structural and functional role in the 19S regulatory subunit. Different variants listed in Table 1.1 were expressed and purified as described (see Section 1.3.1). The purified proteins were subjected to crystallization and the crystal structure of Rpn13 variant 4, the N-terminal domain of murine Arm1/Rpn13 (residues 1–150, hereafter termed Pru domain) was determined by X-ray crystallography. As described above, the Pru domain was fused to a PreScission Protease cleavable GST-tag, overexpressed in a soluble form in *E. coli* and purified by different chromatography steps to more than 95% purity (Figure 1.10a). The typical yield of purified soluble Rpn13 Pru from 3 l of culture was approximately 4 mg. Size exclusion chromatography revealed that this domain forms monomers, which was consistent with the structural results.

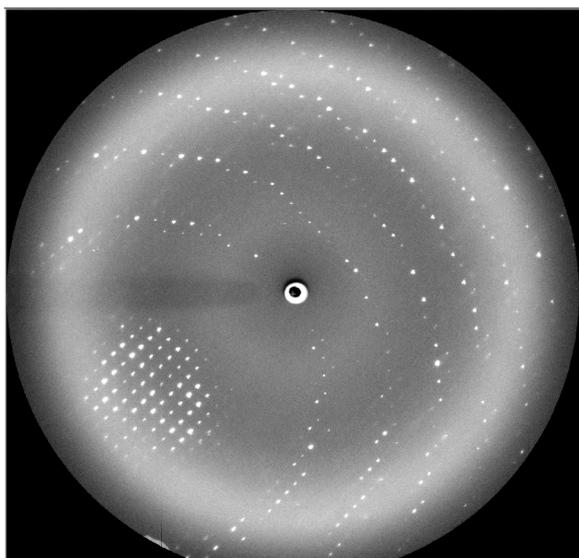
Initial crystallization trials were performed with the purified mRpn13 Pru domain by hanging drop vapour diffusion method using commercial screening solutions (see Materials and Methods) and crystals appeared at 20°C within two days. After optimization the protein was crystallized using equal volumes of protein (5 mg/ml) and reservoir solution (15% (w/v) PEG 4000, 200 mM NaOAc, 1 mM DTT, 100 mM Tris–HCl, pH 8.5) and grew to the size of 500x 200x 200  $\mu\text{m}^3$  (Figure 1.10b). Addition of DTT was critical for growth and diffraction quality of the crystals. Before exposure to X-rays, crystals were soaked in a solution of mother liquor including 15% PEG 400 and subsequently frozen in a stream of cold nitrogen gas at 100 K. Crystals were then measured using synchrotron radiation at the BW6 beamline at the DESY–centre in Hamburg and native data were collected (Table 1.2) using a MarCCD detector. Crystals of Rpn13 Pru were well ordered and diffracted to 1.7 Å resolution (Figure 1.11).



**Figure 1.10:** Crystallization of Rpn13 Pru domain (variant 4) from *M. musculus*. (a) SDS–PAGE of purified Pru domain. (b) Crystals of the Rpn13 Pru domain. (c) Crystals of the SeMet derivative of variant 4.

Data were processed with DENZO and SCALEPACK. The space group of Rpn13 Pru was  $P2_12_12_1$  with the unit cell dimensions  $a=42.4 \text{ \AA}$ ,  $b=56.2 \text{ \AA}$ ,  $c=63.2 \text{ \AA}$ .

Phase determination was performed by crystal structure analysis of the Rpn13 Pru–SeMet–derivative. For selenomethionine labelling the Pru domain was overexpressed in the methionine auxotroph *E. coli* strain B834(DE3) and cells were grown in New Minimal Medium (see Materials and Methods). Purification and crystallization of selenomethionine–labelled protein was performed according to the same procedure as for native protein and relevance of DTT for crystallization was even more pronounced (Figure 1.10c).



**Figure 1.11:** *X-ray diffraction image of Rpn13 Pru crystals.*

SAD data were collected at 2.7 Å resolution using anomalous diffraction experiments (absorption wavelength scan of Se identified wavelength  $\lambda=0.9793$  Å, see Table 1.2). Two Se-sites were localized using SHELXD<sup>100</sup>. Subsequent phasing with MLPHARE (the CCP4 suite: programs for protein crystallography) and solvent flattening with DM<sup>101</sup> resulted in an interpretable electron density map, which was traced and phase extended to the native data set at 1.7 Å resolution using ARP-WARP<sup>102</sup>. The model has been completed using the interactive three-dimensional graphic program MAIN<sup>103</sup> and refined with REFMAC5<sup>104</sup>. Temperature factors were optimized with restraints between bonded atoms and between non-crystallographic symmetry related atoms. The finalized structure revealed well defined electron density for the core domain of Rpn13 Pru domain, except for the first 21 N- and the last 20 C-terminal residues of the molecule which were structurally disordered (Figure 1.14). In successive rounds of model building and refinement, the model was completed and included 81 water molecules. Temperature factors were refined with restraints between bonded atoms and between non-crystallographic symmetry related atoms yielding R-values of  $R_{\text{work}} = 19.9\%$  and  $R_{\text{free}} = 22.4\%$  (Table 1.2).

	Rpn13 Pru	Rpn13 Pru SeMet
<i>Crystal parameters</i>		
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell constants	a=42.4 Å; b=56.2 Å; c=63.2 Å	a=42.7 Å; b=56.4 Å; c=64.2 Å
Heavy metal	–	2
<i>Data collection</i>		
Beamline	BW6, DESY	BW6, DESY
Wavelength (Å)	1.05	0.9793
Resolution range (Å) <sup>a</sup>	99–1.7 (1.70–1.67)	99–2.62 (2.7–2.66)
No. observations	1007679	227878
No. unique reflections <sup>c</sup>	25165	8927
Completeness (%) <sup>b</sup>	99.5 (99.9)	98.9 (100)
R <sub>merge</sub> (%) <sup>b, c</sup>	5.1 (27.2)	10.0 (35.3)
I/σ (I) <sup>b</sup>	19.1 (2.9)	32.6 (11.4)
<i>Refinement (REFMAC5)</i>		
Resolution range (Å)	15–1.7	
No. reflections working set	16267	
No. reflections test set	813	
No. non hydrogen	973	
Solvent water	106	
R <sub>work</sub> /R <sub>free</sub> (%) <sup>d</sup>	19.9 / 22.4	
R <sub>work</sub> /R <sub>free</sub> (2.74–1.70 Å) (%) <sup>d</sup>	24.5 / 27.7	
rmsd bond lengths (Å) / (°) <sup>e</sup>	0.012 / 1.75	
Average B-factor (Å <sup>2</sup> )	30.67	
Ramachandran Plot (%) <sup>f</sup>	92.7 / 7.3 / 0.0	

**Table 1.2:** Data collection and refinement statistics.

<sup>a</sup>The values in parentheses of resolution range, completeness, R<sub>merge</sub> and I/σ (I) correspond to the last resolution shell.

<sup>b</sup>Friedel pairs were treated as different reflections.

<sup>c</sup> $R_{merge}(I) = \frac{\sum_{hkl} \sum_j |I(hkl)_j - I(hkl)|}{\sum_{hkl} I_{hkl}}$ , where  $I(hkl)_j$  is the  $j$ th measurement of the intensity of reflection  $hkl$  and  $\langle I(hkl) \rangle$  is the average intensity.

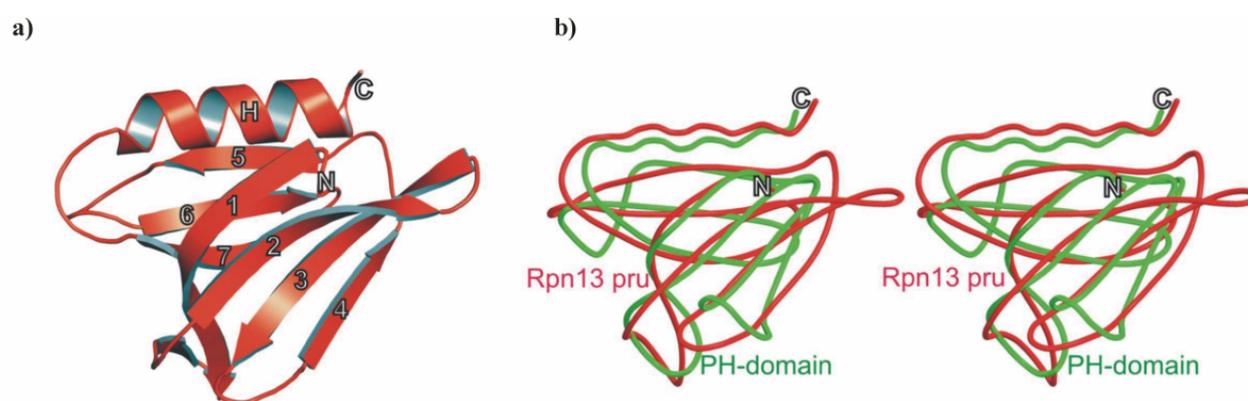
<sup>d</sup> $R = \frac{\sum_{hkl} ||F_{obs}| - |F_{calc}||}{\sum_{hkl} |F_{obs}|}$ , where  $R_{free}^{105}$  is calculated without a sigma cutoff for a randomly chosen 5% of reflections, which were not used for structure refinement, and  $R_{work}$  is calculated for the remaining reflections.

<sup>e</sup>Deviations from ideal bond lengths/angles.

<sup>f</sup>Number of residues in favoured region / allowed region / outlier region.

### 1.3.3.2 Rpn13 Pru adopts Pleckstrin Homology (PH) architecture

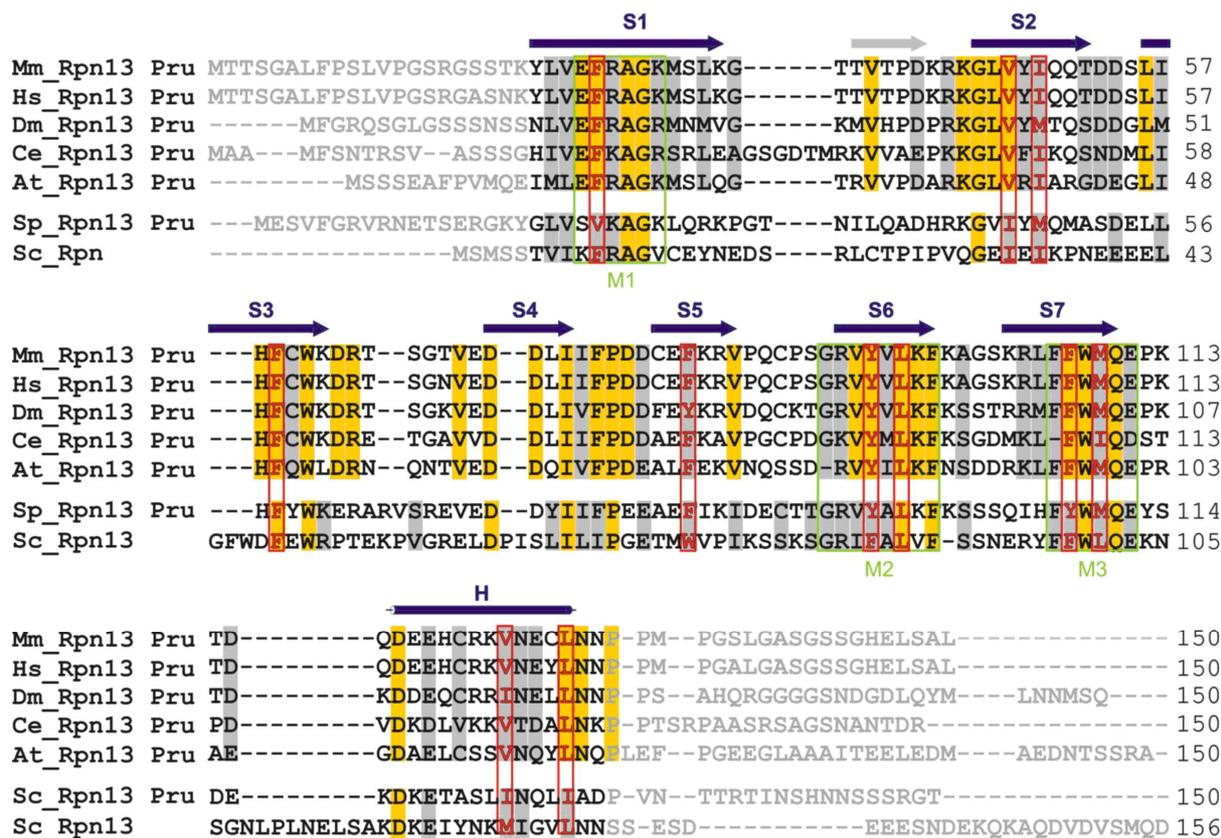
Primary sequence alignment could not identify any sequence similarities of Rpn13 Pru with other proteins, and no predictions of domain structure of Rpn13 could be made by using pattern and profile search programs. Therefore, a new folding pattern of this protein class was expected. Surprisingly, the crystal structure of Rpn13 Pru revealed that the N-terminal domain of Rpn13 has the topology typical for the Pleckstrin Homology (PH) domain fold (Figure 1.12a, b). Thus, we named this domain Rpn13 Pru, for Pleckstrin-like receptor for Ub. The determination of the structure allowed comparison with the secondary structure prediction. Although pattern and profile search program did not identify the fold of the N-terminal domain, the PROFsec prediction in Section 1.3.2.1 detected all secondary structure elements of the PH fold pretty well.



**Figure 1.12:** Crystal structure of murine Rpn13 Pru (variant 4) revealed typical Pleckstrin homology (PH) domain fold. (a) Ribbon drawings of Rpn13 Pru. The PH-fold consisting of seven-stranded  $\beta$ -sandwich structure (1–7) capped by the C-terminal  $\alpha$ -helix (H), indicated in white. (b) Stereoview of the structural alignment of Rpn13 Pru (red) and the PH domain (green) from Pleckstrin (PDB accession code 1PLS)<sup>106</sup>.

PH-domains are small modules (100–120 amino acids) found in a multitude of intracellular proteins with very widespread functions<sup>106-110</sup>. The N-terminal Pru domain of Rpn13 represents the first identified example of a PH fold among subunits of the 26S proteasome. Being very divergent at their sequence level, all PH domains share a common  $\beta$ -sandwich topology. The PH fold of Rpn13 Pru is composed of a four-stranded twisted antiparallel  $\beta$ -sheet ( $\beta_{1-4}$ : residues 22–34, 45–52, 56–62, 71–74) which packs almost orthogonally against a second triple stranded  $\beta$ -sheet ( $\beta_{5-7}$ : residues 80–85, 92–98, 103–110) (Figure 1.12 and 1.13). Rpn13 Pru also forms a hydrophobic core which is typical for PH domains. All conserved hydrophobic residues forming this core, such as Phe26 (S1), Val47 (S2), Ile49 (S2), Phe50

(S3), Phe82 (S5), Tyr94 (S6), Leu96 (S6), Phe107 (S7) and Met109 (S7) are located within secondary structure elements ( $\beta$ -sheets) and positioned towards the barrel of the core (residues are indicated by red boxes in Figure 1.13). One end of the  $\beta$ -sandwich is capped by a long C-terminal amphipathic  $\alpha$ -helix (residues 117–128), which is stabilized by interactions between residues Val124 and Leu128, whereas the other corner of the hydrophobic core is closed by three loops formed by the residues located between strands S1/S2, S3/S4 and S6/S7 (Figure 1.12).



**Figure 1.13:** Sequence alignment of Rpn13 Pru homologues from *Mus musculus* (Mm), *Homo sapiens* (Hs), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), *Arabidopsis thaliana* (At), *Schizosaccharomyces pombe* (Sp) and *Saccharomyces cerevisiae* (Sc). Identical residues of Rpn13 Pru are highlighted against a yellow background, conserved residues against a grey background. Sequence stretches given in grey are disordered in the crystal structure. Residues involved in formation of the hydrophobic core of the PH fold are conserved among all eukaryotes and indicated by red boxes. Motifs (M1, M2 and M3) exhibiting the highest degree of identity among Rpn13 Pru species are marked by green boxes. Secondary structure elements of Rpn13 Pru contributing to the topology of the PH fold are indicated in blue, remaining secondary structure elements in grey.

These loops represent variable (in primary sequence and in fold) regions among PH domains. Although the PH domain family is very divergent at the sequence level, the overall architecture of the molecules is well conserved, with major differences found only in the composition of the loops. These differences might have functional significance, since variable loops were shown to play a central role in specific and selective interactions of PH domains with ligands and proteins<sup>110-114</sup>. Structural superposition of Rpn13 Pru with Pleckstrin revealed high consensus of the backbone of the  $\beta$ -sandwich and the  $\alpha$ -helix (r.m.s. deviation 1.7 Å)<sup>109,110</sup>, but significant differences between the loops (Figure 1.12b).

We performed a structural similarity search in the protein Data Bank using the DALI server (<http://www.ebi.ac.uk/dali/>) to identify PH folds which share similar conformation of the loops. Two proteins were found to share the highest structural homology to Rpn13 Pru: the PH domain of Enabled/VASP Homology 1 protein (EVH1), which is binding to proline-rich motifs (PDB accession code 1EVH, Z-score of 11.3, average r.m.s. deviation calculated over 111 equivalent C $\alpha$ -positions, 2.5 Å)<sup>111</sup> and DAPP1, which specifically binds to phosphatidylinositolphosphates (PIPs) in cell membranes (PDB accession code 1FAO, Z-score of 10.6, average r.m.s. deviation calculated over 95 equivalent C $\alpha$ -positions, 2.5 Å)<sup>113</sup>. Interestingly, despite their structural similarity, EVH1 and DAPP1 have entirely different functions and bind their specific ligands at different sites. This indicates that the PH domain fold is a robust, but at the same time functionally plastic structure which has enough potential versatility to bind a wide range of ligands. Therefore, results of the DALI search alone do not allow any conclusions about function and ligand specificity, as well as about location of putative binding sites in Rpn13 Pru.

### 1.3.3.3 Rpn13 Pru is conserved in eukaryotes

Proteins of the Arm1/Rpn13 family are highly conserved among mammals (human and mouse share 97% sequence identity) with less significant sequence similarities between high and low eukaryotes. Figure 1.13 shows the multiple sequence alignment of Rpn13 from different species. Having solved the crystal structure of the N-terminal Pru domain of murine Rpn13 and considering significant homology of the N-terminal domain among higher eukaryotes (about one third of the residues forming the  $\beta$ -sheets are strictly conserved), we conclude that all eukaryotic Rpn13 homologues possess a PH domain fold at their N-terminus. Flanking regions of the PH core structure are not conserved in length and charge among species (the first 21 N- and the last 20 C-terminal residues of the molecule were structurally disordered). Although the PH fold of Rpn13 from fungi (*S. cerevisiae* and *S. pombe*) shows lower

sequence similarity to murine Rpn13 Pru, residues involved in formation and stabilization of the hydrophobic core are absolutely conserved, providing another indication that the N-terminal domains of Rpn13 from low eukaryotes also contain a PH fold (Figure 1.13). NMR data from cooperation partners revealed that the yeast homologue follows similar topology<sup>99</sup>. Rpn13 from yeast represents the shortest member of the family, since it has only 5 residues preceding the N-terminal PH domain and completely lacks the linker and the C-terminal domain, which was shown to be essential for Uch37 binding. Since yeast does not contain Uch37, there is apparently no need for the C-terminal domain. It has been shown that the first 132 amino acids from human Rpn13 are sufficient for the protein to form a stable complex with subunit Rpn2 of the 19S regulatory particle<sup>78</sup>, which together with structural data would indicate that two distinct domains of Rpn13 might act independently of each other. Rpn13 Pru is able to bind simultaneously to mono- or polyUb chains<sup>99</sup>, as well as to Rpn2 of the 19S regulatory particle, thus acting as a mediator for Ub attachment to the proteasome. The fact that yeast lacks the C-terminal part of Rpn13 indicates that it is the sole function of the PH fold to bind both Rpn2 and Ub, leaving the C-terminus to perform other functions such as binding of deubiquitinating enzymes (DUBs).

### **1.3.4 Ubiquitin docking at the proteasome via the PH domain interaction of Rpn13 Pru**

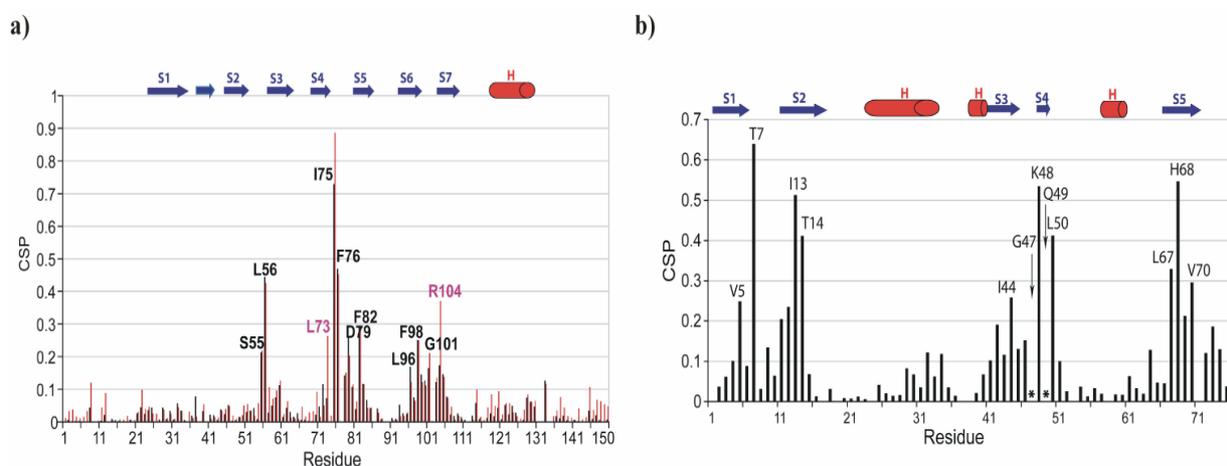
#### **1.3.4.1 Investigation of the Pru interaction with ubiquitin by chemical shift perturbation analysis.**

Functional studies from cooperation partners indicated that the N-terminal Pru domain of Rpn13 is able to specifically bind to ubiquitin and therefore might represent a novel ubiquitin receptor in the 19S regulatory subunit of the proteasome. For this reason, the domain was named Pleckstrin-like ubiquitin receptor (Pru).

In cooperation with the group of Walters (University of Minnesota, USA) NMR spectroscopy was used to confirm an interaction between Rpn13 and ubiquitin and to determine how hRpn13 binds ubiquitin. Towards this aim, we performed optimized triple resonance experiments to assign chemical shift values to hRpn13 Pru's backbone (N,H<sup>N</sup>,C $\alpha$ ,C') and C $\beta$  atoms. Chemical shift assignments were made by using 3D HNCA/HNCOCA and HNCACB experiments on 0.6 mM <sup>15</sup>N-, <sup>13</sup>C- and 70% <sup>2</sup>H-labelled hRpn13 Pru. A <sup>15</sup>N dispersed NOESY spectrum (200 ms mixing time) acquired on <sup>15</sup>N-, 50% <sup>2</sup>H-labelled human Rpn13

Pru confirmed that human Rpn13 Pru in solution has similar topology and architecture as murine Rpn13 Pru in the crystal. Moreover, the first 21 and last 20 amino acids of the hRpn13 Pru are similarly disordered in solution. This spectrum was used to aid in the identification of intermolecular NOE (nuclear Overhauser enhancements) interactions between hRpn13 Pru and Ub, as it was compared to an  $^{15}\text{N}$  dispersed NOESY experiment acquired under identical conditions on  $^{15}\text{N}$ ,  $^{13}\text{C}$ - and 70%  $^2\text{H}$ -labelled hRpn13 Pru mixed with an equimolar ratio of unlabelled Ub. For this experiment, no  $^{13}\text{C}$  or  $^2\text{H}$  decoupling was implemented to enable the ready identification of intermolecular NOE interactions (Figure 1.15).

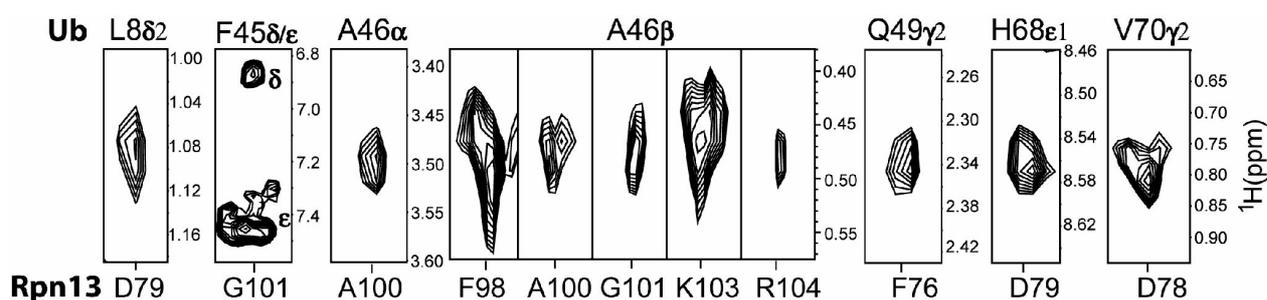
To determine how Rpn13 binds ubiquitin, the chemical shift perturbation analysis and the NOESY experiments were used to directly detect intermolecular contacts. For the chemical shift perturbation analysis, three sets of experiments were performed. In particular unlabelled monoUb or tetraUb was added to  $^{15}\text{N}$  labelled hRpn13 Pru (Figure 1.14a) or alternatively unlabelled hRpn13 Pru was added to  $^{15}\text{N}$  labelled monoUb (Figure 1.14b). In each case, the  $^{15}\text{N}$  labelled protein was observed by  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC experiments as it bound its unlabelled binding partner. The amide nitrogen and hydrogen chemical shift changes for each residue were mapped according to Equation 1 (see Materials and Methods).



**Figure 1.14:** Conserved loops of Rpn13 Pru bind ubiquitin's  $\beta$ -strand face. (a) Chemical shift perturbation analysis revealed residues within Rpn13 Pru that bind mono- (black) and K48-linked tetraubiquitin (red). For this analysis, the amide chemical shift values of Rpn13 alone were compared to values with equimolar quantities of unlabelled ubiquitin (black) or K48-linked tetraubiquitin (red) present. Residues that were significantly affected by mono- or tetraubiquitin are labelled. L73 and R104, which exhibited significant shifting only when Rpn13 binds to K48-linked tetraubiquitin, are labelled in purple. (b) Analysis of the amide chemical shift values of free and Rpn13-bound ubiquitin were compared at equimolar concentration, as described in (a). Resonances that shifted significantly are labelled and those that undergo severe broadening are marked with asterisks and labelled with an arrow.

This information was complemented with twelve intermolecular NOE interactions to define the interaction surface precisely (see Materials and Methods).

The chemical shift perturbation analysis and intermolecular NOEs indicated that Rpn13 binds ubiquitin using a surface formed by the S2–S3, S4–S5, and S6–S7 loops. More specifically, S55, L56, I75, F76, D79, F82, L96, F98 and G101 all exhibited significant amide chemical shift perturbations as Rpn13 bound ubiquitin (Figure 1.14a, black). In addition, explicit intermolecular NOE interactions were detected between F76, D78, D79, F98, A100, G101, K103, and R104 of hRpn13 and ubiquitin (Figure 1.15).



**Figure 1.15:** Representative interactions identified between Rpn13 Pru and ubiquitin. Each panel contains a selected region of an  $^{15}\text{N}$  dispersed NOESY experiment recorded on  $^{15}\text{N}$ ,  $^{13}\text{C}$  and 70%  $^2\text{H}$ -labelled human Rpn13 mixed with equimolar quantities of unlabelled ubiquitin. All of the resonances displayed in this panel could be unambiguously assigned as intermolecular NOE interactions with ubiquitin. Ubiquitin and Rpn13 Pru assignments are provided at the top and bottom of the expanded regions, respectively.

The amino acids within these three loops form a contiguous surface with the sequence  $^{76}\text{FPD}^{78}$  located at the surface's centre. This motif is strictly conserved in higher eukaryotic Rpn13 proteins (Figure 1.13). Furthermore, in all Rpn13 species, the neighbouring residue of D79 is either an aspartic or glutamic acid. Interestingly, the residues within these loops share little sequence homology between mammalian and yeast Rpn13. P77, however, is strictly conserved. It is worth noting that prolines are excluded from amide chemical shift perturbation analysis; however, P77 introduces a turn in the loop connecting S4 and S5, and thereby plays an important role in defining the architecture of this region (Figure 1.16).

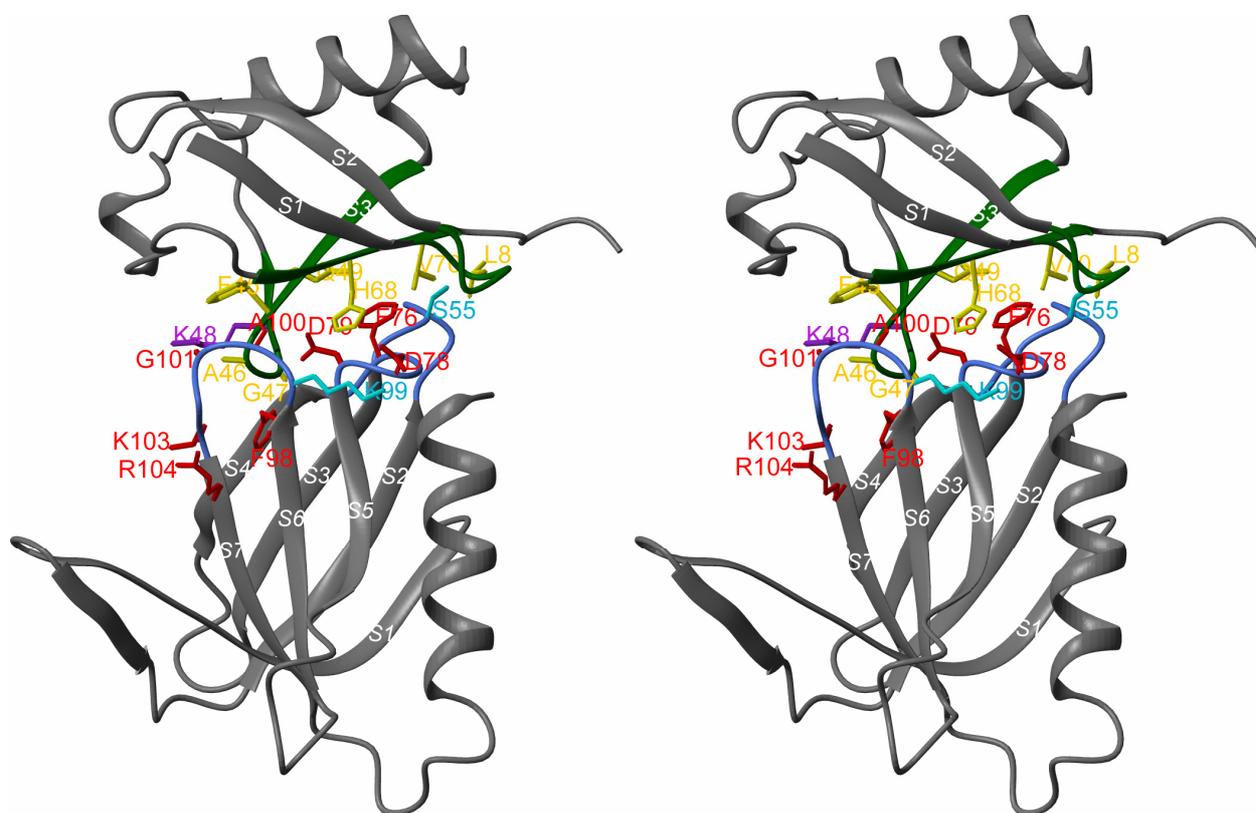
The Ub-binding surface of Rpn13 is rich in hydrophobic residues and therefore was expected to bind to Ub's hydrophobic patch, a  $\beta$ -strand surface that includes L8, I44, and V70 and is involved in binding most Ub receptors<sup>115-117</sup> including proteasome component Rpn10/S5a<sup>118</sup>. This hypothesis was validated by explicit intermolecular NOE interactions between Rpn13 Pru and L8, F45, A46, Q49, H68, and V70 of Ub (Figure 1.15). K48 of Ub is proximal to A46

and Q49 and exhibited significant chemical shift changes upon hRpn13 addition (Figure 1.14b). Altogether, these data suggest it to be within the Rpn13-binding surface. Ub chains formed by K48 linkages serve as effective signals for proteasomal degradation<sup>119,120</sup>, this finding was pursued further, and described in Section 1.3.4.5.

### 1.3.4.2 Novel Ub binding mechanism of Rpn13 Pru

Only five amino acid residues differ between the N-terminal Ub-binding domain of human and murine Rpn13 Pru (Figure 1.13). None of these are implicated by the NMR data as binding Ub and three of these, A18, N20, and A136, are located in the structurally disordered regions. As described above, the <sup>15</sup>N dispersed NOESY experiment confirmed that hRpn13 Pru in solution is similar to the crystal structure of mRpn13 in terms of topology and architecture, as well as in the disordered state of the N-terminal 21 and C-terminal 20 amino acids. Therefore, the mRpn13 Pru crystal structure (Figure 1.12a) was used as a starting structure along with the NMR data on hRpn13 to determine the structure of the mRpn13 Pru:Ub complex (Figure 1.16). mRpn13 Pru:Ub complexes were generated by using HADDOCK1.3 (*High Ambiguity Driven protein-protein DOCKing*)<sup>121</sup> in combination with CNS<sup>122</sup> as described in Materials and Methods. Intermolecular NOE interactions between hRpn13 Pru and Ub were obtained by comparing the <sup>15</sup>N-dispersed NOESY experiments recorded on hRpn13 Pru alone with that of the hRpn13 Pru:Ub complex. In addition, chemical shift assignment of Ub sidechain atoms was obtained by following the assigned resonances of the free protein upon addition of increasing molar ratios of unlabelled hRpn13 Pru. All atom pairs exhibiting an intermolecular NOE interaction were constrained to be within 1.8–6.0 Å of each other. In addition, the constraints from the hydrogen bonds published for Ub (PDB code 1D3Z)<sup>123</sup> were used to maintain its fold, as the protein does not undergo any gross structural change upon binding. For the first step of rigid-body energy minimization, 1000 structures were generated. 200 structures having the lowest energy from the rigid-body docking were subjected to semi-flexible simulated annealing in torsion angle space followed by refinement in explicit water. During semiflexible simulated annealing, atoms at the interface were allowed to move but constrained by the AIRs and unambiguous NOE-derived distance constraints. After water refinement, the resulting structures were sorted according to intermolecular energy and clustered using 1.5 Å cut-off criteria. This treatment resulted in an ensemble of 200 structures with the average RMSD of  $0.94 \pm 0.31$  Å for all backbone atoms and  $0.75 \pm 0.23$  Å for the backbone atoms at the interface. The ten lowest energy structures in the lowest energy cluster were evaluated according to the following criteria: all

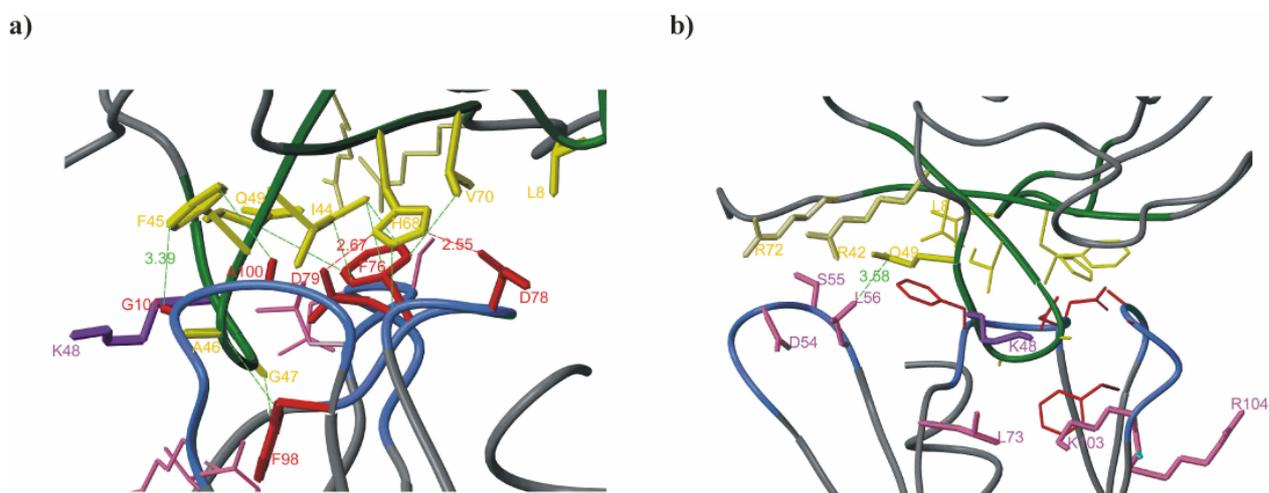
intermolecular hydrophobic contacts between heavy atoms being less than 3.9 Å, intermolecular hydrogen bond distances between proton–acceptor and donor–acceptor pair being within 2.7 Å and 3.35 Å, respectively, and minimum angles of 90° for donor–hydrogen acceptor, hydrogen–acceptor–acceptor antecedent, donor–acceptor–acceptor antecedent atoms of hydrogen bonds<sup>124</sup>.



**Figure 1.16:** Structure of Rpn13 Pru:Ub complex defines a novel Ub–binding motif. Stereo representation of the mRpn13 Pru:Ub complex oriented with Ub at the top. At the interaction surface secondary structural elements of Ub and Rpn13 Pru are displayed in green and blue, respectively. Residues at the contact surface with intermolecular NOEs are in yellow (Ub) or red (Rpn13 Pru) whereas those suggested to be at the contact surface only by the NMR titration experiments are displayed in purple (Ub) or cyan (Rpn13 Pru).

The structure of the mRpn13 Pru:Ub complex reveals a novel Ub binding mode in which residues of the S2–S3, S4–S5, and S6–S7 loops capture Ub (Figure 1.16). Neither Rpn13 Pru nor Ub undergoes structural rearrangements to bind each other: the r.m.s. deviations between the free and complexed state of Rpn13 Pru and Ub are 0.91 and 0.75 Å for backbone atoms and 1.20 and 1.15 Å for all non–hydrogen atoms, respectively. Slight changes were observed in Ub’s S1–S2 loop to enable interaction between one of L8’s methyl groups and the C $\beta$  group of Rpn13 Pru’s D78. This interaction is reflected in the NMR data, as an NOE

crosspeak was observed between D79's amide group and L8's methyl group in Ub (Figure 1.15). In addition, congruent chemical shift changes were observed for this loop and the adjoining strands (Figure 1.14b). Human Rpn13 Pru's interaction with Ub occurs predominantly through residues in the S4–S5 and S6–S7 loops (Figure 1.16). In this region, D78's O $\delta$ 1 and D79's O $\delta$ 2 form hydrogen bonds with Ub's N $\epsilon$ 2 and N $\delta$ 1 of H68, respectively, whereas F76 is engaged in hydrophobic interactions with I44, Q49, and V70 of Ub (Figure 1.17a).

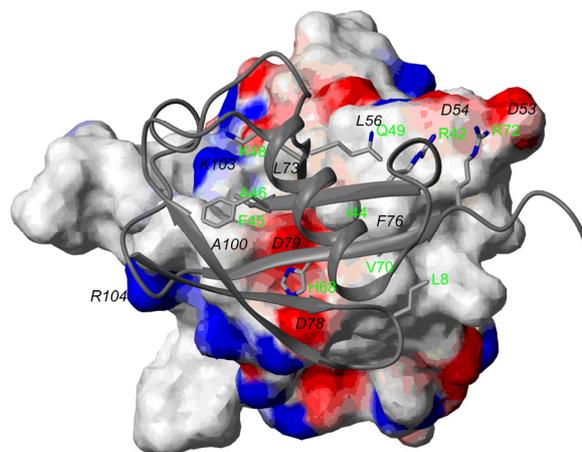


**Figure 1.17:** Structure of Rpn13 Pru:Ub complex defines a novel Ub-binding motif. (a) Zoomed view of Figure 1.16 to highlight several key hydrophobic (green dashed lines) or hydrogen bond (red dashed lines) interactions. Several contact residues from Ub (yellow) or Rpn13 (red) are displayed. (b) Zoomed view of Figure 1.16 rotated by 90° to display relative location of Ub's K48 (purple) and mRpn13 Pru's L73, K103 and R104 (pink). Additional contact residues from Rpn13 (pink) or Ub (yellow) are also highlighted.

These contacts are enabled by the strictly conserved P77, which causes the S4–S5 loop to turn. The resulting change in direction positions F76, D78 and D79 of Rpn13 Pru in an orientation favourable for Ub binding, and generates the core of the Ub-binding surface. The importance of D79 is supported by the strong reduction in binding observed after its substitution to asparagine (see Section 1.3.4.3). A series of additional hydrophobic contacts between Rpn13 Pru and Ub were observed. In particular, the sidechain methyl group of A100 and the C $\alpha$  group of G101 partially bury Ub's F45, which is solvent-exposed in the free protein. Rpn13's strictly conserved F98 located on S6 also becomes less solvent accessible through interactions with A46 and G47 of Ub (Figure 1.17a). Moreover, the methyl sidechain of L56 (S2–S3 loop), which is conserved in higher eukaryotes, contacts Ub's Q49 and is in proximity to R42 and R72 (Figure 1.17b). In agreement with the NMR data, mutation of L56 to alanine significantly reduced Rpn13 Pru binding to Ub (see Section 1.3.4.3). Calculation of

the electrostatic potential of mRpn13's Ub binding surface indicates that a hydrophobic region containing L56 and F76 is available to interact with Ub's L8, I44 and V70 (Figure 1.18).

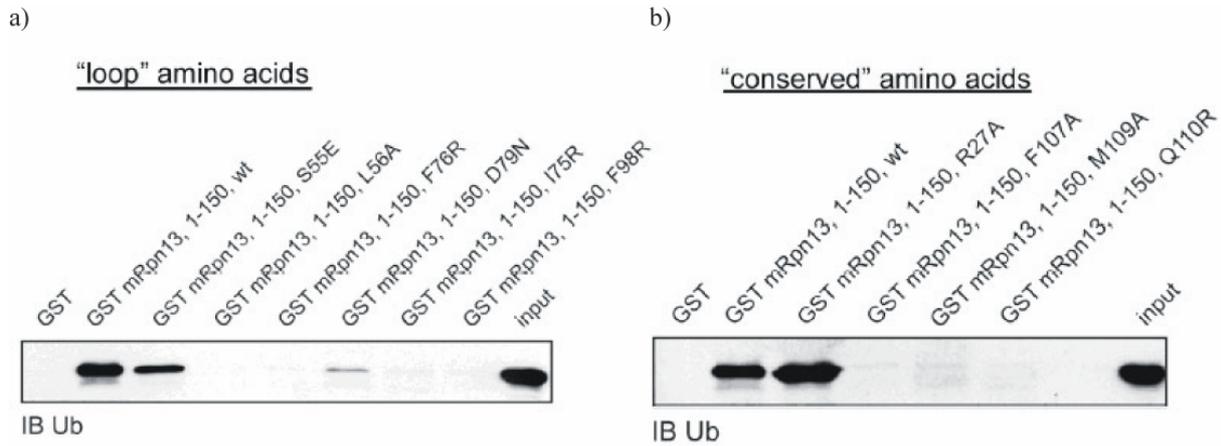
**Figure 1.18:** Representation of the electrostatic surface potential of mRpn13 Pru's Ub-binding surface with several of its key Ub-binding residues labelled in black. The surface is colour coded according to electrostatic potential, contoured from  $-25kT/e$  (intense red, negative charges) to  $+25kT/e$  (intense blue, positive charges). A ribbon representation of Ub is displayed with certain contact residues in ball-and-sticks representation and labelled in green.



Complementary electrostatic interactions between Rpn13 Pru and Ub also stabilize the complex, including interactions of D78 and D79 of Rpn13 Pru with Ub's H68, as well as D53 and D54 with Ub's R42 and R72. In total, the contact surface of Rpn13 Pru and Ub comprises  $1256 \text{ \AA}^2$ , which is large for Ub binding receptors.

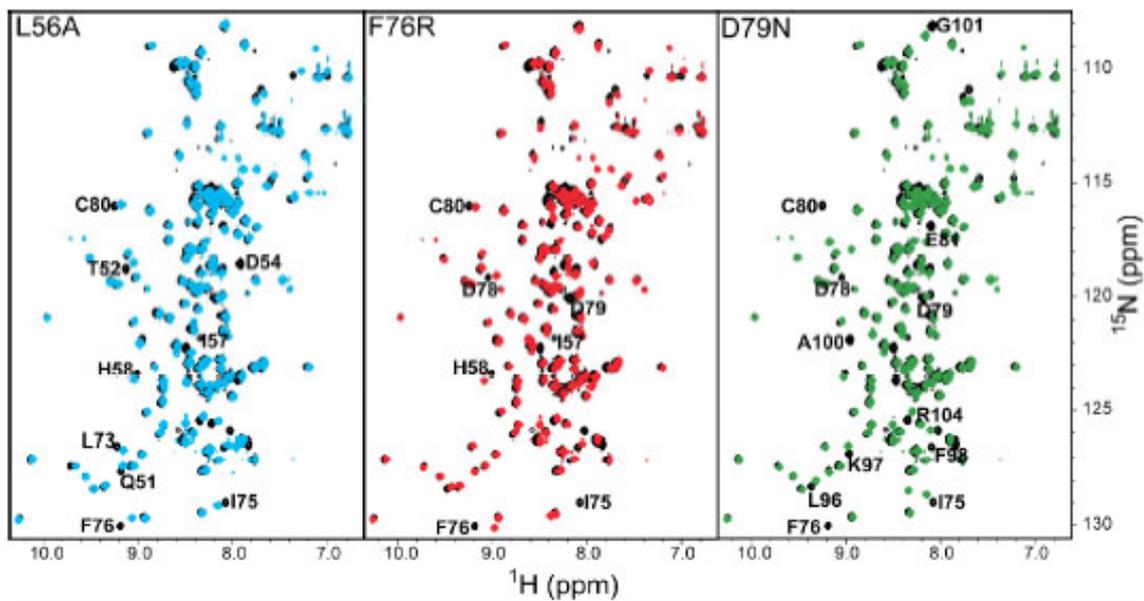
### 1.3.4.3 Mutational analysis of Pru domain binding to ubiquitin

The NMR data acquired on hRpn13 implicated the S2–S3, S4–S5, and S6–S7 loops as forming the Ub recognition surface, but these regions share only 22% sequence identity with the homologue from yeast. To analyze the significance of specific residues for Rpn13 binding to Ub, several amino acid substitutions were made, based either on the NMR Ub binding data or their conservation among all Rpn13 species. Amino acid substitutions within N-terminal (1-150), wildtype mRpn13 (R27A, F107A, M109A, Q110R, S55E, L56A, I75R, F76R, D79N, and F98R) were created in cooperation with the group of Dikic (Goethe University Medical School, Frankfurt) by *in vitro* mutagenesis. The protein products were expressed as GST-fusions and used in GST pull-down assays. The results support and extend the NMR experiments since they revealed specific residues in the S2–S3 (L56A), S4–S5 (I75R, F76R, D79N), and S6–S7 (F98R) loops to be essential for Ub binding as their amino acid substitutions abolished binding to 4 x Ub in the pull down assays (Figure 1.19a).



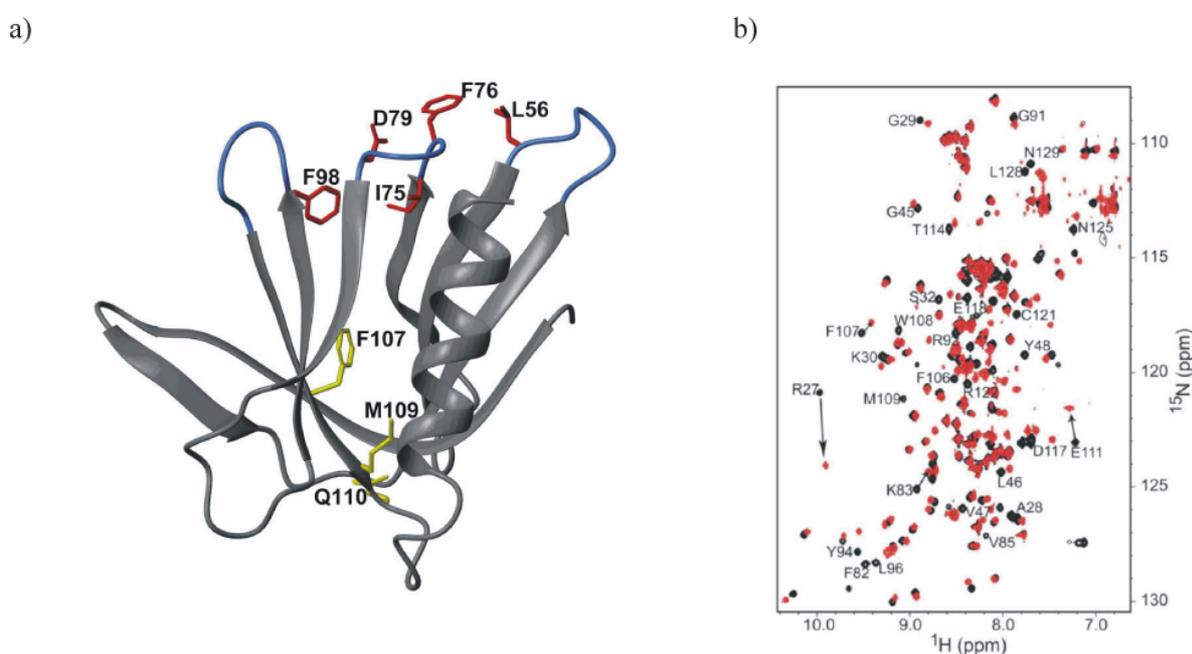
**Figure 1.19:** Identification of *Rpn13* Pru residues critical for Ub binding. (a) Specific amino acid substitutions were made within the S2–S3, S4–S5, and S6–S7 loops of *mRpn13* Pru by *in vitro* mutagenesis and the protein products were expressed as GST–fusions and used in GST pull-down assays to highlight the importance of these loops for tetraUb binding. (b) Experiments that parallel (a) were performed to demonstrate the importance of a subset of conserved residues for the structural integrity of the hydrophobic core and the PH fold.

Furthermore, NMR experiments on *mRpn13* with the incorporated L56A, F76R or D79N mutations indicated that the loss of Ub binding was not caused by loss of structural integrity (Figure 1.20).



**Figure 1.20:** The L56A, F76R and D79N point mutations do not compromise the structural integrity of the *mRpn13* Pru domain. Only local effects are observed in superimposed <sup>1</sup>H, <sup>15</sup>N HSQC spectra of <sup>15</sup>N labelled *mRpn13* Pru wild-type (black) and L56A (cyan), F76R (red), or D79N (green) protein. Shifted resonances are labelled in black.

The results also highlight the importance of structural integrity of the PH fold for Ub binding, since amino acid substitutions that disturbed the PH structure were unable to support Ub binding and those residues that are most conserved among Rpn13 species are important for structural integrity (Figure 1.13 and 1.19b). More specifically, sequence analysis identified three motifs, E25–K30 (M1), G91–F98 (M2), and F106–E111 (M3), as exhibiting the highest degree of identity among Rpn13 species (Figure 1.13, green boxes). The structural integrity of Rpn13 was readily compromised by substitutions in M3. For example, alanine substitution of F107 yielded unstable protein that bound only weakly to Ub (Figure 1.19b). F107 is not surface exposed and is directly involved in forming the Rpn13 hydrophobic core (Figure 1.21a, see also Section 1.3.3.2).



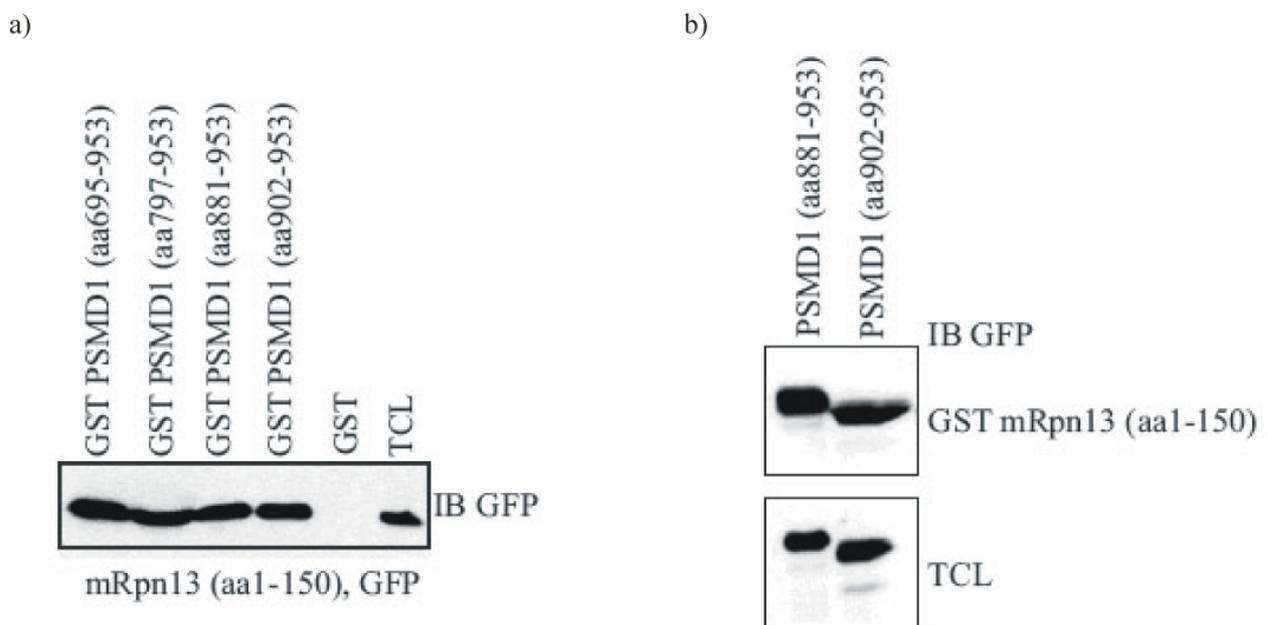
**Figure 1.21:** (a) Location of key mutated residues of mRpn13 Pru from Figure 1.19a and b displayed in red and yellow, respectively; Ub-binding loops are in blue. (b) Superposition of  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of wild type (black) or M109A (red) mRpn13 Pru indicates that residues remote from M109 were significantly shifted by its mutation suggesting that the architecture of the PH domain was distorted. Shifted resonances are labelled in black.

Similarly, substitution of M109 (M109A) or Q110 (Q110R) largely abolished Ub binding (Figure 1.19b). These residues are also located within the hydrophobic core (Figure 1.21a) and a  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC experiment was used to evaluate the structural consequences of the M109A mutation in mRpn13 Pru. This experiment provided evidence of structural rearrangements in Rpn13, as a large number of residues shifted dramatically, including R27, Y48, F82, K83, Y94, and C121 (Figure 1.21b). Altogether, the data suggest that F107, M109 and Q110 abolished Rpn13's Ub binding capacity by disrupting its structural integrity and

thereby highlight the importance of the PH fold for Rpn13 Ub binding. In contrast, substitutions located in the loop regions abolished binding to Ub but did not harm the structural integrity.

### 1.3.4.4 Rpn13 binds proteasome subunit Rpn2 independently of Ub

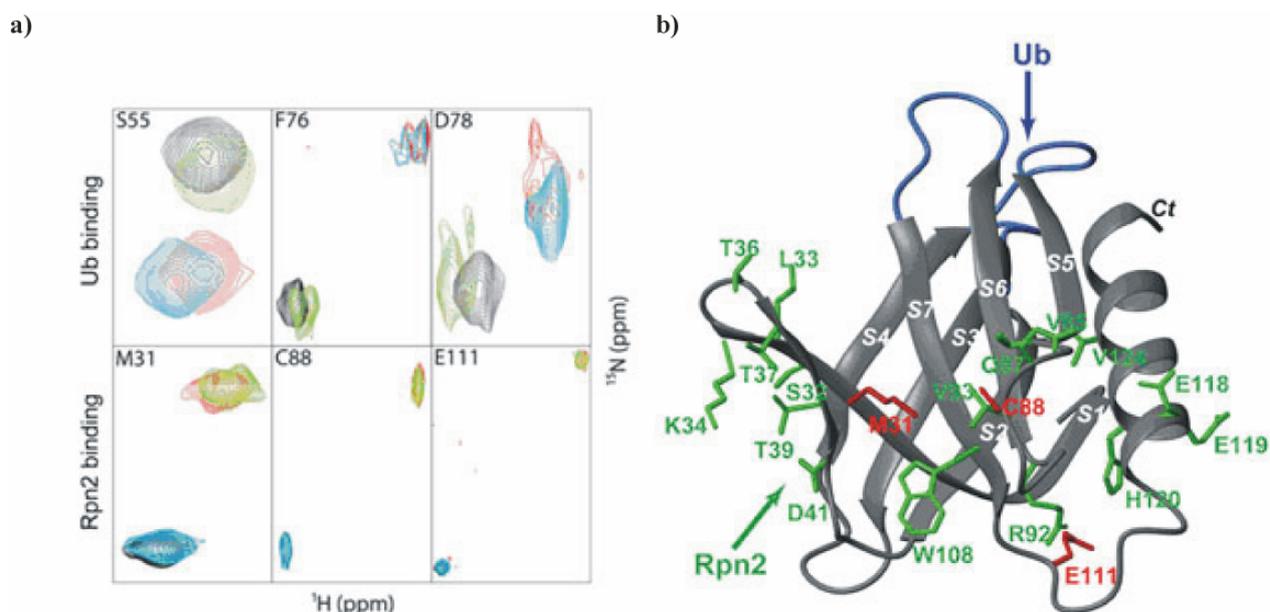
To function as a proteasomal Ub receptor, Rpn13 must bind Ub and proteasome components simultaneously. In both yeast and mammals, Rpn13 binds to Rpn2 through its Pru domain<sup>80,92,93</sup>. To confirm these specific binding requirements of Rpn13 Pru for Ub and Rpn2 a subset of N-terminal deletion mutants of human Rpn2/PSMD1 (aa695–953, aa797–953, aa881–953, and aa902–953) were designed and subcloned into pEGFP–C1 and pGEX–4T1 vectors. The deletion mutants were expressed as GST–fused proteins and tested for their ability to bind mRpn13 (1–150) transiently expressed as myc–tagged protein in HEK293T cells (Figure 1.22).



**Figure 1.22:** Interaction of Rpn13 Pru with base subunit hRpn2/PSMD1. (a) Different deletion mutants of hRpn2 (695–953, 797–953, 881–953 and 902–953) were expressed as GST–fused proteins and tested for their ability to bind mRpn13 (1–150) transiently expressed as myc–tagged protein in HEK293T cells. All deletion mutants of hRpn2 could bind mRpn13, indicating that the last 52 amino acids of hRpn2 are sufficient for mRpn13 binding. (b) Two deletion mutants of Rpn2 (881–953 and 902–953) were transiently expressed as GFPtagged proteins in HEK293T cells and tested for their ability to bind GST–mRpn13 (1–150).

Additionally, two deletion mutants of Rpn2 (881–953 and 902–953) were expressed as GFPtagged proteins in HEK293T cells and tested for their ability to bind GST–mRpn13 (1-150). All deletion mutants of hRpn2 could bind mRpn13, showing that the last 52 amino acids of hRpn2 are sufficient for mRpn13 binding (Figure 1.22).

Using NMR spectroscopy, we could map the surface of Rpn13 Pru for Rpn2 binding which is distinct from the Ub binding site (Figure 1.23a). More specifically, only Ub, and not hRpn2 (797–953), bound to residues in the S2–S3, S4–S5, and S6–S7 loops of hRpn13, as apparent from chemical shift data for residues S55, F76, and D78 (Figure 1.23a). In contrast, M31, C88, and E111 were not affected by Ub binding, but shifted after hRpn2 (797–953) addition. Furthermore, when both Rpn2 and Ub were added, S55, F76, and D78 contacted Ub while M31, C88, and E111 contacted Rpn2 (Figure 1.23a), indicating that the two binding surfaces are largely independent. M31, C88, and E111 are conserved in mRpn13 (Figure 1.23b) and map to S1, the S5–S6 loop, and the region linking S7 to H1, respectively.



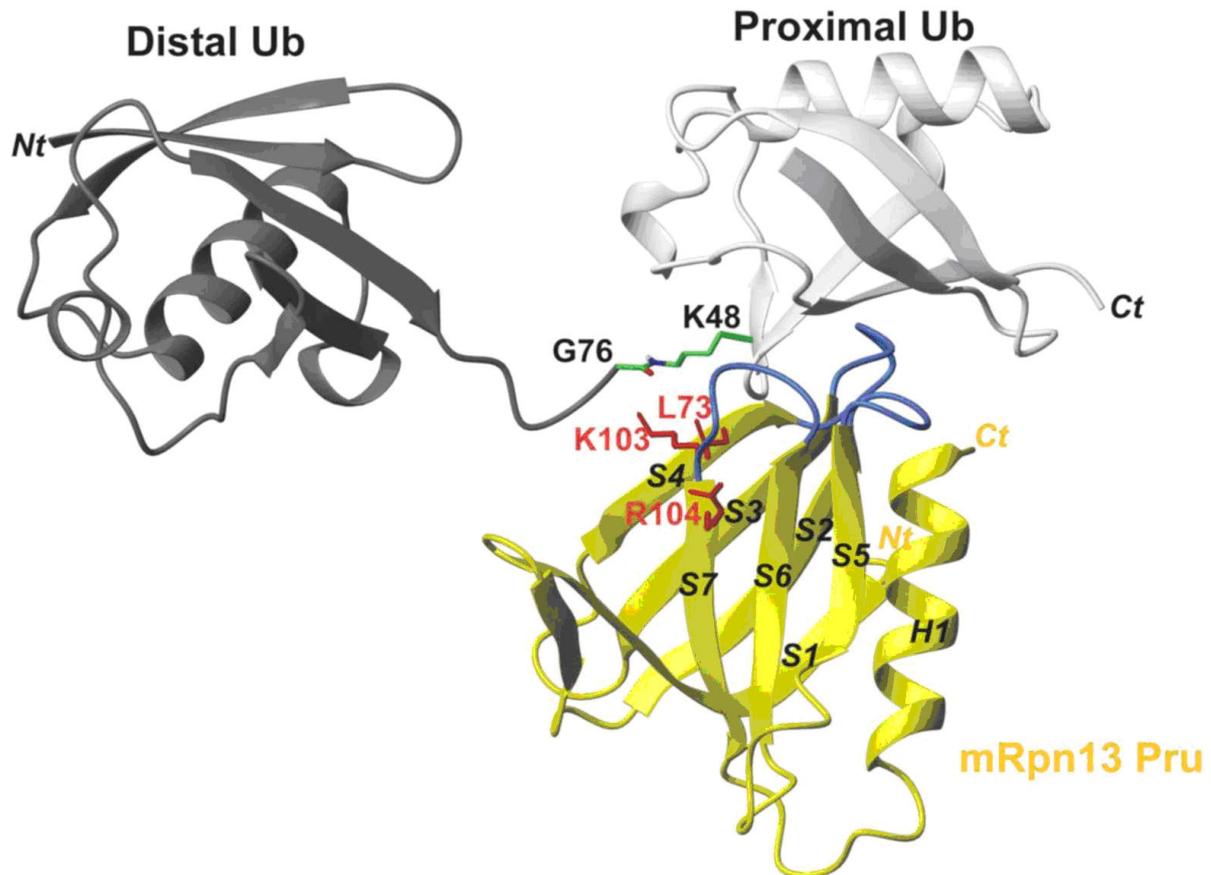
**Figure 1.23:** *Rpn13 Pru binds simultaneously Ub and Rpn2. (a) Ub (blue), hRpn2 (797–953) (green), or Ub and hRpn2 (797–953) (red) were added to hRpn13Pru domain, which was monitored by  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC experiments. Comparison with the spectrum acquired on the protein alone (black) indicates that S55, F76 and D78 bind Ub in a manner that is independent of hRpn2 (797–953), whereas M31, C88 and E111 bind hRpn2 in a Ub-independent manner. (b) The hRpn2-binding surface of mRpn13 was mapped from NMR titration experiments. The sidechain residues attenuated by Rpn2 (797–953) are displayed in green, and those of M31, C88, and E111 highlighted in red. The Ub binding loops, which are remote from the Rpn2-binding residues, are indicated in blue. Location of key mutated residues in mRpn13 Pru from Figure 1.19a and b displayed in red and yellow, respectively; Ub-binding loops are in blue.*

Mutational analysis of these three amino acid residues would provide further insight about the Rpn13:Rpn2 interaction.

To define the Rpn2-binding surface more precisely, hRpn2 (797–953) was added to  $^{15}\text{N}$  labelled hRpn13 and its amide resonances were monitored. At least nineteen hRpn13 crosspeaks were shifted or attenuated, corresponding to residues within the S1–S2 and S5–S6 loops, as well as at the C-terminal end of S7 and the beginning of H1. They form a contiguous surface opposite to the Ub binding loops and include conserved (V85, R92, V93, E111, and E119) and strictly conserved (W108) residues. Overall, binding surfaces of both the Rpn13 Pru and Rpn2 contain many charged amino acid residues, with charges being most likely distributed in a complementary fashion to facilitate binding.

#### 1.3.4.5 K48-linked tetraUb binds hRpn13 Pru

26S proteasomes exhibit high affinity for polyubiquitinated substrates<sup>125</sup>. Ub chains linked via isopeptide bonds between K48 and the C-terminal glycine of neighbouring Ub molecules are known to trigger proteasomal degradation of the labelled protein<sup>1,120,121</sup>. Chemical shift perturbation analysis, shown in Section 1.3.4.1 (Figure 1.14a, red) revealed that Rpn13 Pru binds K48-linked tetraUb. Importantly, K48-linked tetraUb affected hRpn13 residues in a manner comparable to that of monoUb (Figure 1.14a, red versus black). More specifically, tetraUb and monoUb caused chemical shift changes to the same hRpn13 residues, including L56, F76, and F98, and shifted them almost identically. It was demonstrated that hRpn13 binds diUb elements of K48-linked tetraUb<sup>99</sup> and the comparison in Figure 1.14a therefore indicates that one of the Ub moieties within a diUb element binds hRpn13 in a manner similar to monoUb. Only two residues in hRpn13 exhibited changes that were specific to K48-linked tetraUb, namely L73 and R104 (Figure 1.14a). These residues and the sidechain atoms of neighbouring K103 are proximal to each other and in the mRpn13:monoUb structure directed towards K48's sidechain atoms (Figure 1.17b). This arrangement is congruent with Rpn13 forming interactions with K48-linked tetraUb that are not present in the monoUb complex, as suggested by the chemical shift perturbation data. To gain further insight into how Rpn13 binds neighbouring ubiquitins in K48-linked chains, the Rpn13:monoUb structure was used to build a model with diUb (Figure 1.24). In this model, the Rpn13 bound Ub is conjugated to another Ub molecule by an isopeptide bond, which is positioned to avoid steric clashes and the resulting complex is energy minimized.

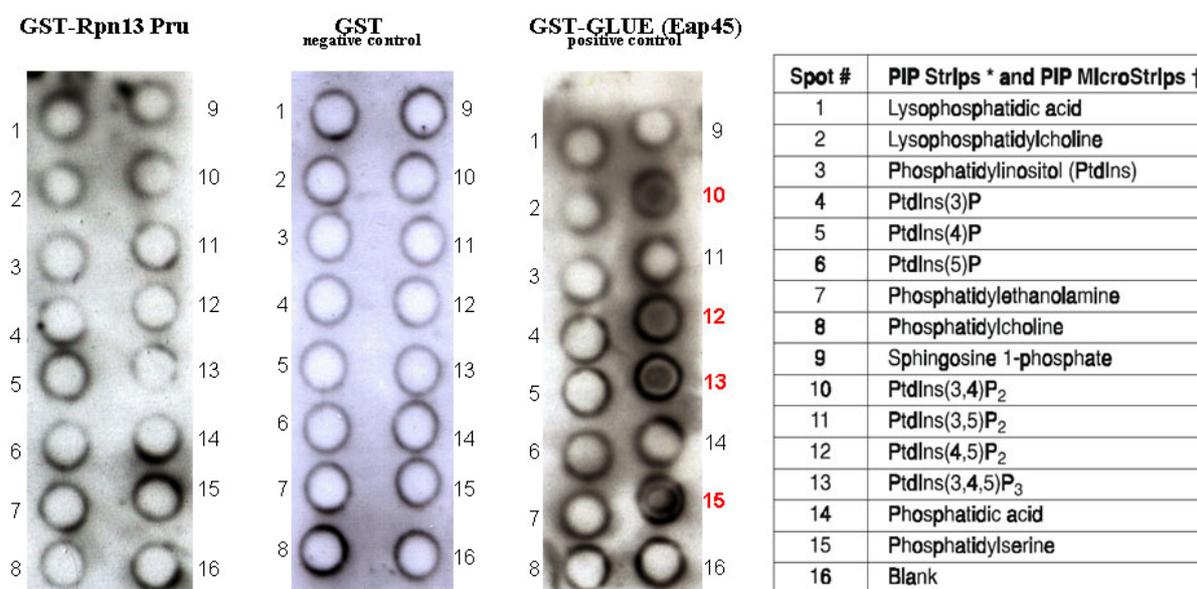


**Figure 1.24:** Model of *mRpn13 Pru* complexed with *diUb*. White and grey ribbon diagrams display the proximal and distal Ub, respectively, whereas a balls-and-sticks representation is used for the K48–G76 isopeptide bond linkage. *Rpn13 Pru* is coloured in yellow and loops recognizing Ub in blue, whereas L73, K103 and R104 are displayed in red. *DiUb* was created by *Insight II* software based.

As expected, the model revealed putative interactions between the Ub–linker region and *Rpn13*'s L73 and K103. In particular, when the proximal Ub binds *Rpn13*, the K48–G76 isopeptide bond is located  $<4 \text{ \AA}$  away from *Rpn13*'s L73, which becomes partially buried. The separation of the C $\delta$  groups of Ub's linked K48 and *Rpn13*'s L73 is reduced by  $\sim 1 \text{ \AA}$  as compared to in the complex with monoUb, due to rotations in the K48 sidechain that avoid steric clashes between the distal Ub subunit and *Rpn13*. This rotation similarly places K103's sidechain atoms within  $4 \text{ \AA}$  of the carbonyl oxygen of the isopeptide bond, suggesting possible electrostatic interaction. To summarize, the data indicate that *Rpn13* readily binds K48–linked chains and most likely affects the configuration of the linkage between conjugated ubiquitins.

### 1.3.5 Investigation of the interaction of Rpn13 Pru with PIPs

As has been reported recently, a small fraction of human Rpn13 was found to be membrane associated, similar as has been shown for 26S proteasomes<sup>86</sup>. Interestingly, PH domains are best known for their ability to recruit proteins to membranes through their association with membrane phosphatidylinositolphosphates (PIPs)<sup>111,113</sup>, although they can simultaneously function as general protein-binding modules, like the GLUE domain which can bind to Ub and PIP independently<sup>126,127</sup>. A protein-lipid overlay assay was performed using PIP Strips<sup>TM</sup> membranes containing different phosphatidylinositides and lipids to test whether mRpn13 Pru can bind to PIPs. As shown in Figure 1.25, GST tagged Rpn13 Pru did not specifically bind to PIPs. In contrast, the GLUE domain, which is a PH domain known to bind PIPs, exhibited signals for PIP binding.

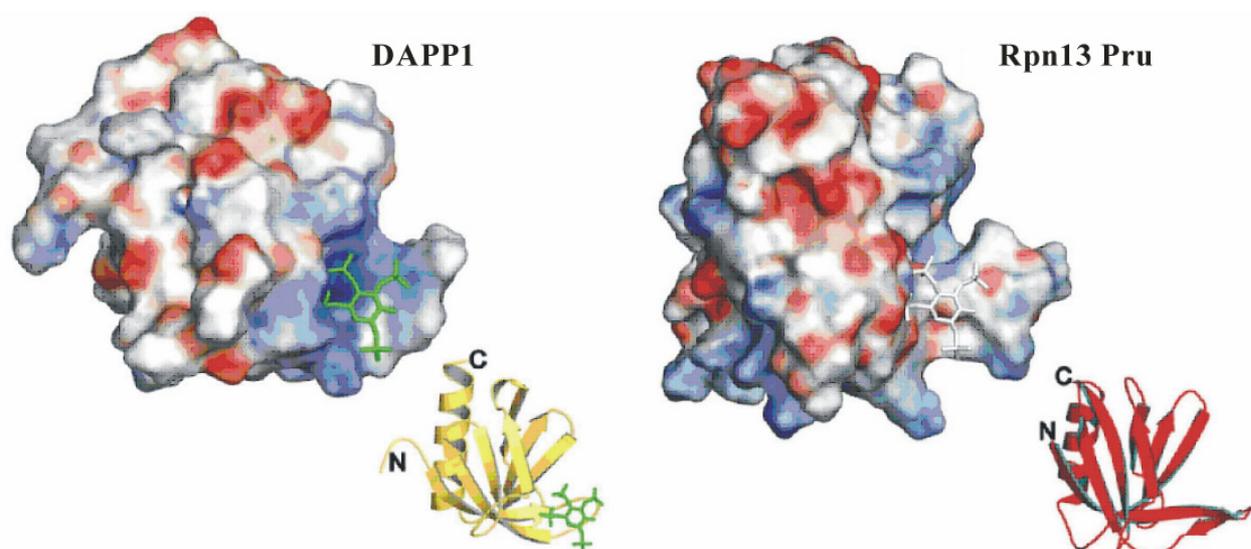


**Figure 1.25:** Protein-lipid overlay assay of mRpn13 Pru using PIP Strips<sup>TM</sup> membranes containing 100 pmol of phospholipids. The membranes were incubated with GST alone, GST-tagged Rpn13Pru and GST-tagged GLUE. Membranes were decorated with GST antibody and bound protein was detected by chemiluminescence. In contrast to GST-GLUE as positive control, GST (negative control) and GST-tagged Rpn13 did not show any signal.

We compared structures of Rpn13 Pru and other PH domains bound to PIPs in order to get insights about structural requirements for binding to PIPs.

Most PH domains bind PIPs at a canonical site located between the S1-S2 and S3-S4 loops. A typical example is DAPP1 with PIP bound to the canonical site (PDB accession code 1FAO)<sup>113</sup>. Rpn13 Pru shows a similar architecture of the putative PIP binding grove as found at the PIP binding site in DAPP1:PIP complex (Figure 1.26). In DAPP1 it is the loop between

strands S1 and S2 which is mainly responsible for the interaction, and which is strictly conserved among PH domains binding PIPs<sup>110</sup>. Although the conformation of the loop between S1 and S2 in Rpn13 Pru is similar to that of DAPP1, the overall charge distribution on the surface of the Rpn13 Pru differs from the positively charged profile of DAPP1:PIP binding site (Figure 1.26). The positively charged surface of the molecule is essential for binding the negatively charged phosphate side chains of PIPs. Additionally, this region is not conserved among members of the Rpn13 family, thus excluding a possibility that these molecules carry a canonical PIP binding site. Together with the result from the protein–lipid overlay experiment, it is very unlikely that the Pru domain can specifically interact with PIPs.

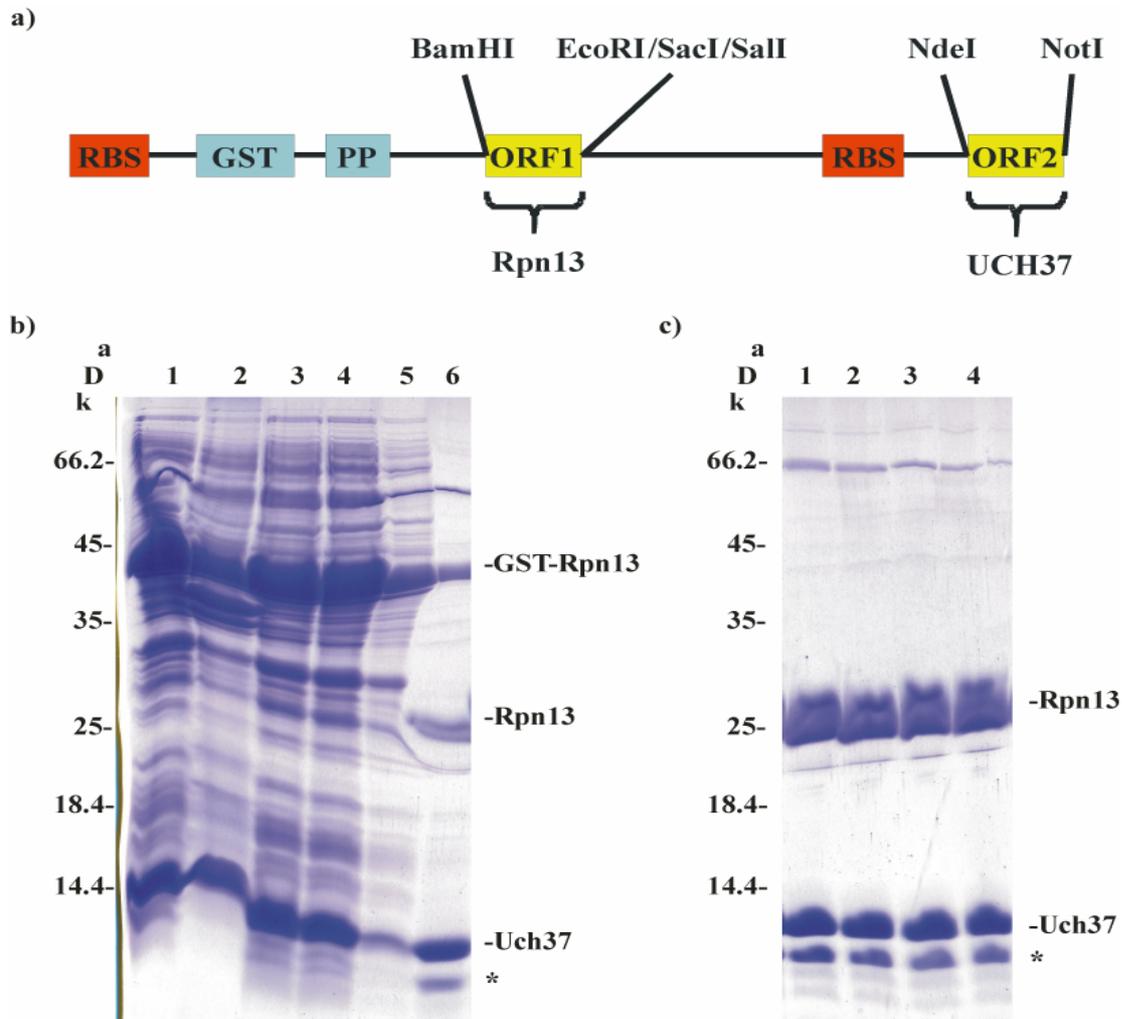


**Figure 1.26:** Structural comparison of Rpn13 Pru and DAPP1 bound to inositol 1,3,4,5-tetrakisphosphate ( $P_5I$ ) (PDB accession code 1FAO). Electrostatic potential surface representations of Rpn13 Pru and DAPP1 are colour coded according to their electrostatic potential, contoured from  $-25kT/e$  (intense red, negative charges) to  $+25kT/e$  (intense blue, positive charges). Ribbon presentations of molecules in the same orientation are shown as insets.

### 1.3.6 Interaction of Rpn13 with the Ubiquitin C-terminal hydrolase (UCH) 37

The C-terminal part of Rpn13 is supposed to recruit the deubiquitinating enzyme Uch37 to the ubiquitinated substrates for subsequent release of the polyUb chains and therefore might also function in processing of substrate proteins. Rpn13 is not only reported to bind to the C-terminus of Uch37, but was also shown to significantly stimulate the enzymatic activity of the enzyme upon binding<sup>76-78</sup>. To structurally investigate the interaction of Rpn13 with Uch37, different C-terminal variants of Rpn13 were expressed and subjected to crystallization. The Rpn13 variants 5, 6, 7 and 8 listed in Table 1.1 were cloned and expressed as described in Section 1.3.1. Truncated versions of Uch37 (aa 6–228, aa 6–325) were cloned in pGEX–6P–1 vector and purified according to the same procedure as Rpn13. So far, no crystals were obtained.

To gain more insight into the interaction modus of Rpn13 with Uch37 the two proteins were coexpressed using a bicistronic expression system. Therefore, a vector with a bicistronic operon was constructed by introducing a second ribosomal binding site (RBS) and a second open reading frame (ORF2) into the pGEX–6P–1 vector (Figure 1.27a) as described in Materials and Methods. The C-terminal domain of Rpn13 was cloned into the first ORF1, thereby introducing a PreScission Protease cleavable GST-tag, whereas Uch37 was cloned into the second ORF2. Both proteins were expressed and purified from bacterial lysate by affinity chromatography using Glutathione Sepharose and eluted with PreScission Protease according to the same procedure as in Section 1.3.1. The complex was further purified via size exclusion chromatography (see Materials and Methods) and subjected to crystallization. Rpn13 and Uch37 formed a stable complex *in vitro*, since only Rpn13 carried an affinity tag and the complex remained stable upon gel filtration. Figure 1.27 shows the Rpn13 (aa 202–407):Uch37 (aa 230–325) construct as an example for the copurification. The DNA sequence encoding this Uch37 construct contains an unexpected Shine–Delgarno sequence in proximity to the transcription start, which led to the additional expression of a shorter version of Uch37 (marked with an asterisk). First results indicate that this shortened version (aa 249–325) is still able to bind to Rpn13. As demonstrated in Section 1.3.2.1 (Figure 1.7), the linker domain in Rpn13 lacks secondary structure elements and is rather unlikely to be involved in interaction with Uch37. Therefore, a shortened variant (300–407) of hRpn13 was constructed, which was sufficient for binding to Uch37 and represents the shortest known Uch37–interaction module on Rpn13 so far. In future, further truncation mutagenesis has to be performed to specifically map the interaction module and to improve crystallization tendency.



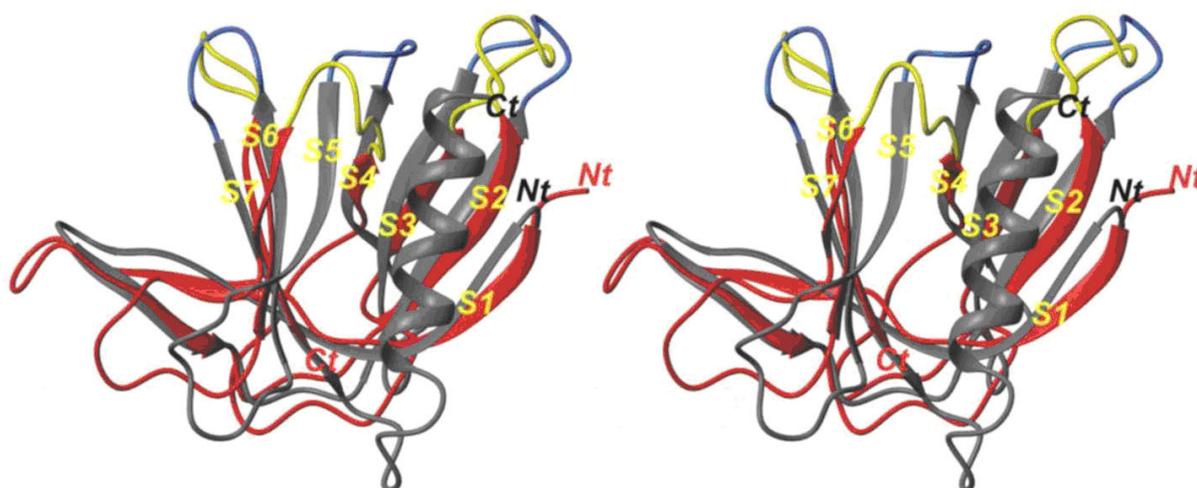
**Figure 1.27:** Bicistronic expression and purification of Rpn13:Uch37 complex. (a) Schematic view of the constructed bicistronic expression vector based on pGEX-6P-1. Rpn13 was cloned into ORF1 and carried a cleavable GST-tag, whereas UCH37 was cloned into ORF2. (b) Coomassie stained SDS-PAGE (17.5%) of purification of the Rpn13(202–407):Uch37(230–325) complex with Glutathione-Sepharose. 1=after sonification, 2=pellet, 3=flow through, 5=wash, 6=elution with PP. The asterisk marks a Shine-Delgarno product of Uch37. (c) The Rpn13 (202–407):Uch37 (230–325) complex after run with Superdex 200 gel filtration column, 1–4=fractions.

## 1.4 Discussion

### 1.4.1 The N-terminal Pru domain of Rpn13

The structure of the CP has been solved by X-ray crystallography<sup>11-13</sup>, whereas structural insights on the RP are limited. The complexity of substrate recognition by the proteasome has hindered the development of general mechanistic models for how protein degradation is triggered and regulated. Proteasomal Ub receptors are well characterized, with the structures of proteasomal subunit Rpn10/S5a<sup>118,128</sup> as well as UBL/UBA family members<sup>129-132</sup>. This study could identify Rpn13 as a novel Ub receptor for the proteasome, which is unrelated to Rpn10 and the three UBL/UBA proteins, and moreover defines a new class of Ub recognition surface.

The Ub-binding region of Rpn13 adopts a PH fold and the complex structure of Rpn13 Pru with Ub unveiled a new Ub-binding mode. The PH fold is present in a remarkably large number of proteins<sup>110</sup>, but Rpn13 Pru is the first example of a PH domain structure within the 26S proteasome. NMR structural analysis of Rpn13 from yeast<sup>99</sup> revealed that it adopts a PH fold despite its low sequence identity with Rpn13 from higher eukaryotes (Figure 1.13 and 1.28). Residues that are conserved between fungi (*S. cerevisiae* and *S. pombe*) and eukaryotic Rpn13 are largely involved in forming and stabilizing the typical PH hydrophobic core (Figure 1.13).



**Figure 1.28:** Stereo representation of the structural superposition of murine (grey) and yeast (red) Rpn13 Pru. Ub binding loops for murine and yeast Rpn13 are coloured in blue and yellow, respectively<sup>99</sup>.

Structural superposition of secondary structure elements of murine and yeast Rpn13 resulted in an r. m. s. deviation of 2.55 Å (Figure 1.28). Since scRpn13 is the most divergent in primary sequence from mRpn13, we conclude that all Rpn13 homologues contain N-terminal PH topology.

## 1.4.2 Comparison of Rpn13 from mouse and yeast

Rpn13 from budding yeast is the smallest member of the family, with only five residues preceding the N-terminal PH domain and no C-terminal Uch37-binding domain<sup>76-78</sup>. No Uch37 homologue has been identified in yeast, consistent with the omission of the C-terminal domain in scRpn13. Rpn13 proteins from mouse and yeast bind Ub using the S2–S3, S4–S5, and S6–S7 loops, which are 22% identical in primary sequence. The S6–S7 loop plays an important role in binding Ub for both species, with F91, S93 and R96 of scRpn13 expected to perform analogous roles to F98, G101 and R104 of mRpn13. By contrast, the S4–S5 loop in mRpn13 is conserved in mammals, but not in *S. cerevisiae*, in which it is shortened and directed towards the structural core rather than the Ub binding surface (Figure 1.28). Moreover, important interactions are lost between this loop and Ub in scRpn13; for example, D78, which in mammals forms a hydrogen bond with H68, is replaced with G71 in scRpn13. It is tempting to speculate that mammals take better advantage of Ub's H68. Such differences between scRpn13 and mRpn13 correlate with the significantly higher affinity of hRpn13 for Ub as compared to scRpn13<sup>99</sup>. This alteration in Rpn13, together with the difference in its association with Uch37, suggests interesting evolutionary adjustments of the role of Rpn13 in proteasome function, the details of which remain to be described.

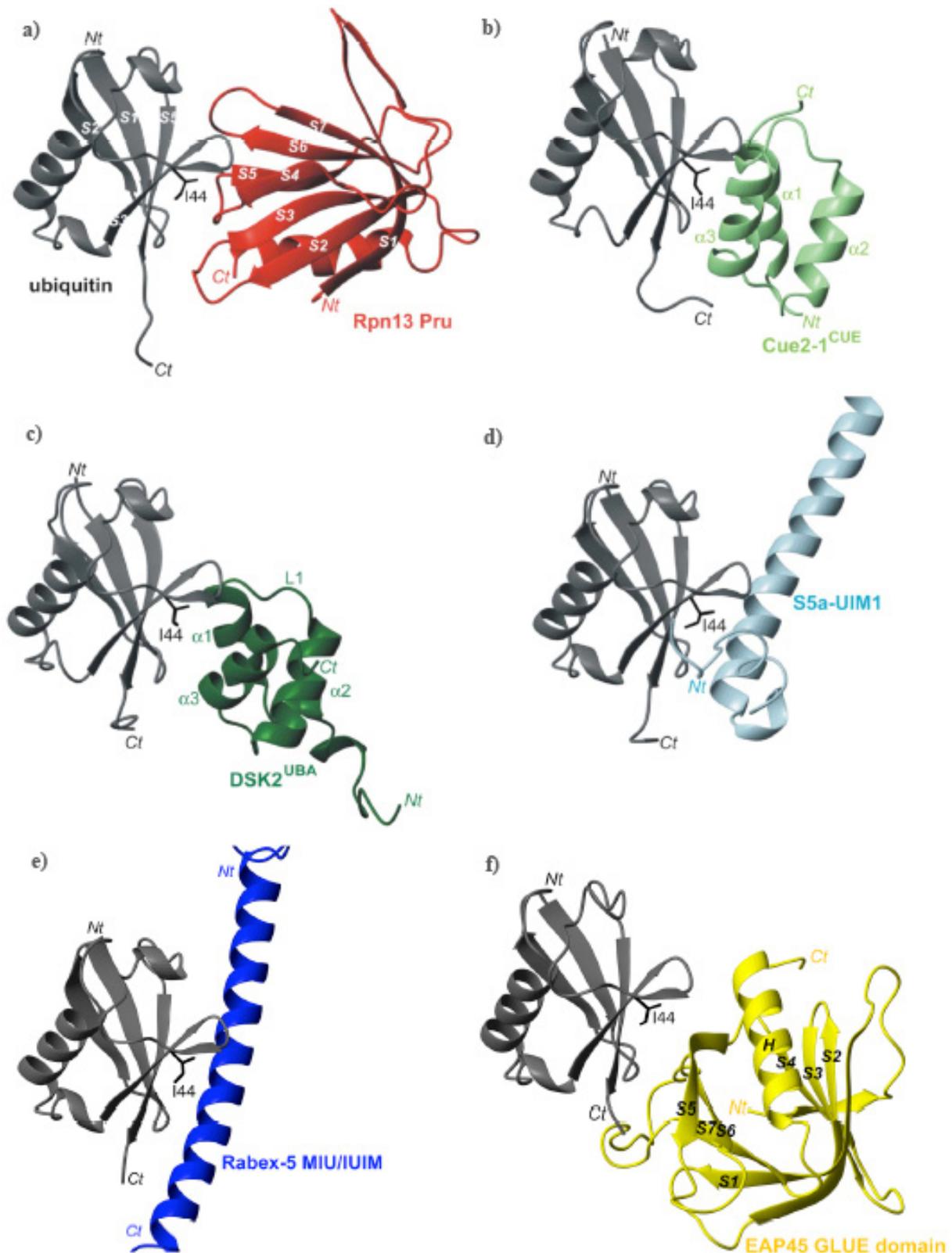
### 1.4.3 Rpn13 loops bind Ub

Rpn13, like many other Ub receptors, binds to the L8, I44, and V70 hydrophobic pocket of Ub. However, the Pru domain of Rpn13 is the first to bind this region using exclusively loops (Figure 1.29). Ubiquitin-binding domains, which complex structures with ubiquitin were solved, are listed in Table 1.3.

Ub-binding domain	Full name	Source protein	PDB ID
UBA	<u>u</u> biquitin- <u>a</u> ssociated	Dsk2	1WR1
		hHR23A	1ZO6
		Mud1	1Z96
		Ede1	2G3Q
CUE	<u>c</u> oupling of <u>u</u> biquitin conjugation to <u>e</u> ndoplasmatic reticulum degradation	Vsp9	1P3Q
		Cue2	1OTR
GAT	<u>G</u> olgi-localized, gamma-ear-containing, ADP-ribosylation-factor-binding protein and target of myb	GGA3	1YD8
			1WR6
		TOM1	1WRD
UEV	<u>U</u> bc <u>E</u> 2 <u>v</u> ariant	Vps23	1UZX
		Tsg101	1S1Q
Ubc	<u>u</u> biquitin- <u>c</u> onjugating enzyme	UbcH5	2FUH
UIM	<u>u</u> biquitin- <u>i</u> nteracting <u>m</u> otif	Vsp27	1Q0W
			1O06
		S5a	1YX5
			1YX6
DUIM	<u>d</u> ouble-sided <u>U</u> I <u>M</u>	Hrs	2D3G
MIU	<u>m</u> otif <u>i</u> nteracting with <u>u</u> biquitin	Rabex-5	2FID, 2FIF
			2C7N
NZF	<u>N</u> pl4 <u>z</u> inc <u>f</u> inger	Npl4	1Q5W
A20 ZnF	<u>A</u> 20 <u>z</u> inc <u>f</u> inger	Rabex-5	2FID, 2FIF
			2C7N
ZnF UBP	<u>z</u> inc <u>f</u> inger <u>u</u> biquitin-specific processing protease	Isopeptidase T	2G45
GLUE	<u>G</u> RAM-like <u>u</u> biquitin binding in <u>E</u> LL-associated protein 45	Vsp36/EAP45	2HTH
			2DX5
PH	Pleckstrin homology	Rpn13	2Z59

**Table 1.3:** Ubiquitin-binding domains. Protein data bank codes (PDB ID) are only listed for co-ordinates containing ubiquitin-binding domains complexed with ubiquitin.

Most of the Ub receptors characterized to date use  $\alpha$ -helices to bind this surface of Ub, including UBA, CUE, UIM, DUIM, MIU and GAT<sup>133</sup>. Among them, the UBA and CUE domains are structurally homologous showing common 3-helical bundle architecture. The CUE domain of Cue2-1 binds Ub through the  $\alpha$ 1 and  $\alpha$ 3 helices (Figure 1.29b)<sup>116</sup>, whereas the UBA-domain of Dsk2 uses the loop between  $\alpha$ 1 and  $\alpha$ 2, as well as the C-terminal part of  $\alpha$ 3 (Figure 1.29c)<sup>131</sup>. M342 of Dsk2's UBA and M19 of Cue2-1's CUE domain play an integral role in recognizing Ub by forming hydrophobic contacts with its I44-centred pocket and a hydrogen bond to G47<sup>116,131</sup>. Structural characterization of the UIMs demonstrated that a single  $\alpha$ -helix is sufficient for binding this region of Ub<sup>118,134,135</sup>. The UIM helix includes a conserved alanine neighbored by a bulky hydrophobic residue, each of which packs against Ub's I44 as demonstrated in the S5a UIM1/Ub complex (Figure 1.29d)<sup>118</sup>. The MIU/IUIM domain of Rabex-5 (Figure 1.29e)<sup>136,137</sup> and the pol  $\eta$  UBZ domain<sup>138</sup> similarly bind this region in Ub through a single  $\alpha$ -helix, with a contact surface formed by hydrophobic residues proximal to a central invariant alanine. In this special case the helix is in the reverse orientation. The GLUE domain of ESCRT-II EAP45, which exhibits a split pleckstrin-homology topology, is the only previously known Ub-binding PH domain<sup>139</sup>. However it binds Ub in a different manner: the I44 containing surface of Ub is contacted by residues within secondary structural elements including the EAP45 C-terminal helix corresponding to H1 in Rpn13<sup>139</sup> (Figure 1.29f). Moreover, although the longer S6-S7 loop of EAP45 is involved in binding Ub, the S2-S3 and S4-S5 loops are not; instead, contacts are formed by residues from S5 and S6. Like these other receptors, Rpn13 binds the I44 hydrophobic patch of Ub, but does so using three loops to generate a relatively large buried surface area, which comprises 1256 Å<sup>2</sup> for both molecules in complex. Thus, the relatively high affinity of 7.5  $\mu$ M for Rpn13 binding to monoUb<sup>99</sup> is explained by the enlarged contact surface as compared to that of EAP45 GLUE (1000 Å<sup>2</sup> in total)<sup>140</sup>. Published values for the total buried surface of Cue2-1<sup>cue</sup> and Dsk2<sup>UBA</sup> upon Ub binding are even smaller: 960 and 800 Å<sup>2</sup>, respectively<sup>116,131</sup>.



**Figure 1.29:** Structural comparison of Ub receptors complexed with Ub. (a)–(f) Complex structures of Ub and specific receptors displayed with Ub in the same orientation (grey) and Ub's I44 shown in black sticks. Each receptor has unique colour coding: (a) Rpn13 Pru (red), (b) Cue2-1<sup>CUE</sup> (PDB code 1OTR; light green), (c) Dsk2<sup>UBA</sup> (PDB code 1WR1; dark green), (d) S5a-UIM1 (PDB code 1YX5; light blue), (e) Rabex-5 MIU/IUIM (PDB code 2FIF; dark blue), (f) EAP45 GLUE domain (PDB code 2HTH; yellow). All structures are compared by a best-fit superposition of bound Ub (grey). In (c), L1 denotes the loop connecting  $\alpha 1$  and  $\alpha 2$ .

## 1.4.4 The first proteasomal Pleckstrin Homology domain

As shown in this work, Rpn13 Pru adopts the fold typical for Pleckstrin Homology (PH) domains, which are present in a remarkably large number of proteins from yeast to humans<sup>110</sup>. So far, Rpn13 is the first example which adopts a PH topology within the 26S proteasome. PH domains are small modules (100–120 amino acids) found in a multitude of intracellular proteins with very widespread functions<sup>106-110</sup>. Although PH domains show low homology at the primary sequence level and have diverse cellular functions as well as ligand/protein binding properties, their overall architecture and composition of the hydrophobic core are well conserved. Each PH domain is electrostatically polarised, with variable loops forming a charged surface. These loops are supposed to determine variability and specificity of protein:protein and protein:ligand interactions of different PH domains and indeed, binding of Rpn13 Pru to ubiquitin is exclusively mediated by its loops. Wide distribution and functional diversity of the PH domain family originates from their ability to bind to virtually any target by simple adaptation of their molecular surface and by displaying highly variable residues throughout the protein surface. This feature of surface adaptation is comparable to the modularity of other protein binding scaffolds like the ankyrin-repeat<sup>141</sup> or the immunoglobulin fold. The NMR data showed that Rpn13 Pru binds to Rpn2 and Ub from opposite directions, thus displaying another example of the protein adaptor functions of PH domains. The very high degree of flexibility in PH domains makes it rather difficult, to predict putative interaction partners or binding motifs of Rpn13 by comparison with homologous PH domains. The GLUE domain of ESCRT-II EAP45, which is so far the only identified PH domain known to bind Ub<sup>139,140</sup> in contrast to Pru, forms a split pleckstrin-homology domain, but binds ubiquitin along the S5 edge of the  $\beta$ -sandwich<sup>139</sup>. Rpn13's ubiquitin binding mode is novel, and it will be interesting to learn whether other receptors use this unique binding mode.

Many PH-domains act in pathways of membrane trafficking and were shown to bind phosphatidylinositolphosphates (PIPs) through a conserved and positively charged pocket<sup>110,142</sup>. Additionally, it has been shown that a small fraction of Rpn13, like proteasome, appears to be membrane associated<sup>86</sup>. However, protein-overlay assay and the structural superposition performed in this study exclude an interaction of Rpn13 Pru with PIPs.

### 1.4.5 Rpn13 and its neighbours in the proteasome

Rpn1 and Rpn2, which abut the hexameric ATPase complex, have been proposed to play scaffolding roles in the RP, as they bind multiple proteins involved in Ub chain docking and processing, including deubiquitinating enzyme Ubp6<sup>143</sup>, Ub ligase Hul5<sup>144</sup>, the arsenite-inducible factor AIRAP<sup>145</sup> and UBL/UBA family members<sup>41</sup>. Rpn10 may also reside in this region of the proteasome<sup>77,144</sup>. The extent of coordinate action on an Ub chain will depend on its length, but, given the presence of both an Ub ligase and two deubiquitinating enzymes in this assembly, chain length is likely to be highly dynamic for any conjugate docked in this area. The proteasome is typically viewed as a molecular machine with hard-wired connections between discrete parts whose movements are driven by the energy of ATP hydrolysis. However, the Ub receptors and Ub chain processing enzymes lack any obvious ATPase motifs. Many (perhaps most) contain notably flexible segments<sup>118,130</sup>, which may allow their active sites to explore large spaces in the periphery of their docking site in the proteasome. Such flexibility may enable several enzymes to work simultaneously on a given Ub chain. The identification of a new Ub receptor in this region of the proteasome, together with its intimate linkage with a deubiquitinating enzyme, emphasizes the importance of this region as well as the significance of Ub chain dynamics for substrates docked here. Rpn13 has a unique Ub binding preference for diUb elements within K48-linked chains<sup>99</sup> and most likely interacts directly with the isopeptide bond within an Ub chain. This Ub binding mode is consistent with the functional relationship of Rpn13 with Uch37, which it adds to this region's collection of chain processing enzymes<sup>76-78</sup>. Binding to Rpn13 both facilitates the deubiquitinating activity of Uch37<sup>76,77</sup> and links Uch37 to the proteasome, suggesting that Rpn13 plays a major role in Ub chain disassembly at the proteasome. We propose that, in docking Ub conjugates at the proteasome, Rpn13 performs two functionally distinct but related roles. The substrate component of the conjugate is exposed to an unfolding site in the proteasome, while, perhaps at the same time, the Ub component is presented to a chain disassembly site. The unfolding site is presumably the hexameric ATPase ring composed of Rpt1–6, while the disassembly site is proposed to be Uch37. These components are all found clustered within the base subassembly of the proteasome. Our findings suggest that the Ub receptor function of Rpn13 does not serve simply to enhance Uch37 activity, but, as other Ub receptors, it promotes protein degradation by participating in the substrate docking step. Uch37 has an explicit enzymatic directionality whereby it only deconjugates chains from the distal end, removing one Ub moiety at a time<sup>74</sup>. This strict cleavage specificity complements that of Ubp6 and Rpn11, as Ubp6 can deconjugate multiple Ub's in a single cleavage event<sup>143</sup> and Rpn11 performs “en bloc” deubiquitination from the proximal end<sup>54,55</sup>. Deubiquitinating

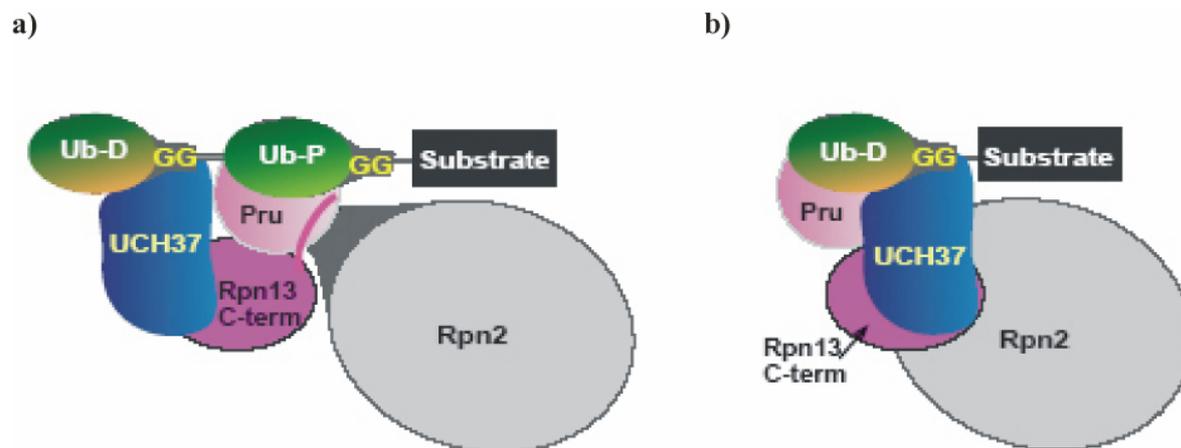
activities, particularly Ubp6, are antagonized by another Rpn2-bound protein, the chain elongation factor Hul5<sup>144</sup>. With so many receptors and chain-processing enzymes bound to Rpn1 and Rpn2, some of them mutually antagonistic, the detailed pathway by which a substrate is degraded may be subject to many stochastic variations. Whether this unanticipated design promotes high substrate flux through the proteasome is unclear, but it seems well suited to allow the cell to fine-tune proteasome activity<sup>143,145</sup>.

### 1.4.6 Rpn13 recruits Uch37 to the proteasome

Uch37 was proposed to function as an “editing isopeptidase”<sup>74</sup> and to disassemble chains bound to the proteasome, and, given that it removes one Ub at a time from the substrate-distal tip of the chain, to complete the removal of short chains more rapidly than long ones. In this way, Uch37 could help to suppress the degradation of oligoubiquitinated substrates by proteasomes, thereby functioning in a proofreading process. The attrition of long chains by Uch37 may have a more modest effect on their affinity for the proteasome, because they arrive with more Ub groups than are needed to maintain interaction with proteasomal Ub receptors. Recent studies<sup>76-78</sup> showed that Rpn13 recruits Uch37 to the proteasome and acts as a facilitator of Uch37 by the stimulation of its catalytic activity. If the Ub receptor activity of Rpn13 simply served to present chains to Uch37, then Rpn13 would likewise serve only as an antagonist of degradation. However, Rpn13 seems to be a positive regulator in this process<sup>71,77,96</sup>. The ubiquitin-binding data suggests that, by binding to Rpn13, Uch37 is brought into the proximity of a high-affinity complex involving Ub, and thus activated. Other mechanisms of activation by Rpn13 may also operate, and it has been postulated that Uch37’s C-terminal region acts as a competitive inhibitor, as its deletion enhances Uch37 deubiquitinating activity and binding of Rpn13 to this C-terminal domain of Uch37 relieves its autoinhibition. Additionally, the presence of Rpn13 might just enhance Uch37 catalytic activity principally by serving as a bridge between Uch37 and Ub.

Altogether the data<sup>1,99</sup> confirm that Rpn13 binds diUb elements of K48-linked Ub chains and engages interactions with K48 and G76 that are linked by an isopeptide bond as shown by NMR spectroscopy. We propose that by abrogating interactions between the Ub moieties of a diUb unit and by changing the orientation of the isopeptide linker region, the cleavage site of Uch37, Rpn13 facilitates substrate recognition by Uch37. This model is similar to one proposed for substrate ubiquitination, whereby an E3 enzyme can align the E2-Ub thiolester in a favourable orientation<sup>146</sup>. Whereas Uch37 exhibits distal end cleavage specificity, the interaction of Rpn13 with the Ub linker is likely to place it preferentially at the penultimate,

rather than most distal Ub subunit of a chain (Figure 30). This localisation of Rpn13 to the more proximal Ub subunit is supported by NMR data using diubiquitin in which either the proximal or distal subunit is  $^{13}\text{C}$  labelled. This was then monitored by  $^1\text{H}$ ,  $^{13}\text{C}$  HMQC experiments as hRpn13 Pru was added and indeed, hRpn13 Pru binds to the proximal and not the distal subunit. Consequently, the preferred cleavage site when Rpn13 is bound to the penultimate subunit would be at the distal end of the chain (Figure 1.30).



**Figure 1.30:** Model for Rpn13-mediated activation of Uch37 by Ub binding. Ub-P, penultimate Ub, Ub-D; distal Ub. Proposition that Uch37-mediated deubiquitination occurs through a 2-step processive mechanism, in which it cleaves first the distal (a) and then the initially penultimate (b) linkage as the chain remains stationary on the proteasome.

With the exception of yeast, Rpn13 homologues from all organisms contain a linker region which connects the PH domain with the C-terminal Uch37 binding domain. In contrast to the N-terminal parts of hRpn13 and hUch37, secondary structure prediction of the C-terminus of both proteins shows solely helical secondary structure elements. Indeed, coexpression of the two interaction partners in this study revealed that the interaction module is limited to these helical domains located at the end of both polypeptide chains. Moreover, the amino acid residues from 300 to 407 on Rpn13 and from 230 to 325 on Uch37 appear to be sufficient for interaction.

## 1.5 Materials and Methods

### 1.5.1 Materials

#### 1.5.1.1 Oligonucleotides

The following oligonucleotides were used as primers to amplify different constructs of Rpn13 and Uch37 from *Saccharomyces cerevisiae*, *Mus Musculus* and *Homo sapiens*. For the ligation of the amplified genes into the vectors pRSET–GST–PP, pGEX–6P–1 and pGEX–bicistron, all primers carried specific restriction sites. These sites were BamHI (forward and reverse) in case of pRSET–GST–PP, BamHI (forward) and EcoRI (reverse) for pGEX–6P–1. Primers used for cloning into pGEX–bicistron vector carried the restriction sites BamHI/EcoRI and NdeI/NotI for the open reading frames 1 and 2. Additionally, all primers were designed to have an overhang to ensure high cleavage efficiency. The primers are synthesized and purified from Metabion international AG.

<u>Name</u>		<u>Enzyme</u>	<u>Organism</u>
<u>Rpn13</u> mmRpn13_1_fo	5'–GCG GAT CCA TGA CGA CTT CAG GCG CTC TG–3'	BamHI	<i>Mus musculus</i>
mmRpn13_150_re	5'–GCG GAT CCT CAC AGT GCT GAA AGC TC–3'	BamHI	<i>Mus musculus</i>
mmRpn13_151_fo	5'–GCG GAT CCG GCG GTG AGG GTG GCC–3'	BamHI	<i>Mus musculus</i>
mmRpn13_407_re	5'–GCG GAT CCT CAG TCT AGA CTC ATA TCT TCT TC–3'	BamHI	<i>Mus musculus</i>
scRpn13_1_fo	5'– AAA AAA GGA TCC AGT ATG AGT TCA ACT GTA ATT AAA TTC –3'	BamHI	<i>Saccharomyces cerevisiae</i>
scRpn13_156_re	5'– AAA AAA GGA TCC TTA ATC TTG CAT ACT AAC ATC CAC–3'	BamHI	<i>Saccharomyces cerevisiae</i>
scRpn13_123_re	5'– AAA AAA GGA TCC TTA ATT GTA AAT TTC TTT ATC TTT CGC AC–3'	BamHI	<i>Saccharomyces cerevisiae</i>
hsRpn13_202_fo	5'– AAA AAA GGA TCC GGG AGC AGC TCC TCC TCC –3'	BamHI	<i>Homo sapiens</i>
hsRpn13_267_fo	5'– AAA AAA GGA TCC CTC CAG AGC ATC CTG GCC –3'	BamHI	<i>Homo sapiens</i>
hsRpn13_407_re	5'– AAACAAA GGA TCC TTA GTC CAG GCT CAT GTC CTC C –3'	BamHI	<i>Homo sapiens</i>

Rpn13_407_re_bicis	5'– AAA AAA <b>GCG GCC GCC</b> ATA TGT ATA TCT CCT TCT TAA AGT TAA ACA AAA TTA TTG TCG ACG GAG CTC GAA TTC TTA GTC CAG GCT CAT GTC CTC C–3'	NotI	<i>Homo sapiens</i>
<b>UCH37</b>			
hsUch37_6_fo	5'– AAA AAA <b>GGA TCC</b> GGG GAG TGG TGC CTC ATG–3'	BamHI	<i>Homo sapiens</i>
hsUch37_228_re	5'– AAA AAA <b>GAA TTC</b> TTA TTT TCT GTC AGA CAC AAT GGC –3'	EcoRI	<i>Homo sapiens</i>
hsUch37_325_re	5'– AAA AAA <b>GAA TTC</b> TTA AGC TTT CTT TGC GTT CTG TTT TTC–3'	EcoRI	<i>Homo sapiens</i>
hsUch37_234_fo_bicis	5'– AAA AAA <b>CAT ATG</b> ATA TAT GAG CAG AAG ATA GCA G–3'	NdeI	<i>Homo sapiens</i>
hsUch37_407_re-bicis	5'– AAA AAA <b>GCG GCC GCT</b> TAA GCT TTC TTT GCG TTC TGT TTT TC –3'	NotI	<i>Homo sapiens</i>

### 1.5.1.2 Vectors

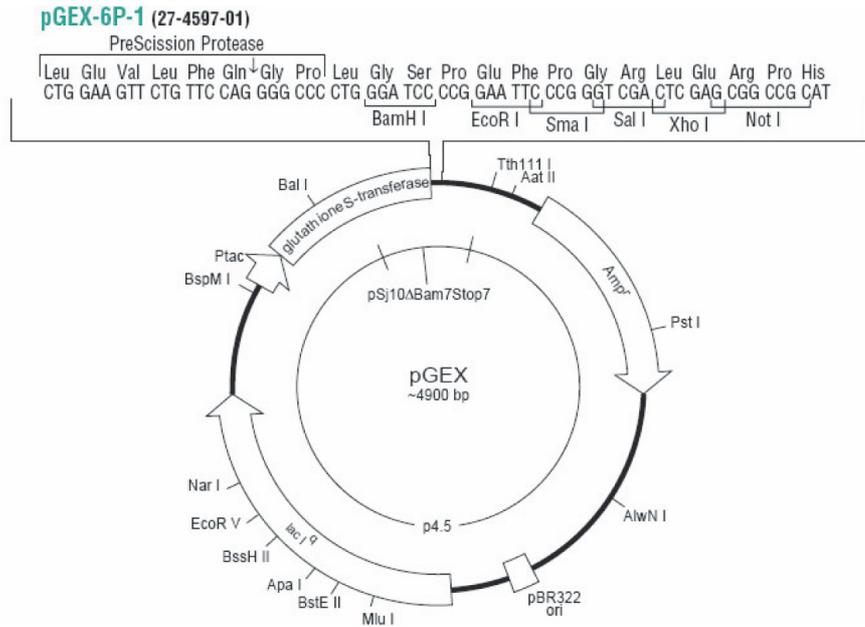
#### 1.5.1.2.1 pRSET–GST–PP

The expression vector pRSET–GST–PP was derived from the pRSET vector (Invitrogen) by cutting out the MCS (Multiple Cloning Site) with restriction enzymes NdeI and BamHI and insertion of the DNA encoding GST–tag followed by a PreScission Protease (PP) cleavage site. Rpn13 constructs (see Table 1.1, 1.3.1) were ligated with BamHI restriction sites into the pRSET–GST–PP vector, which carries a T7 promoter, ampicillin resistance (Amp<sup>r</sup>) and a lac operator, allowing gene expression to be induced by IPTG (isopropyl–b–D–thiogalactopyranosid). Finally, a protein expressed with this vector could be purified by affinity chromatography via the GST–tag which can be cleaved off with PreScission Protease.

#### 1.5.1.2.2 pGEX–6P–1 and pGEX–4T–1

The pGEX–6P–1 vector (Amersham Biosciences) was used for expression of Uch37 variants, which did not express in the pRSET system and also as a platform for construction of the bicistronic vector. Similar to pRSET–GST–PP vector, this vector contains a GST–tag followed by a PreScission Protease site for cleavage of the tag (vector map is shown in Figure 1.31). Uch37 constructs were inserted with the restriction sites BamHI and EcoRI of the MCS. The vector carries ampicillin resistance (Amp<sup>r</sup>) and expression is under the control of the *tac* promoter, which is induced by the lactose analogue IPTG.

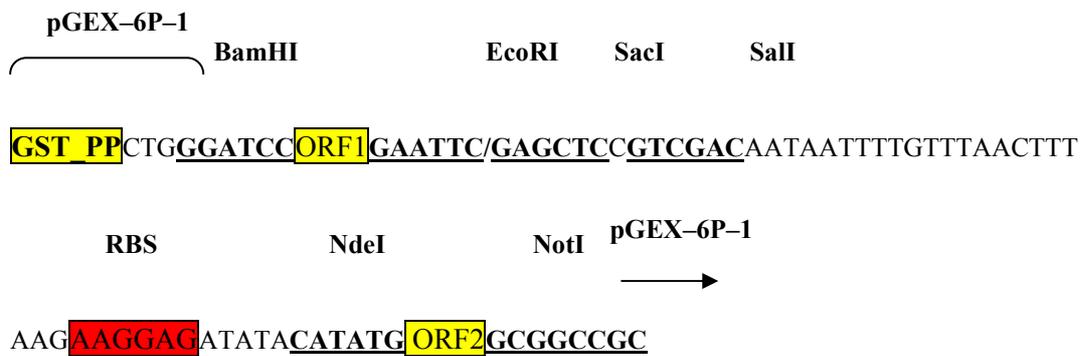
**Figure 1.31:** Vector map of expression vector pGEX-6P-1 (Amerham Bioscience).



The pGEX-4T-1 vector was used for expression of GST-fused tetraUb and different point mutants of Rpn13 Pru domain. It is very similar to pGEX-6P-1 except that it carries a thrombin cleavage site instead of the PP cleavage site. Different deletion mutants of hRpn2 (aa695-953, aa797-953, aa881-953, and aa902-953) were also subcloned into pGEX-4T-1 using BamHI restriction sites and additionally into pEGFP-C1 (Clontech) verified by sequencing. All pGEX vectors are also engineered with an internal *lacIq* gene. The *lacIq* gene product is a repressor protein that binds to the operator region of the *tac* promoter, preventing expression until induction by IPTG, thus maintaining tight control over expression of the insert.

### 1.5.1.2.3 Bicistronic vector system

Based on the pGEX-6P-1 vector (Figure 1.31), an expression vector with a bicistronic operon for simultaneous expression of two different proteins was constructed (Figure 1.32). Therefore, the Rpn13 gene was inserted with the restriction enzymes BamHI and NotI using a reverse primer, which contains several restriction sites, a second ribosomal binding site (RBS) and the second open reading frame (ORF2) resulting in a vector system with an ORF1 bordered by BamHI at the 5'-end and EcoRI at the 3'-end and an ORF2 bordered by NdeI and NotI. A GST-tag with PP cleavage site is preceding ORF1. In case of coexpression of the two proteins, Rpn13 and Uch37 were cloned into ORF1 and ORF2, respectively.



**Figure 1.32:** Inserted sequence of bicistronic expression vector *pGEX-bicistron*. *Rpn13* was cloned in the *ORF1* and carries a cleavable *GST*-tag, whereas *Uch37* was cloned into *ORF2*.

### 1.5.1.3 Bacteria

The following *E. coli* strains were used for molecular biological working and for the expression of the variants of *Rpn13* and *Uch37*. The strains were ordered from Stratagene, Heidelberg.

Strain	Genotype	Application
XL-1 blue strain a	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac{F' proAB lacI <sub>q</sub> ZM15Tn10(Tet <sup>r</sup> )} non inducible promoter	Preparation of high-quality plasmid DNA.
BL21-CodonPlus(DE3)-RIL	B F-ompT hsdS(rB- mB-) dcm+ Tet <sup>r</sup> galλ (DE3) endA Hte {argU ileY leuW Cam <sup>r</sup> }, extra copies of argU, ileY and leuW tRNA genes on a ColE1- compatible plasmid with canamycin resistance marker, protease deficiency, chromosomal T7-polymerase gene	Induced high-yield protein expression with, especially with T7 promoter.
B834(DE3)	F-ompT gal hsdSB (rB-mB-) met dcm lon _DE3	Methionine auxotroph expression host; labelling of proteins with selenomethionine

**Table 1.4:** Different *E. coli* strains which were used in this study.

### 1.5.1.4 Enzymes

Name	Company
Taq Polymerase	New England Biolabs, Schwalbach
Pfu Polymerase	Promega
Shrimp Alkaline Phosphatase	Roche
T4 DNA Ligase	New England Biolabs, Schwalbach
Restriction Enzymes (BamHI, EcoRI.....)	New England Biolabs, Schwalbach
Trypsin	SIGMA, Saint Louis, MI, USA
Chymotrypsin	SIGMA, Saint Louis, MI, USA
Carboxypeptidase Y	Worthington Biochemical Corp., Lakewood, NJ
PreScission Protease	self-made

**Table 1.5:** *List of used enzymes.*

### 1.5.1.5 Equipment and chemicals

ÄKTA Explorer, ÄKTA Purifier and ÄKTA Prime Chromatography Systems	Amersham Biosciences, Freiburg
Amicon Ultra 10K, 30K device	Millipore, Bedford, MA, USA
Autoklav Bioclav	Schütt
Autoklav Varioklav	H+P Labortechnik GmbH
Balance MC1 AC210P	Sartorius, Göttingen
Balance CPG22	Sartorius, Göttingen
Branson Sonifier 250	G. Heinemann
Centrifuges: Avanti J–25, Avanti J–20 XP	Beckman Instruments
Centrifuge 5810 R	Eppendorf
22 mm x 0.96 mm thick Plain circle cover slides	Hampton R., Laguna Niguel, CA,US
22 mm Circular Cover Slides (siliconized)	Jena Bioscience
Cups (1.5 ml and 2.0 ml)	Sarstedt
ThermoTube™ PCR Tubes 0.2 ml	Peqlab Biotechnology GmbH
Digital Camera	Sony
Econo–Column Chromatography Column	Biorad
Electrophoresis Power Supply	Consort
Electrophoresis Power Supply EPS 601	Amersham Biosciences
Freezer –20 °C	Liebherr

Freezer –80 °C	Forma Scientie, Life Sciences International GmbH
Fridge 4 °C	Liebherr
Fraction collector Frac 100	Amersham Pharmacia, Freiburg
GDS–7900 System Transilluminator (Agarose gel documentation)	UVP
Gel chambers	workshop, Institut für Physiologische Chemie, München
Gel Drying Film	Promega Corp., Madison, WI, USA
Glas ware	Schott, Duran
Grease	Hampton R., Laguna Niguel, CA, USA
Greiner 96 well plates	Greiner und Söhne, Frickenhausen
Incubators for molecular biology	Heraeus Christ
Incubator for crystallisation	RUMED
Microscope	Carl Zeiss AG, Oberkochen, Germany
Mikro 22 R centrifuge	Hettich
Mikrowave	Bosch
Mono Q HR 10/10	Amersham Pharmacia, Freiburg
Multichannel Pipette Research Pro 10 and 300	Eppendorf
Multitron II shaker	Infors AG, Bottmingen
Nanosep Centrifugal Devices (3 and 10K)	Pall Life Sciences (New York, USA)
NMR spectrometer	Varian
Overhead shaker	workshop, Institut für Physiologische Chemie, München
Mastercycler gradient (PCR)	Eppendorf
Petriplates	Greiner und Söhne, Frickenhausen
Peristaltic–Pump P–1	Amersham Biosciences
pH–Meter CG 822	Schott
Photometer Gene Quant pro	Amersham Biosciences
Photometer Ultrospec 10	Amersham Biosciences
Pipettes 2, 20, 100, 200, 1000 µl	Gisco
Primus (PCR)	MWG Biotech
PP–Test tubes, 15 and 50 ml (falcons)	Greiner
Quarz precision cuvettes 105.201–QS	Hellma Müllheim
Röntgen films NewRX	Fuji
Sterile benches	BDK, Genkingen

Superdex 200 HR 10/30	Amersham Pharmacia, Freiburg
Table centrifuge 5415 D	Eppendorf–Netheler, Hamburg
Thermomixer compact	Eppendorf
VDX™ 24 well plate (greased)	Hampton R., Laguna Niguel, CA, US
Vortex Mixer	Bender&Hobein, Zürich

***Unless stated otherwise all chemicals were from the following companies in p.A. quality.***

- Amersham Biosciences, Freiburg
- Applied Biosystems, Weiterstadt
- Fluka, Neu–Ulm
- Fermentas, St.Leon–Rot
- Merck, Darmstadt
- New England Biolabs, Schwalbach/Taunus
- Qiagen, Hilden
- Promega
- Roth, Karlsruhe
- Serva, Heidelberg
- Sigma Aldrich, Deisenhofen

### 1.5.1.6 Media and Buffers

All solutions and media used for microbiological work were sterilized either by autoclaving or filtration. Buffers were prepared using bidestillized water (ddH<sub>2</sub>O, Millipore).

NAME	PRESCRIPTION
<b><i>E. coli</i> Media</b>	
LB	10 g Trypton, 5 g Yeast Extract, 10 g NaCl, ad 1 l ddH <sub>2</sub> O (for culture plates ad. 1.5% w/v Agar)
<b>Selenomethionine labelling of proteins</b>	
New Minimal Medium	7.5 mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 8.5 mM NaCl, 55 mM KH <sub>2</sub> PO <sub>4</sub> , 100 mM K <sub>2</sub> HPO <sub>4</sub> , 20 mM glucose, 50 µg/l of each amino acid, 0.1% CaCl <sub>2</sub> , 0.1% FeCl <sub>2</sub> , 1 µg/l(w/v) trace elements (each), 10 mg/l thiamine, 10 mg/l biotin, 0.25 mM selenomethionine
<b>Preparation of M9 Minimal Medium</b>	
10xM9 salts	479 mM Na <sub>2</sub> HPO <sub>4</sub> , 220 mM KH <sub>2</sub> PO <sub>4</sub> , 86 mM NaCl
M9 Minimal Medium	10 % (v/v) 10xM9 salts, 1 mM MgSO <sub>4</sub> , 19 mM NH <sub>4</sub> Cl, 17 mM glucose, 0.1 mM

	CaCl <sub>2</sub> , 1% Gibco MEM Vitamin solution
<b>Competent Cells</b>	
electrocompetent	10% Glycerol (v/v)
chemical competent	0.1 M CaCl <sub>2</sub>
<b>Isolation of plasmid DNA</b>	
Resuspension buffer	50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNase
Lysis buffer	200 mM NaOH, 1.0% (w/v) SDS
Neutralization buffer	3 M potassium acetate, pH adjusted to 5.5 with acetic acid
<b>Agarose gels</b>	
5xSample Buffer	50 mM EDTA, 0.05% Bromphenolblue (w/v), 0.05% Xylencyanol (w/v), 30% Glycerol (v/v)
1xTAE – Electrophoresis-Buffer	40 mM Tris/Acetate, 20 mM Natriumacetat, 1 mM EDTA, pH 7.5
<b>Antibiotics</b>	
Ampicillin	100 mg/ml H <sub>2</sub> O
Chloramphenicol	30 mg/ml Ethanol
Tetracyclin	12.5 mg/ml Ethanol
<b>SDS-PAGE</b>	
4xSample Buffer	40% Glycerol (v/v), 240 mM Tris-HCl pH 6.8, 0.02% Bromphenolblue (w/v), 20% β-Mercaptoethanol (w/v), 8% SDS (w/v)
10xSDS Runningbuffer	248 mM Tris-Cl, 1.9 M Glycine, 10% SDS(w/v)
<b>Western Blot</b>	
Transfer buffer	20 mM Tris-Cl, 150 mM Glycine, 20% (v/v) Methanol, 0.08% (w/v) SDS
Ponceau S solution	0.2% (w/v) Ponceau S in 3% (w/v) Trichloroacetic acid
<b>Staining of Protein gels</b>	
Coomassie Stain	40% Ethanol (v/v), 10% Acetic acid (v/v), 0.2% Coomassie Brilliant Blue R-250 (w/v)
Coomassie Destain	30% Methanol (v/v), 10% Acetic acid (v/v)
<b>Immunodecoration</b>	
TBS	10 mM Tris-Cl pH 7.5, 135 mM NaCl
ECL reagent	0.1 M Tris-Cl pH 8.5, 2.5 mM 3-Aminophthalhydrazide, 0.4 mM P-cumaric acid, 0.0113% H <sub>2</sub> O <sub>2</sub>
<b>Limited proteolysis</b>	
Reaction Buffer	100 mM NaCl, 20 mM Tris-Cl pH 8.0
<b>Protein purification</b>	
Lysis Buffer	100 mM NaCl, 20 mM Tris-Cl pH 8.0, 1mM DTT
<b>Glutathione-Sepharose column</b>	
Wash Buffer	Lysis Buffer
Elution Buffer	Lysis Buffer plus PreScission Protease or plus 10 mM red. Glutathione
<b>TCA precipitation</b>	
Precipitant	72% Trichloroacetic acid

TCA wash buffer	100 mM Tris Tris–Cl pH 8.8, 70% (v/v) Ethanol, small amount (spatula tip) Bromphenolblue
<b>Mono Q Column</b>	
Low salt buffer	20 mM Hepes pH 7.0, 1 mM DTT
High salt buffer	Low salt buffer with 1 M NaCl
<b>Size Exclusion Chromatography</b>	
Gel filtration buffer	20 mM Tris–Cl pH 8.0, 40 mM NaCl, 1 mM DTT
<b>Crystallisation</b>	
Nextal Screen System	see manufacturer`s buffer composition table
Crystallisation condition for mRpn13 Pru	15 % (w/v) PEG 4000, 200 mM NaOAc, 1 mM DTT, 100 mM Tris–HCl, pH 8.5
<b>Protein overlay assay</b>	
Blocking buffer	10 mM Tris–Cl, pH 8.0, 150 mM NaCl and 0.1% (v/v) Tween–20, 0.1% ovalbumin (w/v)
<b>NMR spectroscopy</b>	
Buffer C	20 mM NaPO <sub>4</sub> pH 6.5, 30 mM NaCl, 2mM DTT, 0.1% (w/v) NaN <sub>3</sub> , and 10% (v/v) D <sub>2</sub> O
<b>Biochemical pull down assay</b>	
Thrombin cleavage buffer	20 mM Tris–Cl pH 8.4, 150 mM NaCl, 2.5 mM CaCl <sub>2</sub> , 1 mM DTT
Incubation buffer	Thrombin cleavage buffer plus 10% Glycerol and 1% Triton X–100.

**Table 1.6:** *Media and buffers.*

### 1.5.1.7 DNA Isolation and Preparation Kits

Name	Application
Pure Yield™ Plasmid MidiPrep System	Isolation of plasmid DNA from <i>E. coli</i> .
Wizard® SV Gel and PCR Clean Up System	Purification of PCR–products and extraction of DNA from agarose gel

Company: Promega Corp., Madison, WI, USA.

### 1.5.1.8 Columns

NAME	VOLUME	V <sub>0</sub> = void volume	LOADING RANGE	MMW–Range	TECHNIQUE
Mono Q 5/50 GL	1 ml		50 mg		Anion exchange
Superdex 200 HR 10/30	24 ml	7.77 ml		10 – 600 kDa	Size exclusion

**Table 1.7:** Columns.

### 1.5.1.9 Bioinformatics

DNA and protein sequences were obtained from the „National centre for Biotechnology Information” (<http://www.ncbi.nlm.nih.gov>) and SWISS-PROT (<http://www.expasy.ch>). Primary sequence predictions were obtained from PredictProtein. PredictProtein is an Internet service (accessible via SWISS-PROT portal) for sequence analysis and the prediction of protein structure and function. Users submit protein sequences or alignments and PredictProtein returns multiple sequence alignments, PROSITE sequence motifs, low-complexity regions (SEG), nuclear localisation signals, regions lacking regular structure (NORS) and predictions of secondary structure, solvent accessibility, globular regions, transmembrane helices, coiled-coil regions, structural switch regions, disulfide-bonds (DISULFIND), sub-cellular localization, and functional annotations. For all services, users can submit their query either by electronic mail, or interactively from World Wide Web. The secondary structure prediction was made by PROFsec, an advanced method of PHDsec. The computation of various physical and chemical parameters for a given protein by entering its sequence was performed by the ProtParam tool (SWISS-PROT portal). The computed parameters include the molecular mass, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY)

## 1.5.2 Methods

### 1.5.2.1 Primer design

The length of the primers varied from 26 to 94 bases and the primers had at least one G or C at the 3'-end to obtain better binding and elongation. They were designed to have a melting temperature of approximately 60°C to allow sufficiently high annealing temperatures and contained the required restriction sites. In case of the pRSET-GST-PP, pGEX-6P-1 and pGEX-bicistron vector (for ORF2 a start codon was necessary) no start codon had to be introduced because it is already included in the vector. For all constructs the reverse primers had to contain a stop codon (TGA) before the restriction site (see Section 1.5.1.1).

### 1.5.2.2 Methods in molecular biology

#### 1.5.2.2.1 Polymerase chain reaction (PCR)

Since the beginnings of the PCR new applications were constantly found for this method, which allows amplifying DNA sequences of any sequence. For the PCR two oligonucleotide primers complementary to sequences at the ends of the required DNA, a mixture of all four dNTP's and a for high temperature stable polymerase with the suitable buffer and the right concentrations of salt are required. After the denaturizing of the dsDNA (denaturizing), the primers hybridize with the complementary DNA (annealing) and the polymerase elongates the ssDNA construct starting with the 3'-end of the primer (elongation). Thus, the number of dsDNA constructs increases exponentially with each cycle. Either the enzymes Taq or Pfu were used as polymerases.

#### Reaction mixture:

	Substance	Volume
1)	DNA template (approx. 0.5 µg/µl)	1 µl
2)	Forward primer (100 pmol/µl)	0.5 µl
3)	Reverse primer (100 pmol/µl)	0.5 µl
4)	dNTP's (5 mM per nucleotide)	2 µl
5)	Taq or Pfu buffer (10x)	5 µl
6)	ddH <sub>2</sub> O sterile	40 µl
7)	Taq (5 U/µl) or Pfu polymerase (3 U/µl)	1 µl

The reaction mixture was pipetted into 500 µl PCR tubes and PCR reaction was carried out in a thermocycler.

PCR program:

	Temperature	Time
1)	95 °C	10 min
2) a)	92 °C	1 min
b)	58 °C	1 –2 min
c)	72 °C (Taq)/ 75 °C (Pfu)	1 min
Step 2 was repeated thirty times.		
3)	72 °C (Taq)/ 75 °C (Pfu)	10 min

Annealing temperature (2b) and elongation time (2c) of the reaction can be varied and had to be adjusted to the respective reaction. PCR products were analyzed by agarose gel electrophoresis and purified using the Wizard<sup>®</sup> SV Gel and PCR Clean Up System. To check for positive clones by PCR on *E. coli* colonies, colonies were picked with a sterile tip and resuspended in 10 µl of sterile ddH<sub>2</sub>O. The reaction volume was then only half of the above mixture, Taq was used as polymerase and the resuspended colony as template. The forward primer of the insert and a reverse primer annealing in the vector served as primers. The mammalian expression vector pcDNA3.1/myc–his containing the full–length cDNA of mRpn13 was kindly provided by Dr. Lamerant, Centre de Biophysique Moléculaire, France,<sup>147</sup> and served as template for cloning of mRpn13 constructs. A pGEX–4T–1 vector containing DNA of human Uch37 (kindly provided by Ivan Dikic, Institute of Biochemistry II, Goethe University Hospital, Frankfurt (Main)) was used as a template for all Uch37 cloning.

#### 1.5.2.2.2 Digestion of DNA with restriction endonucleases

For analytical and preparative purposes plasmid DNA and PCR products were digested with restriction endonucleases, which digest the DNA at specific, mostly palindromal sequences. All used enzymes produced an overhang, thus only ligation of sticky ends occurred. The incubation time, temperature and the buffer conditions were applied according to the manufacturer's recommendations. The obtained fragments were analyzed by agarose gel electrophoresis. Before the final sequencing of the resulting plasmid DNA, an additional restriction digestion was carried out to see whether the clone contained the required insert.

PCR product	Vector	Control Restriction
5 µl PCR construct	3 µg vector	2 µl vector
5 µl BSA (10x)	5 µl BSA (10x)	2 µl BSA (10x)
0.5 µl Enzyme 1	0.5 µl Enzyme 1	0.5 µl Enzyme 1
0.5 µl Enzyme 2	0.5 µl Enzyme 2	0.5 µl Enzyme 2
5 µl 10xBuffer	5 µl 10xBuffer	2 µl 10xBuffer
	ad 50µl H <sub>2</sub> O steril	13 µl H <sub>2</sub> O sterile

**Table 1.8:** Restriction digestion of PCR product or vector as well as control restriction.

All constructs for insertion into the pRSET–GST–PP/pGEX–4T–1 vector as well as the vector were digested solely with BamHI, whereas all constructs for ligation into pGEX–6P–1 vector were digested with endonucleases BamHI/EcoRI. According to Figure 1.27, constructs for ORF1 and ORF2 of the pGEX–bicistron vector were digested with the enzymes BamHI/EcoRI and NdeI/NotI, respectively. The analysis of the desired DNA fragments and cleaved vectors was done by gel electrophoresis (see 5.2.2.3).

Self ligation of the linearized plasmid was suppressed by dephosphorylating the 5'–end using salmon alkaline phosphatase (SAP). Therefore 1 µl of SAP (10 U/µl) and the appropriate amount of 10 fold SAP buffer were added to the purified, digested vector and incubated for one hour at 37 °C. The reaction was stopped through gel electrophoresis or purification with Wizard<sup>®</sup> SV Gel and PCR Clean Up System. This reaction was not used in all cases because the dephosphorylation could reduce the number of positive clones after transformation.

### 1.5.2.2.3 Ligation

If PCR products and vector are restricted with the same restriction enzymes the desired DNA construct can be inserted into the open plasmid and ligated with T4 DNA ligase. Linearized DNA vector and an approximately 10 fold molar excess of DNA fragment to be inserted were incubated with T4 DNA Ligase and buffer. The ligation mixture was prepared on ice and incubation was performed either at RT for 1 h or at 4°C over night. The self ligation of the vector served as negative control. 0.5 to 1 µl of this mixture were used for transformation into competent *E. coli* cells.

Ligation		Negative-Control	
1 $\mu$ l	vector	1 $\mu$ l	vector
10 x molar excess	PCR-construct	16 $\mu$ l	H <sub>2</sub> O sterile
2 $\mu$ l	10xLigase-buffer	2 $\mu$ l	10xLigase-buffer
1 $\mu$ l	T4-DNA-Ligase	1 $\mu$ l	T4-DNA-Ligase

**Table 1.9:** *Ligation mixture.*

#### 1.5.2.2.4 DNA agarose gel electrophoresis

The agarose gel electrophoresis was used to separate and analyze DNA fragments between 0.1 and 8 kb. Voltages were in the range of 80 to 120 V and the concentration of agarose between 1–2%. Detection was done by the DNA intercalating dye ethidium bromide which can be observed in the ultraviolet. In case of 1% (2%) agarose gel, 1 g (2 g) of agarose was solved under heat in 100 ml 1xTAE-buffer and supplemented with 1:10000 fold dilution of ethidium bromide (Sigma). The rigid gel was placed in a gel chamber and covered with TBE-buffer. Samples were mixed with 5 x sample buffer and loaded onto the gel. For the marker a 1kb or a 100bp DNA ladder (New England Biolabs) was used. DNA constructs and linearized vectors that were utilized for ligation were excised from the gel under UV light ( $\lambda = 254$  nm) and purified with the Wizard<sup>®</sup> SV Gel and PCR Clean Up System.

#### 1.5.2.2.5 Competent *E. coli* cells

##### 1.5.2.2.5.1 Chemical competent *E. coli* cells

500 ml LB medium supplemented with tetracycline (XL1blue) or chloramphenicol (BL21–(DE3)–RIL) were inoculated with 5 ml of over-night culture of the desired *E. coli* strain. The cells were grown at 37°C to OD<sub>600</sub> of 0.4. After incubation on ice for 15 minutes cells were centrifuged for 15 minutes at 2580 x g (4000 rpm, JA10 rotor) and 4°C. All following steps were done at  $\leq 4^\circ\text{C}$ . The pellet was resuspended in 40 ml 0.1 M CaCl<sub>2</sub>, incubated on ice for 30 min and harvested by centrifugation at 2580 x g and 4°C for 15 min. This pellet was resuspended in 20 ml 0.1 M CaCl<sub>2</sub> supplemented with 4 ml of glycerol. After an additional incubation on ice for 2 h, the suspension was dispensed in 200  $\mu$ l aliquots and after shock-freezing in liquid nitrogen stored at  $-80^\circ\text{C}$ .

#### 1.5.2.2.5.2 Electro-competent *E. coli* cells

500 ml LB media supplemented with tetracycline (XL1blue) or chloramphenicol (BL21–(DE3)–RIL) were inoculated with 5 ml of over-night culture of the desired *E. coli* strain. The cells were grown at 37°C to OD<sub>600</sub> of 0.5 and then harvested by centrifugation for 15 minutes at 2580 x g (4000 rpm, JA10 rotor) and 4°C. All following steps were done at ≤4°C. The pellet was washed three times with 500, 250 and 50 ml 10% (v/v) glycerol. After the washing steps, cell pellet was resuspended in 500 µl 10% (v/v) glycerol, dispensed in 40 µl aliquots and after shock-freezing in liquid nitrogen stored at –80°C.

#### 1.5.2.2.6 Transformation

Transformation of plasmid DNA into chemical competent *E. coli* cells was mostly done with the heat shock method. Therefore 100 µl of the competent cells were thawed on ice, mixed with 4–10 µl of the ligation suspension (or 1 µl of purified plasmid) and incubated on ice for 10–30 minutes. Afterwards, cells were heated at 42°C for 1 min. After a short recovery on ice for 2 minutes, 500 µl of fresh LB medium were added and the mixture was shaken at 37°C for 45 min. Finally, the bacteria were harvested by centrifugation, resuspended in 200 µl of fresh LB media and plated on a selective media (LB/Amp<sup>r</sup> for all used vectors). The plates were incubated at 37°C over night.

For the transformation of plasmid DNA into electrocompetent cells 0.5–1 µl of ligation suspension (or purified plasmid) were added to 40 µl of thawed competent cells. The suspension was transferred to a pre-cooled cuvette and the cuvette was introduced in an electroporation apparatus (settings: 2.0 kV, 400 Ω, 25 µF, time constant 8–9 ms). After a brief application of a high electric voltage, the suspension was diluted with 800 µl LB-medium, and shaken for 30–60 min at 37°C. The experiment proceeded according to the transformation into chemical competent cells (see above).

#### 1.5.2.2.7 Isolation of plasmid DNA from *E. coli*

To analyse whether the PCR constructs have been properly inserted into the vector, the plasmid DNA was isolated from XL1blue cells and digested with restriction enzymes. Alternatively, *E. coli* colonies were tested for the right insert by PCR as described in 5.2.2.1. To isolate plasmid DNA, colonies were picked with pipet tips from the selective media plate,

transferred into 2 ml LB media containing the suitable antibiotic (at a dilution of 1:1000) and placed into a 37°C-shaker over night.

The cultures were transferred to a 2 ml reaction tube, harvested by centrifugation for 1 min at 16,000 g. Plasmid DNA was isolated according to the following procedure:

- 1) Resuspension of cell pellet in 300 µl of resuspension buffer.
- 2) Addition of 300 µl of lysis buffer and incubation at RT for 5 min.
- 3) Neutralization with 300 µl of neutralization buffer
- 4) Centrifugation at 21900 x g for 10 min.
- 5) Precipitation of DNA by mixing supernatant of centrifugation with 600 µl isopropanol.
- 6) Centrifugation at 21900 x g for 10 min and 4°C.
- 7) Washing with 500 µl of ice-cold 70% ethanol.
- 8) Drying of DNA pellet at 50°C for at least 10 min.
- 9) Dissolving of DNA in 30 µl of sterile ddH<sub>2</sub>O.

For large scale plasmid DNA isolation from 50 ml of culture media the Pure Yield™ Plasmid MidiPrep System from Promega was used.

#### **1.5.2.2.8 Mutagenesis**

Point mutants within full-length, wild-type mRpn13 R27A, R92A, F107A, F107S, W108A, M109A, M109C, F107A/W108A, S55E, L56A, I75R, F76R, D78N, D79N, F98R, G101F, K103R, R104A were created by *in vitro* mutagenesis using the QuickChange Site-directed Mutagenesis Kit (Stratagene) and verified by sequencing.

#### **1.5.2.2.9 DNA sequencing**

The concentration of isolated plasmid DNA was determined at 260 nm using a photometer and send to Eurofins Medigenomix GmbH for sequencing.

### 1.5.2.3 Methods in protein biochemistry

#### 1.5.2.3.1 Gene expression

After transformation of the desired plasmid DNA in *E. coli* BL21–(DE3)–RIL cells, 100 ml of LB media containing ampicillin (100 µg/ml) and chloramphenicol (30 µg/ml) were inoculated with the transformed cells (500 µl suspension, see 5.2.2.6) and incubated over night in a 37°C-shaker. Three litres of LB<sub>Amp,Cam</sub> medium were inoculated with the preculture at a ratio of 1:40 and incubated at 37°C. After cell density reached OD<sub>600</sub> 0.6 to 0.8, expression was induced with 0.5 mM IPTG (end concentration). The cells were incubated at 24°C over night and harvested by centrifugation at 4°C and 2580 x g (4000 rpm, JLA 8.1000 rotor) for 10 min. Finally the supernatant was discarded and the resulting pellet was gently resuspended in cold lysis buffer at 1/100 of the culture volume. For storage, the suspension was frozen and stored at –20 °C. For test expressions 100 ml of LB media as well as different expression times and temperatures were used (normally 3 h at 37°C, O/N at 37°C and O/N at 24°C).

#### 1.5.2.3.2 Cell lysis

The thawed suspended cells were disrupted on ice using an appropriately equipped sonicator for the suspended volume (parameters: big tip, pulse maximum and duty cycle 60) for 10 min. After centrifugation at 24000 x g (20000 rpm in Beckman JA20 rotor) and 4°C for 20 min, the supernatant including the protein of interest was used for further purification steps like affinity chromatography. An aliquot of the cleared lysate and the remaining pellet, which was treated with 8 M urea, were saved for analysis by SDS–PAGE.

#### 1.5.2.3.3 Purification of GST-fused proteins

GST–fused proteins were purified from bacterial lysate by affinity chromatography using Glutathione Sepharose 4B (Amersham Biosciences). Therefore, an appropriate amount of Glutathione Sepharose 4B suspension (mostly 5 ml for 3 l culture volume) was transferred to a Biorad column and washed with 50 ml of cold lysis buffer. The cleared bacterial lysate was mixed with the equilibrated beads and shaken at 4°C for 1 h under skaking. After this incubation the mixture was transferred to the column and washed with 4 x 50 ml of cold lysis buffer. The flow through and the washing fractions were collected and analyzed by SDS–

PAGE analysis. Recombinant protein was eluted from glutathione beads by cleavage with PreScission Protease at 30 °C for 1 h in 10 ml of lysis buffer (alternatively over night at 4°C). To analyze protein, which is still bound to the Glutathione Sepharose after treatment with PreScission Protease, the beads were incubated for 5 min with lysis buffer supplemented with 10 mM of reduced glutathione. Samples for SDS–PAGE analysis were taken of both elutions. If PreScission Protease was not able to cleave the fusion protein on the resin, the protein was eluted with reduced glutathione and the GST–tag was removed in solution. In this case, the protein was separated from GST via anion exchange chromatography.

#### **1.5.2.3.4 Anion exchange chromatography (MonoQ)**

The principle of ion exchange chromatography is the interaction between charged proteins and the contrarily charged groups of the column matrix. The binding of the proteins to the anion exchange column occurs under low salt conditions and the bound protein can be eluted under high salt conditions. Therefore, eluted protein from the Glutathione Sepharose column was loaded on the equilibrated MonoQ column in low salt buffer using an Äkta Chromatography System, washed with 5 CV (column volumes) of low salt buffer and eluted with increasing salt concentration using a linear gradient spanning 25 CV. The purification was monitored by UV spectrum at 280 nm and collected fractions containing the eluted protein were analyzed by SDS–PAGE.

#### **1.5.2.3.5 Size exclusion chromatography**

The protein of interest was further purified in gel filtration buffer on a Superdex 200 gel filtration column. Size exclusion chromatography is used to separate proteins according to their size. The column material consists of a cross–linked network of sugar molecules so that larger proteins migrate faster through the column than smaller ones, additionally depending of the shape of the protein. Thus, the elution volume ( $V_e$ ) from the column is a characteristic property of each protein. Eluted protein fractions from glutathione sepharose or anion exchange columns were pooled and concentrated by centrifugation using Amicon Ultra devices (Millipore). 500  $\mu$ l of the concentrated protein solution were injected on the preequilibrated Superdex 200 column via an Äkta Chromatography System. After elution with 1.2 CV, the target protein containing fractions were pooled and either purified by an additional gel filtration run or directly used for crystallisation trials. Through comparison with

marker proteins which were run on Superdex 200 under the same conditions as the purified protein, this method also gives information about the approximal size and oligomerisation of the examined protein.

### 1.5.2.3.6 Expression and purification of selenomethionine labelled protein

For selenomethionine labelling Rpn13 Pru was overexpressed in the methionine auxotroph *E. coli* strain B834(DE3) (Novagen). The gene expression took place in New Minimal Medium (NMM)<sup>148</sup> supplemented with selenomethionine and ampicillin. All stock solutions for the medium were either autoclaved or sterile filtered. For the 10x amino acid stock, 0.5 g of each amino acid except Tryptophan and Tyrosin were mixed. Tryptophan and Tyrosin had to be dissolved in 1 M HCl. The New Minimal Medium was prepared as follows:

#### Preparation of 1 l New Minimal Medium (5.1.6):

1) Mix together in autoclaved Fernbach culture flasks:

7.5 ml	1 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	
1.7 ml	5 M NaCl	
55 ml	1 M KH <sub>2</sub> PO <sub>4</sub>	
100 ml	1 M K <sub>2</sub> HPO <sub>4</sub>	
9 ml	40% glucose	
1 ml	1 mg/ml CaCl <sub>2</sub>	
1 ml	1 mg/ml FeCl <sub>2</sub>	
1 ml	10 mg/ml thiamine	
10 ml	10 mg/ml biotin	soluble at slightly basic pH (with NaOH)
9 ml	40% glucose	
1 ml	100x trace element stock	Ca <sup>2+</sup> , Zn <sup>2+</sup> , Mn <sup>2+</sup> , MoO <sub>4</sub> <sup>-</sup> in H <sub>2</sub> O
0.85 g	10x amino acid stock	
10 ml	0.5% Tryptophan	dissolved in 1 M HCl
10 ml	0.5% Tyrosin	dissolved in 1 M HCl
ad 1 l	sterile ddH <sub>2</sub> O	

2) Set pH between 7.4 and 7.0.

3) Add 1 ml of 1000x ampicillin.

4) Add 50 mg of selenomethionine.

100 ml of overnight cultures of LB medium were harvested at 3000 x g for 10 min, washed with 50 ml of NMM supplemented with 100 mg/l Ampicillin. Pellets were resuspended in 3 L of NMM, supplemented with 0.25 mM selenomethionine and 100 mg/l ampicillin. After cells were grown at 37 °C to an OD<sub>600nm</sub> of ~0.6, expression was induced with 0.5 mM IPTG. Incubation was continued at 24°C over night. The purification of selenomethionine-labelled protein was performed according to the same procedure as for unlabelled protein.

#### **1.5.2.3.7 Transfer of proteins to nitrocellulose membrane (Western-blot)**

Proteins separated via SDS-PAGE were transferred onto nitrocellulose membrane using the semi-dry blotting method<sup>149,150</sup>. The gel, the membrane, and six sheets of Whatman filter paper (3MM) were incubated in transfer buffer. Three sheets of filter paper were placed on the anode electrode followed by the membrane and on top of it the gel. This was covered with other three filter papers and with the cathode electrode. The transfer was performed at 200 mA for 1.5 h. To verify transfer efficiency, the nitrocellulose membranes were reversibly stained with Ponceau S solution. Afterwards, the membrane was immunodecorated.

#### **1.5.2.4 Detection and quantitative determination**

##### **1.5.2.4.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)**

Proteins can be separated according to their molecular mass by discontinuous sodiumdodecylsulfat polyacrylamide gelelectrophoresis (SDS-PAGE) as described<sup>151</sup>. Smaller proteins migrate faster through the gel towards the negative pole than larger ones. Although the electrophoretic mobility depends not only on molecular mass but also on the shape and the charge of the protein, the molecular mass could be estimated through comparison to a marker because the protein solution is denaturated under reducing conditions ( $\beta$ -mercaptoethanol) by adding SDS (it binds to the unfolded peptide and has a high charge). Two different electrophoresis systems were used. Normally, big gels were used (running gel: 9 x 15 x 0.1 cm; stacking gel: 1 x 15 x 0.1 cm) and in some cases Mini gels (Mini-PROTEAN

II, Bio–Rad) (running gel: 7 x 7.2 x 0.075 cm; stacking gel: 1 x 7.2 x 0.075 cm). The concentration of acrylamide and bis–acrylamide in the separating gel was chosen according to the molecular size of the proteins to be separated (mostly 16 or 17.5%). For the big gels, glass plates of 160x140 mm and spacers of 1 mm thickness were used. The samples were dissolved in appropriate volumes of 4 x sample buffer and incubated at 95°C for 5 min, before loading. The electrophoresis was performed at 25–45 mA for 1–3 h (30–60 min for mini gels) in 1 x SDS running buffer. Protein molecular mass standards were used in each gel.

Bottom gel	
Separating gel	Separating gel 16% (w/v) acrylamide, 0.16–0.33% (w/v) bis–acrylamide, 380 mM Tris–HCl (pH 8.8), 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.035% (v/v) TEMED.
Stacking gel	Stacking gel 5% (w/v) acrylamide , 0.1% (w/v) bis–acrylamide, 60 mM Tris–HCl (pH 6.8), 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.035% (v/v) TEMED

**Table 1.10:** *Composition of a SDS–PA gel.*

#### 1.5.2.4.2 Staining SDS–PA gels with Coomassie brilliant blue

To visualize the protein on the gel, Coomassie Brilliant Blue dye was used, which allows detection of protein amounts down to 0.1 µg. After SDS–PAGE and removal of the bottom and stacking gels, the separating gel was heated in Coomassie stain for 1 min using a microwave for, incubated for 20–40 min at room temperature, subsequently washed with water and destained in Coomassie destain until distinct bands became visible against a clear background. The gel was scanned and then dried overnight between two gel–drying films.

#### 1.5.2.4.3 Immunodecoration

To visualize the immobilized proteins on PVDF or nitrocellulose membrane, immunodecoration with specific antibodies was carried out. Membranes were first incubated for 1 h in 5% (w/v) milk powder in TBS to block all non–specific binding sites. The immunodecoration was done for 1 h at RT or over night at 4°C, with specific antiserum

(1:200 to 1:1000 dilutions in milk/TBS). The membrane was then washed 3 times (each wash lasts 5 min), with TBS, TBS/0.05% (w/v) Triton X-100 and again with TBS, and incubated for 1 h with horseradish peroxidase coupled to secondary goat anti-rabbit-IgG (diluted 1:10.000 dilutions in milk/TBS). The membrane was again washed (as above) and treated with ECL reagents ( $H_2O_2$  was added fresh). The luminescence reactions were detected with Röntgen films.

#### **1.5.2.4.4 Determination of the protein concentration by Bradford assay**

The Bradford assay is based on the fact that the absorption of Coomassie-Brilliant Blue dye shifts from 465 nm to 595 nm when it binds to protein. Its sensitivity allows to measure protein concentrations between 0.1 mg/ml and 30 mg/ml per ml Bradford solution. Protein concentration was determined using the “Bio-Rad-Protein assay” reagent<sup>152</sup>. Protein solution was mixed with 1 ml reagent (1:5 dilution) and incubated for at least 5 min at RT. To increase the accuracy of the assay six samples were measured (twice 2.5  $\mu$ l, 5  $\mu$ l and 7.5  $\mu$ l of the diluted protein solution). The absorbance was measured at 595nm using a 10 mm path length microcuvette. Finally, the protein concentration was calculated according to a standard curve (Bovine from IgGs (Biorad) was used as standard protein).

#### **1.5.2.4.5 Determination of the protein concentration by UV spectroscopy**

Protein concentration could be also determined by measuring the absorbance at 280 nm using a photometer. Therefore, the extinction of an appropriate dilution of 150–200  $\mu$ l of the protein solution was measured in a quartz cuvette with the buffer solution as blank value. The theoretical extinction coefficient was calculated by the ProtParam tool (5.1.9) and used for the determination of the protein concentration.

#### **1.5.2.4.6 Concentration of proteins**

For crystallization high protein concentrations were required. Protein was concentrated by centrifugation using Amicon Ultra 10K or 30K devices or Nanosep Centrifugal Devices (3 and 10K) according to the manufacturer’s instructions.

#### **1.5.2.4.7 Protein precipitation**

Sometimes the concentration of a protein fraction was less than is required for SDS–PAGE. Therefore, the protein was precipitated with 1/5 volume of 72 % trichloroacetic acid (TCA), incubated on ice for 10 min and centrifuged at 14000 rpm for 1 min. The precipitate was washed with 0.5 ml of TCA wash buffer, and dried at 56 °C for 10 min. The precipitated protein was dissolved in 1 x sample buffer and analysed by gel electrophoresis.

#### **1.5.2.5 Limited proteolysis**

Trypsin, Chymotrypsin and Carboxypeptidase Y were used for the proteolytic assay. For the digestion 10 µg of protein were mixed each with different concentrations (100 ng, 50 ng, 25 ng, 12.5 ng) of the respective protease in reaction buffer (total volume of 30 µl) and incubated at 25 °C for 20 min. The reaction was stopped by addition of 10 µl of 4 x sample buffer and samples were analyzed by SDS–PAGE. The experiment shows the stability of a protein variant to proteolytic treatment and degradation products might appear, which represent a stable domain of the protein.

#### **1.5.2.6 Mass spectrometry**

Promising fragments which were found by limited proteolysis were produced by protease treatment as described in 5.2.5 and analysed further by mass spectrometrical analysis. Therefore the enzymatic reaction was stopped by addition of 5 mM PMSF and the solution was sent to the “core facility” in the Max Planck Institute for Biochemistry in Munich. The determination of the molecular mass of the proteolytic fragment mostly allowed localization of the putative cleavage site and mapping of the domain corresponding to the degradation band in the gel.

#### **1.5.2.7 Crystallization**

Highly purified and concentrated protein was subjected to crystallisation trials. Initial screening for the right conditions was done by the sitting drop method at 20 °C using the Nextal Screening system. Therefore, 80 µl of each screening solution were transferred with a

multichannel pipette from the master block to the reservoir well of a 96well plate (if DTT was necessary for crystallization, 1  $\mu$ l of 100 mM DTT was added to each well). In the next step 1  $\mu$ l of the protein solution (for the first screens, protein concentration was adjusted between 5 and 20 mg/ml) was provided into the sitting drop well using a dispenser and 1  $\mu$ l of each reservoir solution was added and mixed. After sealing with tape the plate was incubated at 20°C and analyzed under the microscope in regular periods. For each protein several plates using different screens from Nextal (e.g classics suite, PH clear I and II suites) were set up. Normally, crystals grew after several days and to improve crystal quality, fine screens to optimize the conditions were performed by hanging drop vapour diffusion. 1–2  $\mu$ l of the concentrated protein solution were mixed on a plastic cover slide with the same volume of reservoir solution. The slide was placed over a well containing 300  $\mu$ l reservoir solution (if necessary supplemented with 1 mM DTT) on a 24–well plastic plate and hermetically sealed with grease. Additive screens from Hampton Research were used to improve crystal growth and quality by addition of 0.28  $\mu$ l of each additive to the hanging drop solution.

Rpn13 Pru crystals grew within two days to the size of 500 x 200 x 200  $\mu$ m<sup>3</sup>. The drops contained equal volumes of protein (5 mg/ml) and reservoir solution (15 % (w/v) PEG 4000, 200 mM NaOAc, 1 mM DTT, 100 mM Tris–HCl, pH 8.5). Crystals of the SeMet derivative occurred under the same conditions.

### 1.5.2.8 Structure determination

Before exposure to X–rays, crystals were soaked in a solution of mother liquor including 15% PEG 400 (cryo protectant) for 5 min and subsequently frozen in a stream of cold nitrogen gas at 100 K (Oxford Cryosystems). Single anomalous dispersion (SAD) methods were performed using synchrotron radiation at the BW6 beamline at the DESY–centre in Hamburg, Germany. Native data to 1.7 Å resolution were collected (Table 1.2) using a MarCCD detector. Data were processed with DENZO and SCALEPACK<sup>153</sup>. The space group of Rpn13 Pru was P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with unit cell dimensions of a=42.4 Å, b=56.2 Å, c=63.2 Å. Phase determination was performed by crystal structure analysis of the Rpn13 Pru–SeMet–derivative. SAD data were collected at 2.7 Å resolution using absorption peak wavelengths (0.9793 Å). Two Se–sites were localized with SHELXD<sup>100</sup>. Subsequent phasing with MLPHARE (The CCP4 suite: programs for protein crystallography) and solvent flattening with DM<sup>101</sup> resulted in an interpretable electron density map, which was traced and phase extended to the native data set at 1.7 Å resolution using ARP–WARP<sup>102</sup>. The model has been completed via the interactive three–dimensional graphic programs MAIN<sup>103</sup> and refined with

REFMAC5<sup>104</sup>. Temperature factors were refined with restraints between bonded atoms and between non-crystallographic symmetry related atoms.  $R_{\text{crys}}$  and  $R_{\text{free}}$  have been refined to 0.202 and 0.227, respectively<sup>105</sup>. The final model comprises fully defined residues 21–131 of Rpn13 Pru, whereas the N- and C-terminal parts of the molecule are structurally disordered.

### 1.5.2.9 Protein–lipid overlay assay

The protein–lipid overlay assay was performed to test Rpn13 Pru domain for binding to lipids, especially to phosphatidylinositolphosphates (PIPs). Therefore, PIP Strips<sup>TM</sup> membranes (Molecular Probes) containing 100 pmol of phospholipids were blocked in blocking buffer for 1 h at RT. The membrane was then incubated for 1 h with 0.1 µg/ml of purified GST-tagged protein in the same solution. After washing 3 times over 30 min, the membranes were incubated for 1 h with 1:1000 dilution of anti-GST polyclonal serum (Sigma). The membranes were washed as before, then incubated for 1 h with 1:10000 dilution of goat–anti–rabbit–horseradish peroxidase conjugate (Biorad). Finally, the membranes were washed 3 times over 30 min in blocking buffer and the interaction of GST-tagged proteins was detected by chemiluminescence.

### 1.5.2.10 NMR spectroscopy

#### 1.5.2.10.1 <sup>15</sup>N, <sup>13</sup>C and <sup>2</sup>H labelling of proteins

The following protocol has been used to <sup>15</sup>N, <sup>13</sup>C or <sup>2</sup>H label the proteins used for NMR spectroscopy in the standard expression system. The desired plasmid DNA was transformed in *E. coli* BL21–(DE3)–RIL cells and gene expression took place in M9 Minimal Medium similar to expression in LB medium (see Section 1.5.2.3.1). Expressed proteins were labelled by replacing ammonium chloride, glucose or H<sub>2</sub>O with <sup>15</sup>N ammonium chloride, <sup>13</sup>C glucose or D<sub>2</sub>O respectively as needed. Preparing M9 minimal media begins with preparing an autoclaved 10x stock solution of M9 salts and M9 Minimal Medium was prepared as follows:

#### M9 Salts:

Na <sub>2</sub> HPO <sub>4</sub>	Dibasic sodium phosphate
KH <sub>2</sub> PO <sub>4</sub>	Monobasic potassium phosphate
NaCl	Sodium chloride

**Preparation of 1 l M9 Minimal Medium (5.1.6):**

1) Use autoclaved or sterile filtered solutions

2) Mix together:

100 ml	10xM9 salts	
1 ml	1M MgSO <sub>4</sub>	
1 g	NH <sub>4</sub> Cl	Use <sup>15</sup> N ammonium chloride for <sup>15</sup> N labelling
3 g	glucose	Use <sup>13</sup> C glucose for <sup>13</sup> C labelling
1 ml	100 mM CaCl <sub>2</sub>	
10 ml	GibcoMEM Vitamin Solution	
ad 1l	H <sub>2</sub> O	Use D <sub>2</sub> O for <sup>2</sup> H labelling

2) Set pH between 7.4 and 7.0

3) Transfer into autoclaved Fernbach culture flask.

**1.5.2.10.2 NMR spectra**

NMR spectra were acquired at 25°C on a Varian NMR spectrometer operating at 800 MHz with a cryogenically cooled probe. Processing was performed in NMRPipe<sup>154</sup> and the resulting spectra visualized with XEASY<sup>155</sup>. All NMR samples were dissolved in buffer C. Optimized triple resonance experiments were performed to assign chemical shift values to hRpn13 Pru domain's backbone (N,HN,C $\alpha$ ,C') and C $\beta$  atoms. Chemical shift assignments were made by using 3D HNCA/HNCOCA and HNCACB experiments on 0.6 mM <sup>15</sup>N-, <sup>13</sup>C- and 70% <sup>2</sup>H-labelled hRpn13 Pru in buffer C. An <sup>15</sup>N dispersed NOESY spectrum (200 ms mixing time) acquired on <sup>15</sup>N-, 50% <sup>2</sup>H-labelled human Rpn13 Pru was used to confirm the assignments and that the secondary structural elements match those of the crystal structure. This spectrum was also used to aid in the identification of intermolecular NOE interactions between hRpn13 Pru and Ub, as it was compared to an <sup>15</sup>N dispersed NOESY experiment acquired under identical conditions on <sup>15</sup>N-, <sup>13</sup>C- and 70% <sup>2</sup>H-labelled hRpn13 Pru mixed

with equimolar ratio of unlabeled Ub. For this experiment, no  $^{13}\text{C}$  or  $^2\text{H}$  decoupling was implemented to enable the ready identification of intermolecular NOE interactions.

hRpn13 Pru	Ub
$\text{H}_\text{N}$	
F76	$\text{H}_{\gamma 2}$ of Q49
D78	$\text{H}_{\gamma 2}$ of V70
D79	$\text{H}_{\delta 2}$ of L8 $\text{H}_{\epsilon 1}$ of H68
F98	$\text{H}_\beta$ of A46
A100	$\text{H}_\beta$ of A46 $\text{H}_\alpha$ of A46
G101	$\text{H}_\beta$ of A46 $\text{H}_\delta$ of F45 $\text{H}_\epsilon$ of F45
K103	$\text{H}_\beta$ of A46
R104	$\text{H}_\beta$ of A46

**Table 1.11:** List of NOE-derived distance constraints between hRpn13 Pru and Ub. None of these observations were violated in any of the 200 calculated structures.

### 1.5.2.10.3 Chemical shift perturbation analysis

To determine how Rpn13 Pru binds Ub, chemical shift perturbation analysis and NOESY experiments were used to directly detect intermolecular contacts. Chemical shift perturbation data for Rpn13 Pru binding to ubiquitin and K48-linked tetraubiquitin were obtained for each amino acid residue by comparing the amide chemical shift values of Rpn13 Pru alone and with equimolar quantities of ubiquitin or K48-linked tetraubiquitin. Therefore, three sets of experiments were performed. In particular unlabeled monoUb or tetraUb was added to  $^{15}\text{N}$  labelled hRpn13 Pru or alternatively unlabeled hRpn13 Pru was added to  $^{15}\text{N}$  labelled monoUb. In each case, the  $^{15}\text{N}$  labelled protein was observed by  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC experiments

as it bound its unlabeled binding partner. The amide nitrogen and hydrogen chemical shift changes were mapped for each residue according to Equation 1:

$$\text{CSP} = \sqrt{0.2\Delta\delta_{\text{N}}^2 + \Delta\delta_{\text{H}}^2}$$

In this equation  $\Delta\delta_{\text{N}}$  and  $\Delta\delta_{\text{H}}$  represent the changes in the amide nitrogen and proton chemical shifts (in parts per million), respectively.

#### 1.5.2.10.4 Docking protocol

Rpn13 Pru/Ub complexes were generated by using HADDOCK1.3 (*High Ambiguity Driven protein–protein DOCKing*)<sup>121</sup> in combination with CNS<sup>122</sup>. The atomic coordinates for Ub were obtained from PDB entry 1D3Z<sup>123</sup>. All of the residues with chemical shift perturbation values greater than one standard deviation value above the average and with main chain or sidechain accessibility above 50% were defined as “active” residues (Table 1.12). Neighbours of these residues with greater than 50% accessibility were defined as “passive.” During docking, Ambiguous Interaction Restraints (AIRs) were defined that restrict each active residue to be within 2.0 Å of any atom of the binding partner’s active or passive residues. All active and passive residues localized to distinct surfaces in hRpn13 Pru and Ub that were also identified as the contact surface by intermolecular NOE interactions. “Flexible” segments (Table 1.12) were allowed to move their side chains upon complex formation and defined according to the “active” residues, as described<sup>121</sup>. Intermolecular NOE interactions between hRpn13 Pru and Ub were obtained by comparing the <sup>15</sup>N–dispersed NOESY experiments recorded on hRpn13 Pru alone with that of the hRpn13 Pru:Ub complex. In addition, chemical shift assignment of the sidechain Ub atoms was obtained by following the assigned resonances of the free protein upon addition of increasing molar ratios of unlabeled hRpn13 Pru. All atom pairs exhibiting an intermolecular NOE interaction were constrained to be within 1.8 – 6.0 Å of each other (Table 1.11). In addition, the constraints from the hydrogen bonds published for Ub (PDB code 1DZ3) were used to maintain its fold, as the protein does not undergo any gross structural change upon binding. For the first step of rigid–body energy minimization, 1000 structures were generated. 200 structures having the lowest energy from the rigid–body docking were subjected to semi–flexible simulated annealing in torsion angle space followed by refinement in explicit water. During semiflexible simulated annealing, atoms at the interface were allowed to move but constrained by the AIRs and unambiguous NOE–derived distance constraints. After water refinement, the resulting structures were

sorted according to intermolecular energy and clustered using 1.5 Å cut-off criteria. This treatment resulted in an ensemble of 200 structures with the average RMSD of  $0.94 \pm 0.31$  Å for all backbone atoms and  $0.75 \pm 0.23$  Å for the backbone atoms at the interface. The ten lowest energy structures in the lowest energy cluster were evaluated according to the following criteria: all intermolecular hydrophobic contacts between heavy atoms being less than 3.9 Å, intermolecular hydrogen bond distances between proton-acceptor and donor-acceptor pair being within 2.7 Å and 3.35 Å, respectively, and minimum angles of 90° for donor-hydrogen-acceptor, hydrogen-acceptor-acceptor antecedent, donor-acceptor-acceptor antecedent atoms of hydrogen bonds<sup>124</sup>.

Protein	Active residues	Passive residues	Flexible segments
hRpn13 Pru	55,76,78,99,101,104	53,54,74,77,81,83,100,102,103	51-57,72-85,97-106
Ub	47,48,49,68	6,8,9,10,11,12,16,42,46,51,72	4-18,40-53,66-74

**Table 1.12:** List of active and passive residues defined in hRpn13 Pru and Ub as defined by CSP plots for HADDOCK

### 1.5.2.11 Modelling of mRpn13 Pru:diUb complex

The model structure of mRpn13 Pru:diUb was built from the structures of the mRpn13 Pru:Ub complex and Ub (PDB entry 1D3Z) by using Insight II (Accelrys, Inc). For diUb, the two monoUb molecules were conjugated by forming the G76-K48 isopeptide bond and allowing rotations in the K48 sidechain of proximal Ub to avoid steric clashes between the distal Ub subunit and Rpn13. The structure was subsequently energy minimized by using SYBYL (Tripos, Inc.).

### 1.5.2.12 Biochemical pull-down assays

GST-fused tetraUb subcloned into pGEX-4T1 was coupled to Glutathione Sepharose TM 4B, cleaved with thrombin 4 h at 4°C (Novagen) in Thrombin cleavage buffer. Thrombin was inactivated with addition of PMSF and GST bound to beads was removed by centrifugation. Supernatant containing tetraUb was further used for GST pull-down with GST, GST-fused wildtype and different point mutants of Rpn13 Pru in incubation buffer. After 5 hours of incubation at 4°C, beads were washed three times with incubation buffer. Bound proteins

were boiled for 5 minutes in SDS sample buffer and analyzed by SDS-PAGE and immunoblotting.

### 1.5.2.13 Rpn2 binding assays

Different deletion mutants of hRpn2 (695–953, 797–953, 881–953, and 902–953) were amplified by PCR, subcloned into and pGEX-4TI, verified by sequencing and expressed as GST-fused proteins. The GST-tagged purified proteins were tested for their ability to bind mRpn13 (1–150), which was transiently expressed as myc-tagged protein in HEK293T cells. Two deletion mutants of Rpn2 (881–953 and 902–953) were transiently expressed as GFP-tagged proteins in HEK293T cells and tested for their ability to bind GST-mRpn13 (1–150). The pull-down assays were done as described in Section 1.5.2.12.

For mapping of the Rpn2 binding domain on Rpn13 Pru by NMR spectroscopy, Ub, hRpn2 (797–953) or Ub and hRpn2 (797–953) were added to  $^{15}\text{N}$  labelled hRpn13 Pru domain, which was monitored by  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC experiments. Similar to Section 1.5.2.10.3 the amide nitrogen and hydrogen chemical shift changes were mapped for each residue according to Equation 1.

## 2 Crystallographic studies of the TOM core complex

### 2.1 Abstract

The translocase of the outer mitochondrial membrane TOM mediates transfer of mitochondrial proteins across the outer membrane and plays a fundamental role in the biogenesis of the organelle. Its crystal structure is expected to provide many deep insights for understanding the molecular basis of protein translocation across the outer mitochondrial membrane and membrane trafficking in general.

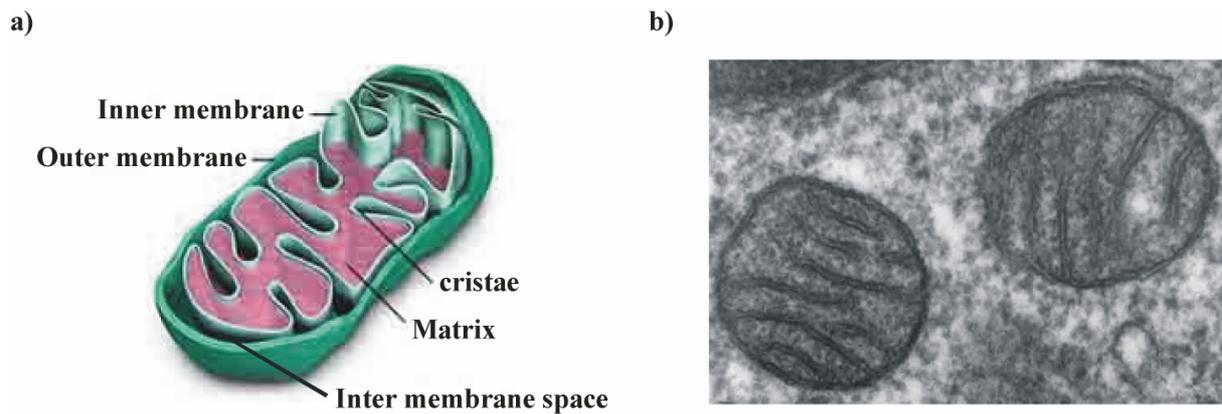
In this study, we present purification and crystallisation of the TOM core complex from *Neurospora crassa* and preliminary crystallographic data. We have been able to determine space group and cell dimensions of crystals and calculate selfrotation, revealing asymmetry of the TOM core complex. Well ordered crystals are essential for the determination of protein structures by X-ray crystallography, and low diffraction qualities of the TOM core complex crystals (8Å) are so far the limiting factor for phase and structure determination. Temperature factor and degree of mosaicity are the most important parameters of crystal order and their improvement will certainly contribute to a better diffraction pattern and increased resolution. Different attempts to improve the diffraction quality have been undertaken including detergent exchange upon purification, usage of additives such as detergents, heavy metal compounds and lipids as well as direct manipulation of the crystals. Since co-crystallization of membrane proteins with monoclonal antibody fragments was reported to promote their crystallization, TOM-directed monoclonal antibody fragments were derived and co-crystallized with TOM core complex. Based on the predicted  $\beta$ -barrel topology of the pore-forming component Tom40, its recombinant expression and crystallization may represent a convenient alternative to the crystallization of the whole complex. We achieved expression and refolding of Tom40 from *N. crassa*, which further purification should result in crystallization-grade protein.

## 2.2 Introduction

### 2.2.1 Mitochondria

The mitochondrion is a membrane-enclosed dynamic organelle found in most eukaryotic cells<sup>156</sup>. These organelles range from 1 – 10  $\mu\text{m}$  in size and are often described as “cellular power plants” since they generate most of the cell’s energy supply by oxidative phosphorylation in terms of adenosine triphosphate (ATP). In addition to providing energy, mitochondria are involved in several fundamental processes such as signalling, cellular differentiation, cell death as well as cell growth and cell cycle control<sup>157</sup>. Important reactions occurring in mitochondria include the citric acid cycle (Krebs cycle), heme biosynthesis,  $\beta$ -oxidation of fatty acids, metabolism of certain amino acids and biogenesis of Fe/S clusters<sup>158</sup>. Their ancestry is not fully understood, but according to the endosymbiotic theory, mitochondria represent a relic of an ancient species of alphaproteobacteria, which were engulfed by the first eukaryotes more than a billion years ago<sup>159</sup>. In consistence with this theory, mitochondria possess, like chloroplasts and the nucleus, their own independent genome, which shows substantial similarity to bacterial ones. During evolution, the bacterial symbiont transferred most of its genome in a gradual process into the nuclear chromosomes<sup>160</sup> and consequently the majority of proteins functioning in the organelle are now coded on nuclear genes (about 99%). For example, the human mitochondrial genome is a circular DNA of about 16 kilobases encoding 37 genes: 13 for subunits of respiratory complexes I, III, IV, and V, 22 for mitochondrial tRNA, and 2 for rRNA<sup>161,162</sup>. The number of mitochondria in a cell varies widely by organism and tissue type. Many cells have only a single mitochondrion, whereas others can include several thousand of the organelle<sup>163,164</sup>. Growth and division of mitochondria are not linked to the cell cycle but are mainly dependent of the energy requirements of the cell. If the energy use is low, mitochondria are destroyed or inactivated. At cell division, they are distributed to the daughter cells randomly during division of the cytoplasm. Mitochondria divide by binary fission similar to bacterial cell division, but unlike to bacteria, mitochondria can also fuse<sup>165</sup>.

Although some anaerobic unicellular eukaryotes contain mitochondria that are atypical with regard to organellar biochemistry and morphology, they all contain organelles with double membrane, whether they are called mitochondria, hydrogenosomes or mitosomes<sup>160</sup>. This double-membraned organization of mitochondria accounts for their partition into four distinct compartments (Figure 2.1): the outer mitochondrial membrane, the intermembrane space (IMS), the inner mitochondrial membrane and the matrix.



**Figure 2.1:** (a) Schematic representation of a mitochondrion showing its compartmentalization. (b) Electron micrograph of mitochondria from mammalian lung.

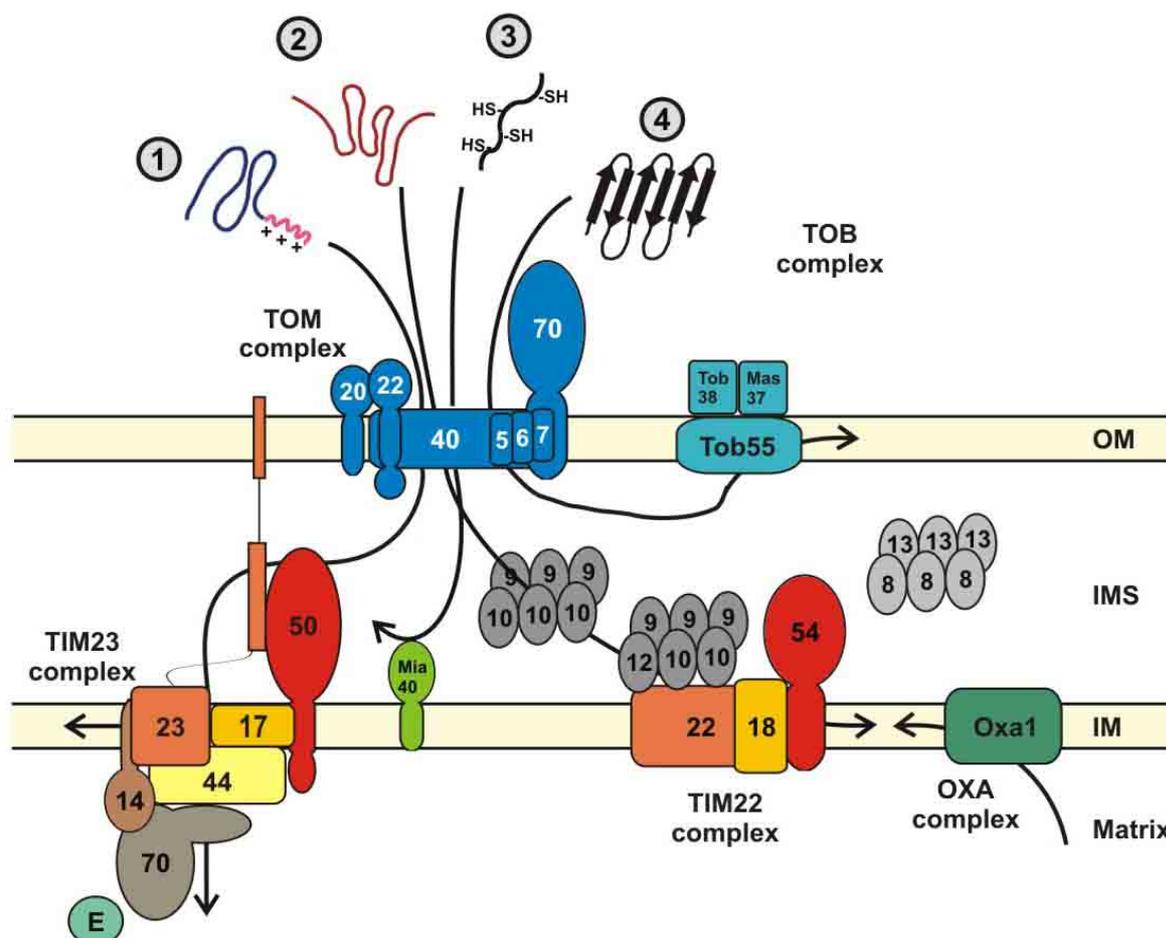
The cristae space (Figure 2.1) is formed by infoldings of the inner membrane, which expand the surface area of the inner membrane, thereby enhancing its ability to produce ATP. These are not simple random folds but rather invaginations that can affect overall chemiosmotic function<sup>166,167</sup>. The outer mitochondrial membrane encloses the entire organelle and shows a protein-to-phospholipid ratio similar to that of the eukaryotic plasma membrane. It comprises large numbers of integral proteins called porins, which form channels that allow free diffusion of molecules with a molecular mass less than 5000 Da. In contrast, large proteins need specific signalling sequences to be transported through the outer membrane. The inter-membrane space (IMS) is between the outer and the inner membrane. Due to the permeability of the outer membrane to small molecules, their concentrations are the same as in the cytosol, whereas the protein composition is different. The inner membrane is rich in the unusual phospholipid cardiolipin<sup>168</sup> and comprises proteins that fulfil the following functions: redox reactions of the oxidative phosphorylation, generation of ATP (ATP synthase), regulation of metabolite passage through the membrane and protein import<sup>163</sup>. Unlike the outer membrane, the inner membrane lacks porins and is highly impermeable to all molecules. This impermeability enables the enzymes of the electron transport chain to maintain a membrane potential across the inner membrane. The matrix comprehends about 2/3 of the total mitochondrial protein content and is important in the production of ATP. It contains distinct enzymes, which major functions include oxidation of pyruvate and fatty acids as well as the citric acid cycle, special mitochondrial ribosomes, tRNA, and several copies of the mitochondrial genome<sup>163</sup>.

With their central place in cell metabolism, dysfunction in mitochondria is an important factor in a wide range of human diseases. Mutations in the mitochondrial DNA (mtDNA) for example are associated with diseases such as cardiomyopathy, MELAS syndrome, Leber's hereditary optic neuropathy and many others<sup>169</sup>. Nuclear genes encoding mitochondrial

proteins can be also mutated and be a cause of human diseases. In addition to the well-documented diseases, several other common human diseases are accompanied by mitochondrial defects. For example, a major contributing factor to Parkinson's and Alzheimer's disease is reduced cellular ATP levels caused by mitochondrial dysfunction.

## 2.2.2 Protein import into mitochondria

The contribution of the mitochondrial genome to the mitochondrial proteome is very limited and does not exceed 1% of its total number of proteins. Since the vast majority of mitochondrial proteins is nuclear-encoded and synthesized in the cytosol, the biogenesis of mitochondria, therefore, depends on the import of these proteins from the cytosol into the organelle (Figure 2.2).



**Figure 2.2:** Schematic representation of different import pathways in mitochondria. 1=presequence import pathway, 2=carrier pathway, 3=intermembrane space import and assembly pathway, 4=outer membrane sorting and assembly pathway.

The processes of recognition and sorting of mitochondrial preproteins are mediated by protein translocase machineries present in both the outer and inner mitochondrial membranes. Most of the mitochondrial precursor proteins, if not all, are in an unfolded conformation and associated with chaperones, which maintain them in a translocation-competent conformation. They are delivered to the organelle by virtue of specific targeting signals, which are recognized by specialized receptor proteins<sup>170,171</sup>. All cytosolic precursor proteins are imported by the general entry gate, the translocase of the outer mitochondrial membrane (TOM, see Section 2.2.2.3). At least four classes of precursor proteins exist that contain distinct targeting signals and are directed to different import routes (Figure 2.2). This classification leads to the description of the following import pathways: the presequence pathway to matrix and inner membrane, the carrier pathway to inner membrane, intermembrane space import and assembly as well as outer membrane sorting and assembly (Figure 2.2)<sup>172</sup>. Additionally, some inner membrane proteins follow a conservative import pathway. They are fully imported into the matrix via TOM and TIM23 complexes (see 2.2.2.1) and are then exported from the matrix into the inner membrane in a process that involves the OXA (oxidase assembly) complex<sup>170</sup>.

### 2.2.2.1 Presequence import pathway

The classic import pathway of mitochondrial matrix proteins involves a characteristic N-terminal presequence that forms a positively charged amphipathic  $\alpha$ -helix (MTS, matrix targeting signal) (Figure 2.2). These proteins were directed across both membranes into the matrix where the presequences are usually proteolytically removed by the matrix processing peptidase (MPP)<sup>172</sup>. After crossing the outer membrane through the TOM complex, the preproteins are transferred to the translocase of the inner mitochondrial membrane TIM23 (also termed presequence translocase). The TIM23 complex can be structurally and functionally subdivided into the membrane-integrated translocation channel, including the components Tim17, Tim21, Tim23 and Tim50, and the matrix-faced import motor PAM (for presequence translocase-associated motor) which comprises at least five proteins, Tim14 (Pam18), Tim16 (Pam 16), Tim44, Mge1 and mtHsp70 (mitochondrial heat-shock protein 70)<sup>170</sup>. Transport of the presequences through the TIM23 complex is driven by the inner membrane potential  $\Delta\psi$  and the hydrolysis of ATP.  $\Delta\psi$  performs a dual role: It activates the channel protein Tim23 and drives translocation of positively charged presequences by an electrophoretic mechanism<sup>173</sup>. Tim23 and Tim17 form the membrane-embedded core of the translocase and share a related membrane domain consisting of four

transmembrane segments<sup>171,174</sup>. Tim50 is anchored into the inner membrane and, as Tim21, was found to participate in preprotein transfer from TOM to Tim23. Tim50 binds to the incoming polypeptide chains as they reach the *trans* site of the TOM complex and presumably passes them to TIM23<sup>174-177</sup>. Tim21 connects TOM and TIM complexes by direct but transient interaction with Tom22. The intermembrane space domain of Tim50 binds to Tim23 and promotes closing of the channel in the absence of preproteins, thus preventing leakage of ions<sup>178</sup>.

The import motor is driven by the ATP requirement of mtHsp70, which interacts with the unfolded preprotein in transit and, supported by membrane-associated co-chaperones, promotes their translocation to the matrix. Two current hypotheses for the mechanism of action of mtHsp70 exist: trapping of preproteins by a Brownian ratchet that prevents back-sliding of the polypeptide chain, or active pulling by a motor (power stroke mechanism)<sup>179-184</sup>. Both models agree that mtHsp70 binds to the inner membrane through Tim44 in an ATP-regulated manner. The identification of the Tim14:Tim16 complex revealed tight regulation of the ATPase activity of mtHsp70 and thereby of translocation<sup>185,186</sup>.

Several integral inner membrane proteins are synthesized with an N-terminal presequence followed by a hydrophobic sorting signal. The sorting signal arrests translocation and induces lateral release of the protein into the inner membrane independently of the import motor<sup>170,172</sup>. This process involves Tim17 and Tim21 and is referred to as the stop-transfer pathway<sup>171,174,187</sup>.

### 2.2.2.2 Carrier pathway

Metabolite carriers such as ADP/ATP and the phosphate carrier form a large class of inner membrane proteins. The carrier or TIM22 pathway (Figure 2.2) is a sorting route selectively used by members of the solute carrier family and by the membrane-embedded TIM subunits Tim17, Tim22 and Tim23. All known substrates of the TIM22 complex represent membrane proteins with even-numbered transmembrane segments that expose their N and C termini to the IMS. Carriers contain usually six transmembrane segments without any cleavable presequence but with an internal targeting signal. However, presequence-carrying preproteins and carrier precursors are imported through the TOM channel, but the former are translocated as unfolded linear polypeptide chains, whereas the latter cross the outer membrane in a loop formation<sup>188</sup>. In the next step of this pathway, the small TIM proteins are employed. These chaperone complexes of the intermembrane space, namely the hexameric Tim9:Tim10 and Tim8:Tim13 complexes, shield the hydrophobic domains of carrier proteins and accompany

them from the TOM to the TIM22 complex<sup>189-194</sup>. Finally, carrier proteins are taken over by TIM22 complex and inserted into the lipid bilayer of the inner membrane in a membrane potential-dependent manner<sup>171</sup>. A modified chaperone complex, which contains Tim12 in addition to Tim9 and Tim10, is located at the surface of the translocase. The core is built by the channel-forming Tim22, responsible for insertion, whereas the exact functions of the associated integral membrane proteins Tim54 and Tim18 are not known<sup>172,195</sup>.

### 2.2.2.3 Intermembrane space import and assembly pathway

Many proteins of the IMS are small (less than 20 kDa) and their folding is stabilized by cofactors or disulfide bridges. Systematic analysis led to the identification of a specific protein import and assembly machinery (Figure 2.2), which central component, the inner membrane protein Mia40, shows characteristic cysteine motifs. Following their translocation across the outer membrane, the precursor forms are recognized and specifically bound by Mia40 through mixed disulfide bonds<sup>196-199</sup>. Mia40 promotes the sequential formation of intramolecular disulfides in the imported proteins and thus initiates the assembly of imported proteins into oligomeric complexes<sup>199,200</sup>. Typical examples of this are the hexameric chaperone complexes formed by small Tim proteins. The activity of Mia40 relies on the function of the sulphhydryl oxidase Erv1, an essential FAD-binding protein that directly interacts with Mia40 and maintains it in an oxidized, active state. The current model suggests a disulfide relay system with electrons being transferred from the precursor protein to Erv1 through Mia40, and from there via cytochrome *c* to the respiratory chain<sup>201-203</sup>. It is now a matter of debate, if Erv1 has further roles and if additional factors might be involved.

An alternative pathway to regulate the import of proteins into the IMS is the stabilization by cofactors. For example, cytochrome *c* is trapped in the IMS by the enzyme cytochrome *c* heme lyase, which inserts the heme factor into cytochrome *c*. This soluble holoform is then released into the IMS<sup>204-206</sup>.

### 2.2.2.4 Outer membrane sorting and assembly pathway

The outer membranes of mitochondria, chloroplasts and Gram-negative bacteria contain  $\beta$ -barrel proteins, a special type of integral membrane proteins. A beta barrel is a large beta-sheet that twists and coils to form a closed structure in which the first strand is hydrogen bonded to the last<sup>207</sup>. So far identified  $\beta$ -barrel proteins are porin (or VDAC), Tom40, Tob55

(Sam50), Mdm10 and Mmm2. Like all other mitochondrial outer membrane proteins, they are synthesized in the cytosol. The  $\beta$ -barrel precursors are imported by the TOM complex and inserted into the outer membrane by a translocase named TOB (topogenesis of mitochondrial outer membrane  $\beta$ -barrel) or SAM (sorting and assembly machinery) (Figure 2.2)<sup>208,209</sup>. Precursors of  $\beta$ -barrel proteins interact mainly with the receptor Tom20 of the TOM complex and are subsequently transported through the translocase to the *trans* side of the outer membrane, where TIM chaperone complexes guide the hydrophobic precursors to the TOB complex<sup>210,211</sup>. Its main component is the essential Tob55 (Sam50), which is homologous to the bacterial outer membrane protein Omp85<sup>208,212</sup>. Several partner proteins of Tob55, namely Tob38 (Sam35), Mas37 (Sam37) and Mdm10 (Mitochondrial distribution and morphology 10), aid in membrane insertion and assembly of the  $\beta$ -barrel precursors. The exact function of Tob38 and Mas37 is not yet known. Mdm10 together with the two other morphology proteins, Mdm12 and Mmm1 (Maintenance of mitochondrial morphology 1) forms a complex that functions in the assembly pathway of  $\beta$ -barrel proteins downstream of the TOB complex<sup>213</sup>. The N-terminal hydrophilic region is exposed to the IMS and forms a characteristic structure, called POTRA (polypeptide translocation associated) domain<sup>214</sup>. The POTRA domain of Tob55 may be involved in the recognition of  $\beta$ -barrel precursors and might pass them on to the membrane-embedded C-terminal region of Tob55, which then facilitates membrane insertion and assembly of the  $\beta$ -barrel proteins<sup>171,215</sup>.

### 2.2.3 The TOM complex

The molecular machine translocating proteins across the mitochondrial outer membrane is the TOM complex (translocase of the outer mitochondrial membrane, Figure 2.2). This large multisubunit membrane protein is responsible for the initial recognition of precursor proteins in the cytosol and subsequent transfer of the polypeptides through pores. It also facilitates release of cytosolic-binding factors and contributes to the unfolding of cytosolic protein domains. The holo-TOM complex is composed of at least seven different subunits. Tom20 and Tom70 are the major receptors that recognize preproteins, whereas the subunits Tom40, 22, 7, 6 and 5 form the stable TOM core complex. Tom20 and Tom70 are both anchored with N-terminal transmembrane segments in the outer membrane and expose hydrophilic domains into the cytosol. They differ in their substrate specificity, but overlap in their function and can partially substitute each other. As shown in the crystal structure of Tom70<sup>216</sup> it contains several conserved tetratricopeptide repeat (TPR) motifs, which are organized in a right-

handed superhelix. These motifs form a peptide-binding groove for specific interaction with cytosolic chaperones like Hsp70, whereas the C-terminal part of Tom70 reveals a putative binding site for precursor proteins. Tom70 shows substrate preference for hydrophobic precursors that comprise internal targeting signals<sup>216,217</sup>. In contrast, Tom20 is the central receptor for N-terminal presequences and structural analysis by NMR demonstrated a binding groove for the hydrophobic face of the MTS (matrix targeting signal, see Section 2.2.2.1)<sup>218</sup>. The binding sites for precursor proteins on the cytosolic site are often referred to as *cis*-binding sites, whereas the IMS-exposed binding surface of the TOM complex represents the *trans*-binding site<sup>171</sup>.

Tom40 represents the pore-forming component of the complex and was proposed, based on theoretical predictions, to adopt a  $\beta$ -barrel fold (see Section 2.2.2.4). Even in the absence of other TOM subunits, purified Tom40 forms pores in artificial membranes that show characteristics to that of the entire TOM complex<sup>219-221</sup>. Tom22, Tom5, Tom6, and Tom7 each have a single  $\alpha$ -helical transmembrane segment that locks them tightly into position on Tom40<sup>222</sup>. Tom22 spans the outer membrane in a N<sub>out</sub>-C<sub>in</sub> orientation exposing a highly negatively charged N-terminal domain to the cytosol and a short C-terminal part to the IMS. It assists the transfer of substrate proteins from the receptors to the pore and may cooperate with Tom20 in the binding and unfolding of precursor proteins. In addition, Tom22 plays a critical role for the general integrity of the TOM complex<sup>171,223-226</sup>. The small TOM proteins (Tom5, Tom6 and Tom7) appear to function in regulating the stability of interactions within the complex, thereby assisting substrate protein transfer to and through the core translocase. The loss of individual small TOM subunits causes only minor effects, but simultaneous deletion of all three is lethal<sup>227-229</sup>. Mutational analysis in *Neurospora crassa* revealed that Tom6 plays a major role in TOM complex stability, whereas Tom7 has a lesser role<sup>229</sup>. The TOM holo complex (purified in the detergent digitonin) has a molecular mass of roughly 490-600 kDA and single-particle imaging of the negatively stained TOM holo complex showed particles with two or three pore-like structures, whereas the TOM core complex, which lacks Tom20 and Tom70, contains two pores with a diameter of  $\sim 2.1$  nm<sup>230-232</sup>. Isolated Tom40 alone is capable of forming homooligomeric structures, which reveal one cavity<sup>220,221,233</sup>. Further studies by electron microscopy report that Tom20 is selectively responsible for the presence of the three pore-like structure. Both subunits Tom22 and Tom20 appear to be critical for the assembly of Tom40 channel units<sup>232</sup>.

In addition to translocation of precursor proteins, the TOM complex is also involved in the integration of outer membrane proteins into the membrane, a process which is only partially understood<sup>171,234</sup>. The structural organization of the TOM machinery is still a subject of

debate and three-dimensional structures will shed light on its overall topology and import mechanism.

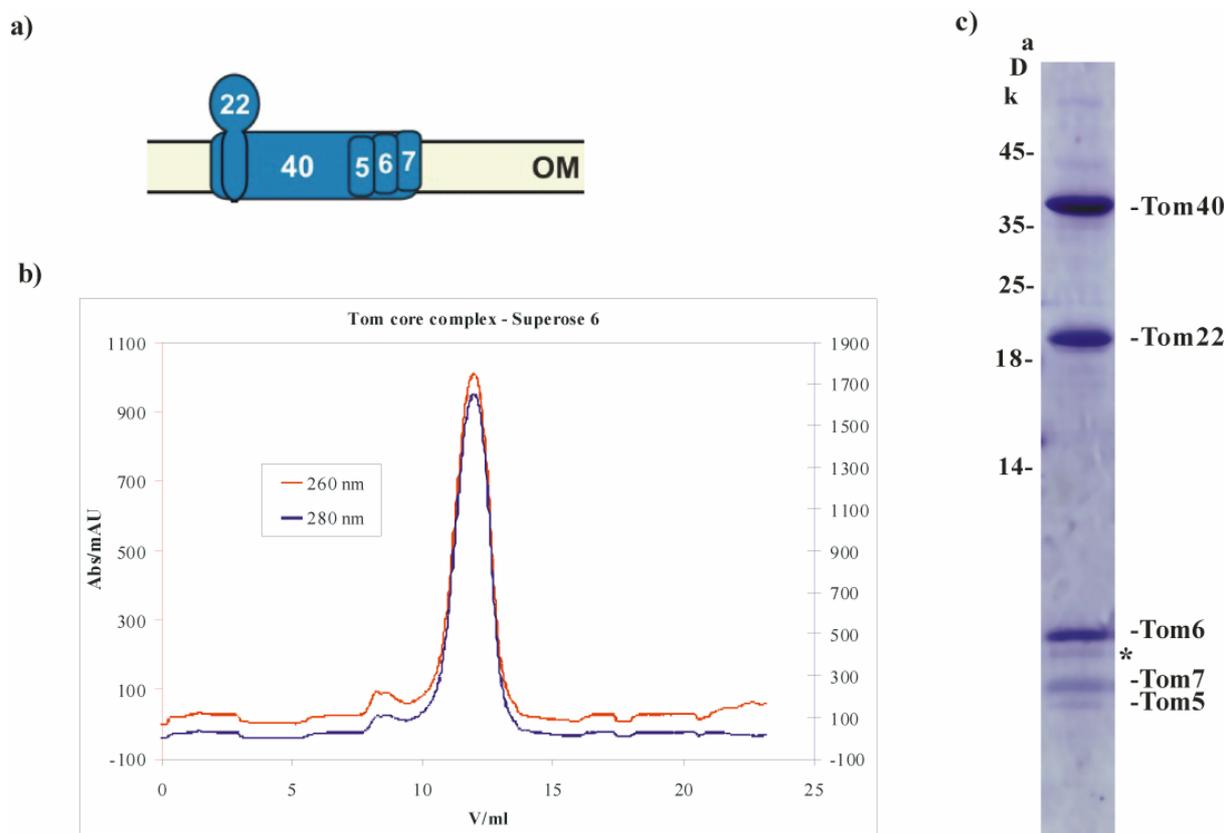
## 2.2.4 Goals of this study

In the outer membrane of mitochondria the TOM complex represents the main entry site for distinct types of mitochondrial precursor proteins. The complex interaction modes of receptor subunits with preproteins and subsequent translocation of unfolded substrate proteins through the membranes are actively investigated, but still very poorly understood. The filamentous fungi *Neurospora crassa* turned out to be an excellent model organism for studying the TOM complex due to its fast growth rate and simple manipulation procedures. The aim of this work was the improvement of crystals of the TOM core complex to get initial structural information of the translocase complex. Therefore, the purification strategy of TOM core complex had to be optimized and the crystal quality should be improved by additive screening and crystal manipulation approaches as well as by detergent exchange. Since crystallization in complex with antibody fragments could facilitate the crystallization and structure determination of membrane proteins, one major goal of this work was to derive TOM-directed monoclonal antibody fragments from classical hybridoma technology to accomplish amelioration of the diffraction quality by co-crystallization. Another aim of this study was the recombinant expression of the main TOM component Tom40 in *E. coli* and its subsequent refolding to obtain highly pure protein suitable for crystallization.

## 2.3 Results

### 2.3.1 Isolation and purification of TOM core complex

The TOM complex which consists of Tom70, 40, 22, 20, 7, 6, and 5 has been shown to be essential for biogenesis of mitochondria. This TOM holo complex can only be isolated in its native state from *Neurospora crassa* using the detergent digitonin, but since digitonin is not stable in water and precipitates in a crystalline form after a few days, this detergent is not suitable for protein crystallization. In contrast, the TOM core complex, which lacks the two receptor subunits of the holo complex, Tom70 and Tom20 (Figure 2.3a), was purified in the detergent *n*-dodecyl- $\beta$ -D-maltoside ( $\beta$ -DDM)<sup>231</sup>. The filamentous fungus *Neurospora crassa* was used as model organism for investigating the TOM complex.

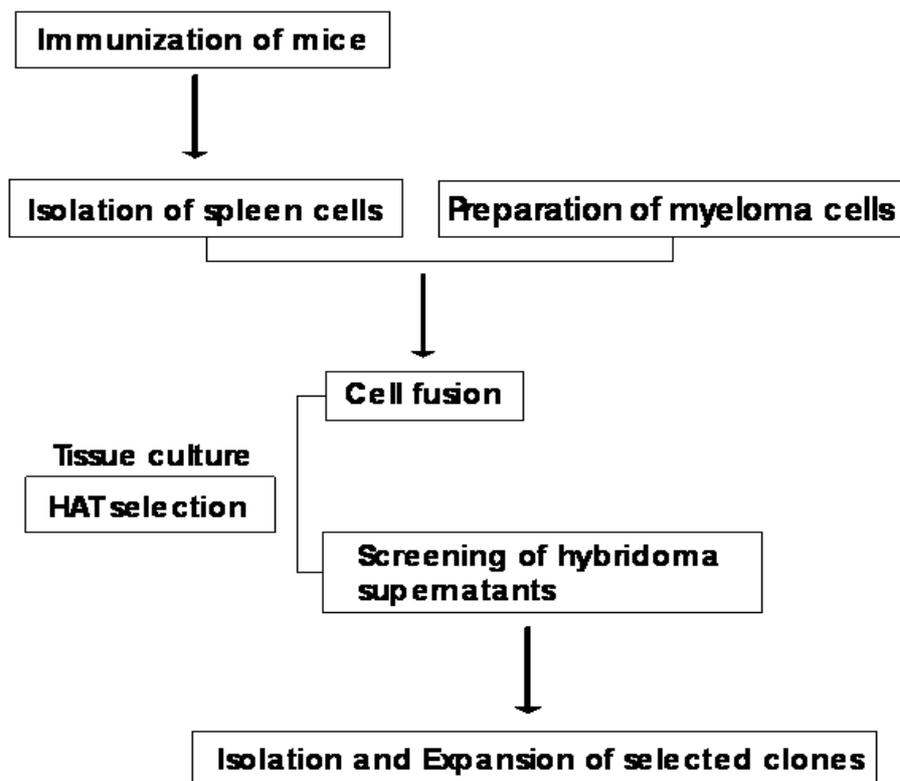


**Figure 2.3:** Purification of TOM core complex from *Neurospora crassa*. (a) Schematic representation of TOM core complex and its subunits Tom40, 22, 7, 6, 5. (b) Chromatogram of a Superose 6 run. (c) Coomassie-stained urea-SDS-PA gel of highly purified protein showing the five subunits of the complex. The asterisk marks a degradation product of Tom6.

In this study, the TOM core complex was isolated from a *Neurospora crassa* strain (GR-107) that carried a version of Tom22 with a hexahistidinyI tag at its C-terminus (Figure 2.3). Growth of the cells and preparation of mitochondria were performed as described previously<sup>230,235</sup>. The TOM core complex was isolated according to Athing et al., 1999, with modifications in the protocol. Isolated mitochondria were solubilized using  $\beta$ -DDM, which removes the transiently bound receptor subunits Tom70 and Tom20 from the TOM holo complex. The core complex was purified from the clarified extract by affinity chromatography via Ni-NTA (for more details, see Materials and Methods). Further purification of the complex by anion exchange (Resource Q) and size exclusion chromatography (Superose 6) resulted in more than 95% pure and stable protein suitable for crystallization trials and functional studies (Figure 2.3). The gel filtration run on Superose 6 and the Coomassie-stained urea-SDS polyacrylamide gel of purified TOM core complex are depicted in Figure 2.3. An average preparation of the protein started with ~1.5 kg of *Neurospora* cells (wet weight), which yielded in 3-5 g of mitochondrial protein and from about 10 mg of mitochondria 2-3 mg of purified TOM core complex were obtained. Isolated TOM core complex demonstrated channel activity<sup>231</sup>, was stable at higher temperatures (45 °C), in urea (4M) and relatively resistant to protease treatment. Upon gel filtration, the TOM core complex ran at a high molecular mass of ~400 kDa. The purified protein was subjected to crystallization and used to generate monoclonal antibodies, which are directed against TOM core complex (see below). The introduction of size exclusion chromatography as an additional purification step could improve the quality of the crystals (see below).

### **2.3.2 Generation of murine monoclonal antibodies recognizing native epitopes of Tom core complex**

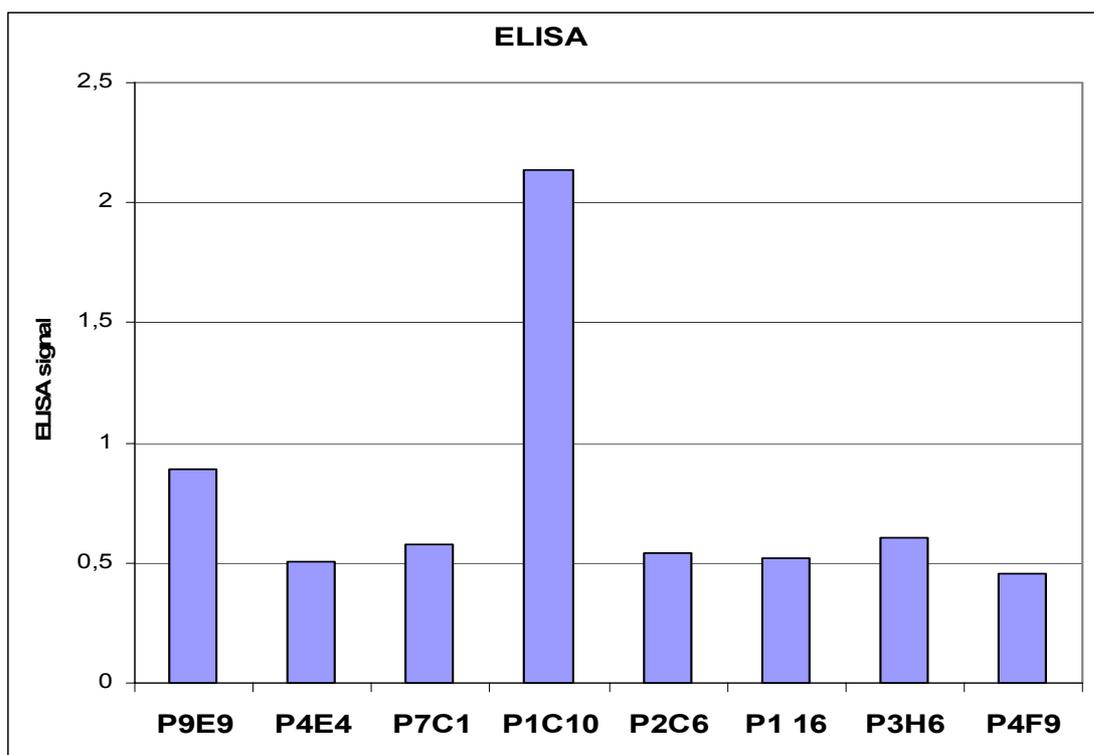
Crystallisation in complex with monoclonal Fv or Fab fragments has been reported to facilitate the crystallization of membrane proteins and to improve the diffraction quality of such crystals<sup>236</sup>. Therefore, monoclonal antibodies directed to TOM core complex were derived from murine cell culture using the highly purified, detergent solubilized TOM core complex. The scheme in Figure 2.4 shows the different steps to generate murine monoclonal antibodies (mABs). It started with the initial immunization of mice with purified TOM core complex. After preparation of spleen cells and their fusion with myeloma cells to obtain permanently living cells, the fusion was distributed in different dilutions on 96well plates and cultivated (see Materials and Methods). Only fused myeloma cells will survive in tissue culture upon HAT selection (for details, see Section 2.5.2.3.2).



**Figure 2.4:** *Generation of TOM-directed monoclonal antibodies (mABs). Schemata showing the individual steps of the manufacturing of the mABs recognizing discontinuous epitops of TOM core complex.*

Screening of the supernatants of the resulting hybridoma clones for conformation-specific mABs was performed using an ELISA (enzyme-linked immunosorbent assay). Clones to be selected clones were isolated via dilution and further tested by ELISA (single cell cloning) to make sure to gain single hybridoma clones (Section 2.5.2.3.3). Positive clones were selected by sandwich ELISA as described in Section 2.5.2.3.5.

Around 80 hybridoma clones were found that produced antibodies which exhibited a positive signal in the ELISA test. However, subtyping revealed that only eight of them produced antibodies of the requested IgG class (see Section 2.3.3.1). The diagram shows the average signal strength of the IgG antibody containing supernatants minus the average background value. Antibody PIC10 exhibited a much higher signal in the ELISA test than the other antibodies. Ni-NTA coated plates to which TOM core complex was bound via its His-tag, were also tested to improve the signal to noise ratio, but the ELISA signals were comparable to the signals of the sandwich approach.



**Figure 2.5:** Hybridoma clones producing murine IgG antibodies with elevated signal strength. The diagram shows the average ELISA signals of each antibody containing supernatant minus the average background value. Antibody P1C10 exhibits the strongest ELISA signal.

### 2.3.3 Characterization of selected monoclonal antibodies

#### 2.3.3.1 Immunological subtyping of murine antibodies

Immunological subtyping experiments were necessary to define the isotype of the immunoglobulines found. They showed that only 10% of all antibodies detected belonged to the IgG class and 90% to the IgM class of immunoglobulines. IgMs form pentamers and therefore reveal high signal strength because of their avidity, but are not suitable for copurification and cocrystallization with TOM core complex due to their lower affinity. To determine the Ig subtype, the mouse hybridoma subtyping kit from Boehringer-Mannheim was used for ELISA experiments. The subtyping was possible using peroxidase conjugated secondary antibodies which were specific for the different subtypes of Igs and for the two subclasses of the light chain (IgA, IgG, IgG1, IgG2a, IgG2b, IgG3, IgM,  $\lambda$ - and  $\kappa$ -chain). The signal strength was detected by an ELISA reader. Table 2.1 shows the results of the subtyping for the eight positive IgG clones.

Name	Subtype
P9E9	IgG1
P4E4	IgG1
P7C1	IgG2a
P1C10	IgG2a
P2C6	IgG2b
P1 16	IgG2b
P3H6	IgG3
P4F9	IgG3

**Table 2.1:** *Subtyping of antibodies from hybridoma supernatants showing positive signals in ELISA tests. All four IgG classes are present*

The IgA isotype was absent, whereas all different IgG subclasses were present. The light chains of all IgG antibodies belonged to the  $\kappa$  subclass.

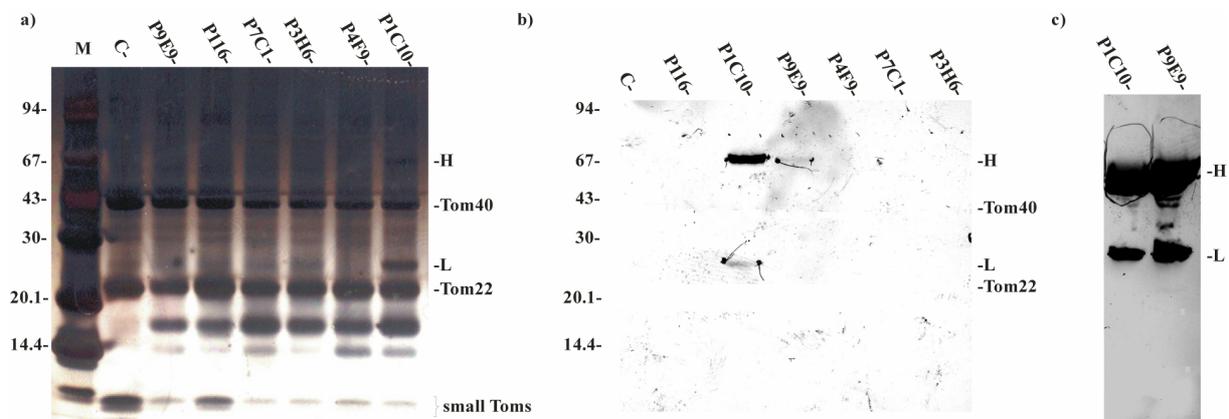
### 2.3.3.2 Western Blot analysis for selection of conformation-specific antibodies

Monoclonal antibody fragments that aid in crystallization should bind a discontinuous epitope on their target protein. Due to their high affinity of antibodies to the natively folded protein they generally exhibit no or only weak signals in Western Blot analysis, performed with unfolded proteins. The selected monoclonal IgG antibodies were tested by Western Blot analysis using the cell culture supernatants and an anti-mouse secondary antibody. TOM core complex was run on a gradient SDS gel, blotted onto nitrocellulose membrane and decorated with the supernatants of the positive hybridoma clones (for details see Materials and Methods). Except for the antibodies P3H6 and P1C10, which exhibited a very weak signal for Tom40, none of the tested antibodies showed activity in Western Blot. In general, all tested antibodies which revealed Western Blot activity appeared to be directed against the pore subunit Tom40.

### 2.3.3.3 Pull-down assay with hybridoma supernatants

Pull-down assays were useful to confirm the *in vitro* interaction of the antibodies from the hybridoma cultures with TOM core complex. Since TOM core complex carries a His-tag on Tom22, purified complex was bound to Ni-NTA magnetic agarose beads and utilized to capture TOM-directed monoclonal antibodies from cell culture supernatants. The eluted

protein mixture should contain the TOM core complex and the binding antibody, respectively. This was monitored by silver-stained SDS-PAGE as well as by Western Blot analysis (Figure 2.6).

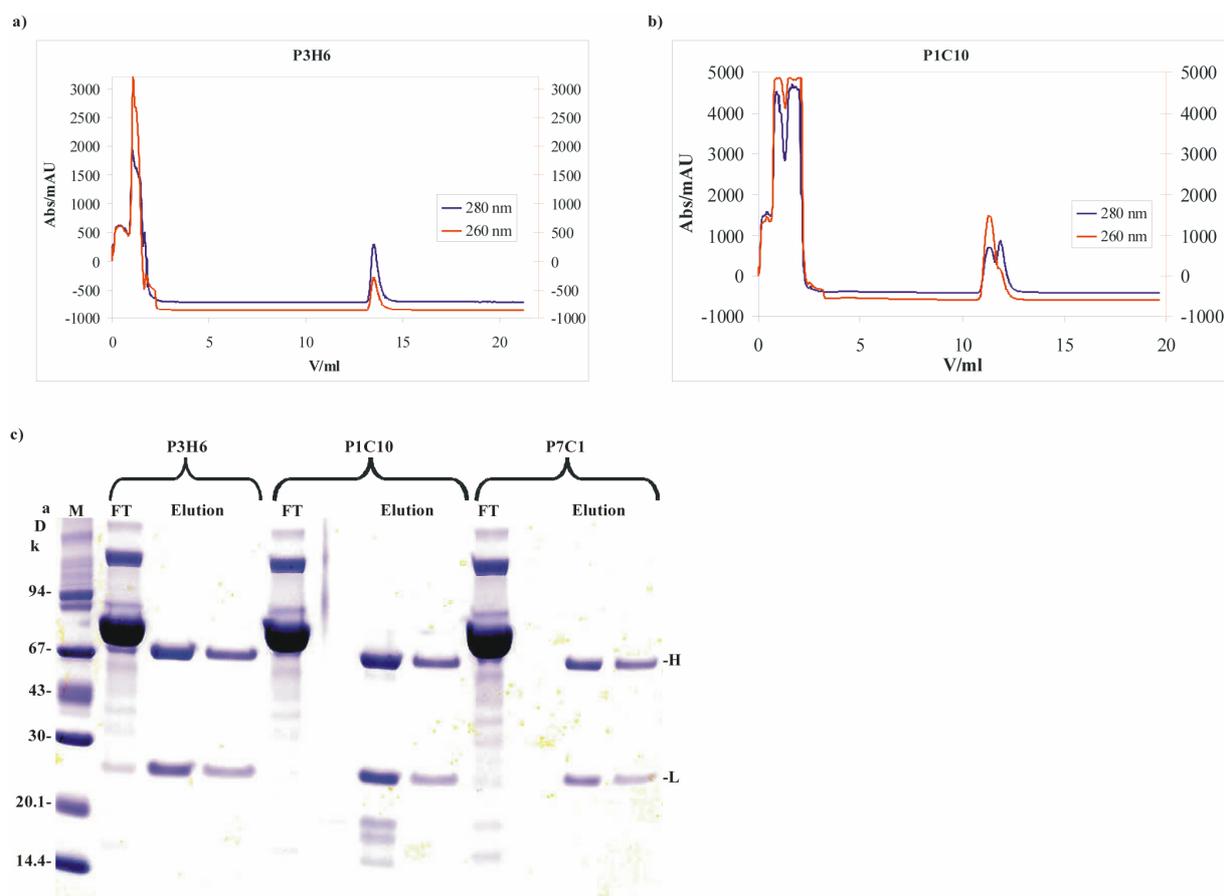


**Figure 2.6:** *Ni-NTA pull-down of monoclonal antibodies with TOM core complex. (a) Silver-stained gradient (4-12%) SDS-PAGE of elution fractions. Six hybridoma supernatants were deployed in this example. P1C10 is the only antibody, which seemed to bind to TOM core complex under the given conditions. The gel shows the heavy (H) and the light (L) chains of P1C10 with the size of about 62 and 26 kDa. M=marker, C=control. (b) Similar experiment as in (a), but monitored by Western Blot analysis. P1C10 was clearly binding. (c) To (b) corresponding flow through fractions of the antibodies P1C10 and P9E9, analyzed by Western Blot and showing the use of a large excess of the antibodies.*

Under the conditions applied P1C10 was the only antibody that was clearly captured from the hybridoma supernatant by TOM core complex as depicted in Figure 2.6a and b. Both the heavy and the light chains (around 26 and 62 kDa) of P1C10 were present in the silver-stained SDS gel and in Western Blot. In comparison to P1C10, pull-down experiment with P9E9 exhibited a very weak Western Blot signal (Figure 2.6b), indicating a possible interaction with TOM core complex, but one with low affinity. In all cases, a large excess of the antibodies was used for incubation with bound TOM core complex (Figure 2.6c) and analysis of the flow through fractions demonstrated that all antibodies were present in the hybridoma supernatant.

### 2.3.3.4 Purification of murine antibodies from hybridoma cell culture

For production and subsequent purification of murine antibodies from hybridoma cell culture, the culture volume and the production rate was scaled up and each mAB (murine antibody) was purified from the supernatant using protein A affinity chromatography. Since the high concentrations of fetal calf serum (FCS) in the medium cause high grade of impurity by its elevated content of antibodies, the first task before purification was to adapt the hybridoma cells stepwise to FCS-free culture medium. Cells were grown on plates for several days and then transferred to cell culture flasks (for details see Materials and Methods), resulting in about 180 ml of culture medium per hybridoma clone with a typical protein concentration between 2 and 20  $\mu\text{g/ml}$ . The supernatant was loaded onto protein A sepharose column to bind the mABs, which were then eluted at low pH (3-4). Figure 2.7a shows the exemplary chromatogram of the antibody P3H6.



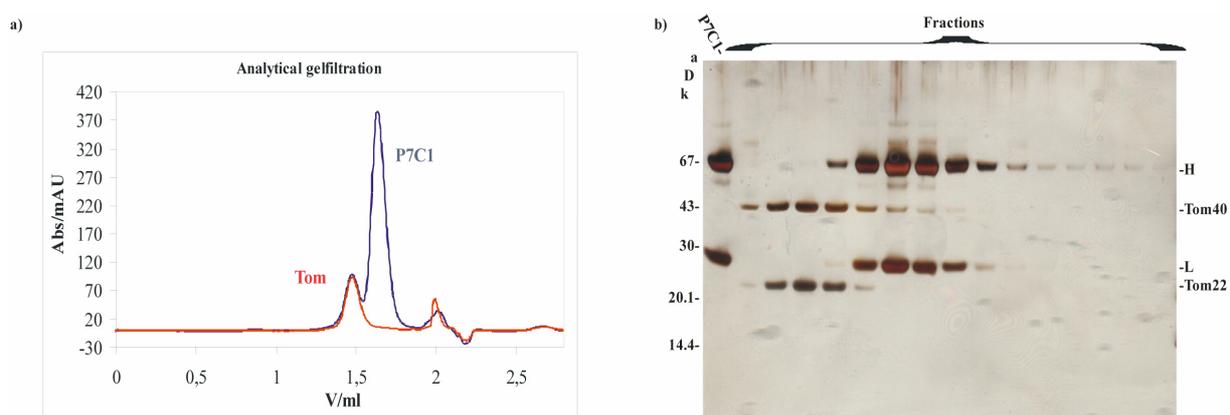
**Figure 2.7** Purification of the mABs from cell culture supernatant. (a) Chromatogram of the protein A purification of antibody P3H6. The first peak with high intensity represents the flow through fraction. (b) Chromatogram of the protein A purification of antibody PIC10. Elution of PIC10 showed a double peak. (c) Coomassie-stained gradient (4-12%) SDS-PAGE of the flow through and elution fractions of the antibodies P3H6, PIC10 and P7C1. M=marker, FT=flow through.

The peak with the high intensity at the beginning represents the protein content in the flow through fraction. In contrast to P3H6, which resembled the standard peak distribution for this kind of purification, the curve of P1C10 exhibited a double peak, indicating a possible disassembly of the antibody at the rather harsh elution condition (Figure 2.7 a and b).

However, between 300 and 500  $\mu\text{g}$  of highly purified mABs were obtained by this purification strategy (Figure 2.7) and subjected to binding studies with TOM core complex (see below).

### 2.3.3.5 Binding studies via analytical gel filtration

To investigate the *in vitro* binding behaviour of the monoclonal TOM-directed mABs and to find the optimal conditions for this interaction, binding assays were performed by analytical size exclusion chromatography. Therefore, the purified mABs (Section 2.3.4.4) were concentrated to 2.5 to 3 mg/ml and the buffer was adjusted to the TOM buffer conditions. The particular antibody and TOM core complex were mixed in a ration 4:1, incubated and run on a SMART-Superose 6 column. None of the tested antibodies was binding to TOM core complex upon gel filtration. No size shift for the TOM:antibody complex in comparison to uncomplexed TOM core could be observed and the antibody fractions ran as the unbound form at an approximate molecular mass of 150 kDa (Figure 2.8).



**Figure 2.8:** Binding assays of TOM core complex with murine antibodies by analytical gel filtration. (a) Superimposed chromatograms of TOM core complex alone (red) and the mixture of TOM with antibody P7C1 (blue). The peak of TOM core complex did not shift and was separated from the antibody fractions. (b) Silver-stained SDS-PAGE analysis of the fractions reflecting the separated peaks of the chromatogram in (a). The first lane shows the purified antibody followed by the peak fractions. H=heavy chain; L=light chain.

SDS-PAGE analysis of the fractions revealed, that TOM core complex and the antibody were separated on the Superose 6 column (Figure 2.8).

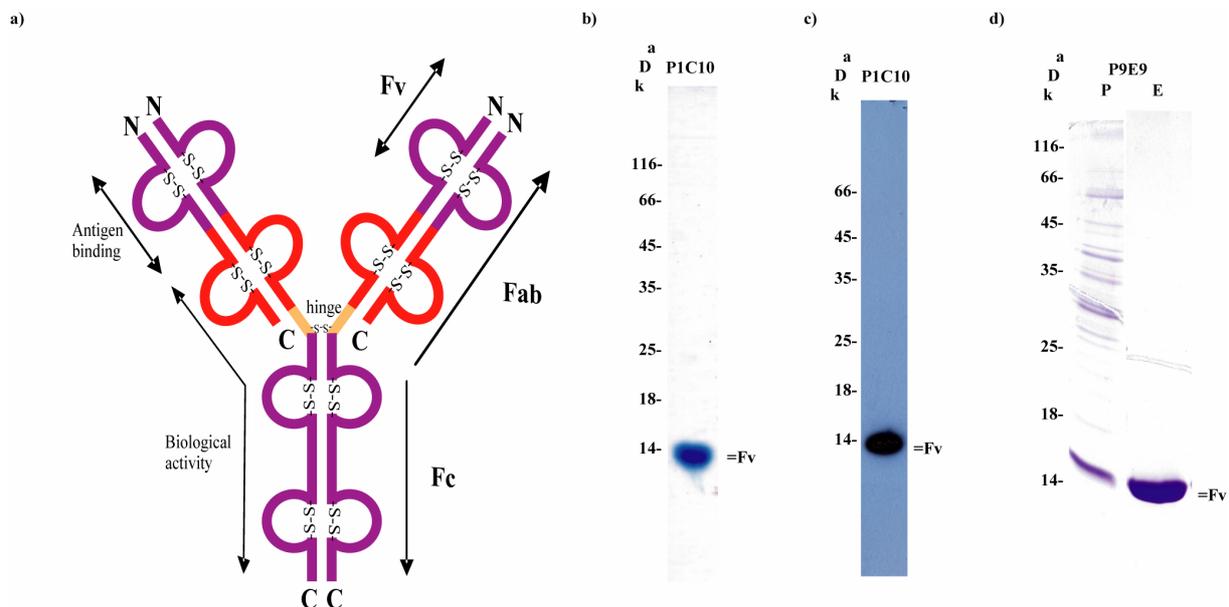
The mAB P1C10 which was the only binding murine antibody in the pull-down assay (see Section 2.3.3.3), eluted at much lower size than expected. Additional experiments using a Superdex 200 column showed that this antibody was unstable and exclusively present in the dissociated form, where the heavy and the light chains are disassembled. This finding is consistent with the protein A purification of P1C10, showing an unusual double peak in the elution (Section 2.3.3.4). Therefore, results from the binding behaviour based on analytical gel filtration are not significant for this antibody.

## 2.3.4 The TOM core in complex with monoclonal antibody fragments

### 2.3.4.1 Expression and purification of Fv fragments

Recombinant Fv fragments consist of the variable domains of the heavy and the light chains. The monoclonal antibodies, which were directed against TOM core complex, were derived from murine cell culture as described above. The mRNA of the selected clones was isolated from the hybridoma cells and transcribed into cDNA. The corresponding genes of the heavy and the light chains, which sequences were unknown, were amplified by PCR using a set of related primers and cloned into the pDRIVE or the TOPO vector (for details, see Materials and Methods). After sequencing, the corresponding genes of the variable domain (Fv) (Figure 2.9a) of the mABs P1C10 and P9E9 were cloned in the bicistronic operon of the plasmid pASK68 allowing periplasmic coexpression of both chains in *E. coli* under the control of the inducible *lac* promoter<sup>237</sup>. The oxidizing environment in the periplasma should ensure proper folding and formation of the disulphide bonds. The antibody containing cleared lysate of the periplasmic fraction was prepared after expression in the *E. coli* strain *JM83*.

Since the vector pASK68 provides a C-terminal fusion of strep-tag I to the heavy chain of the fragment, the Fv domains of P1C10 and P9E9 were purified by streptavidin affinity chromatography (Figure 2.9). The purification using the strep-tag yielded highly pure Fv complex (Figure 2.9b and d). The presence of the light chain ( $V_L$ ) was checked by Western Blot analysis as it carried a myc-tag (Figure 2.9c).



**Figure 2.9:** Purification of the Fv fragments of the antibodies PIC10 and P9E9. (a) Schematic representation of the structural composition of an immunoglobuline. The Fv fragment is the smallest binding part of the antibody. (b) SDS-Page of the purified Fv fragment of the antibody PIC10. Heavy  $V_H$  and light chain  $V_L$  were too similar in size to be resolved. (c) Western Blot against the myc-tag on  $V_L$  showed a positive signal. (d) SDS-Page of the purified Fv fragment of the antibody P9E9. P=periplasma, E=elution.

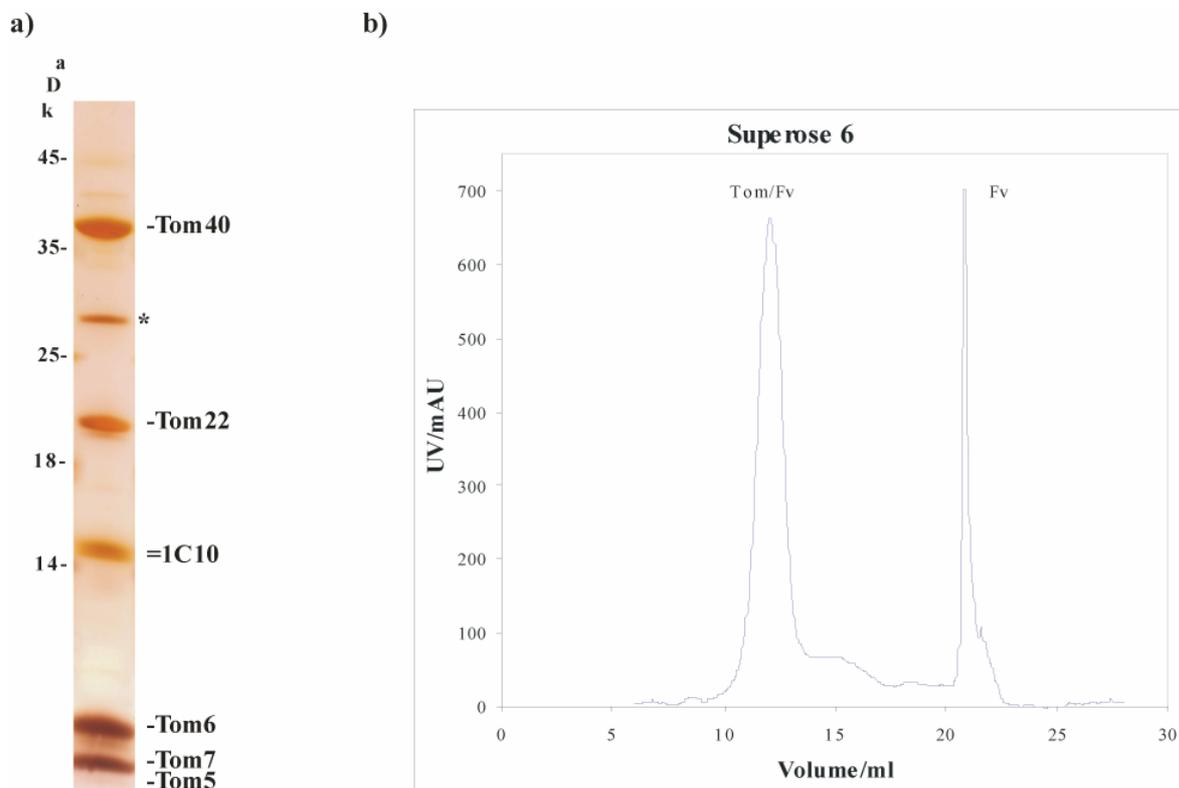
Both Fv fragments P1C10 and P9E9 revealed homogeneous peaks at the right size in gel filtration experiments with a Superdex 75 column. In contrast to P9E9, the expression rate of the Fv fragment of P1C10 turned out to be much lower. This very low expression rate of P1C10 accounted for expression of 48 l culture volume to obtain the required amount of protein necessary for copurification with TOM core complex.

#### 2.3.4.2 Copurification of Fv fragments with TOM core complex

To crystallize TOM protein which is complexed with monoclonal antibody fragments both proteins were mixed and copurified. After the addition of excessive amounts (10:1 molar ratio) of purified Fv fragment to the purified TOM core complex, the resulting TOM:antibody complex was separated from unbound Fv fragment by size exclusion chromatography (Figure 2.10).

The chromatogram of the mixture showed two peaks, corresponding to the TOM core:P1C10 complex and the unbound antibody fraction (Figure 2.10b). It was not possible to define a pronounced size shift of TOM core complex due to its high molecular mass and the limited

resolution of the Superose 6 column. However, the Fv fragment was specifically bound to TOM core complex (Figure 2.10a). Addition of the antibody to the purified TOM complex in the molar ratio of 5:1 was not sufficient to saturate the complex, indicating the presence of more than one antibody binding site. The Fv fragment P9E9 did not bind to TOM core complex under these conditions, which is consistent with all other binding experiments using the corresponding antibody from cell culture (see above). P1C10 which exhibited the highest affinity in the ELISA experiment and which was binding to TOM core complex in the pull-down assay appeared to be the only specific TOM-directed antibody from hybridoma culture. The other six antibodies were not tested, since their affinity to TOM core complex was even lower than for P9E9, according to the ELISA signals.

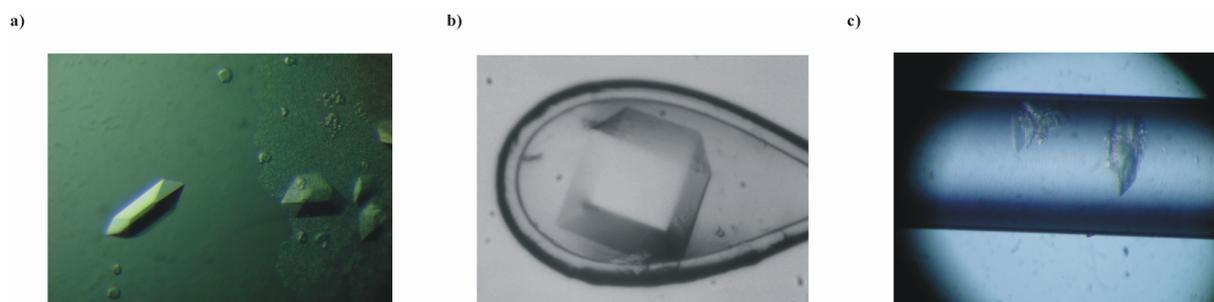


**Figure 2.10:** Copurification of the Fv fragments PIC10 with TOM core complex (a) Silver-stained SDS-urea polyacrylamide gel. The asterisk marks the dimer of the two antibody chains. (b) Chromatogram of the gel filtration using Superose 6. The first peak represents the TOM core:PIC10 complex, whereas the second peak contains the unbound excess of PIC10.

## 2.3.5 Crystallographic analysis of TOM core complex

### 2.3.5.1 Crystallization of TOM core complex

One aim of this study was to determine the three-dimensional structure of the TOM core complex by X-ray crystallography. Therefore, purified TOM core complex was subjected to crystallization experiments according to the protocol established by Simone Schmitt<sup>238</sup>. Improved crystals were obtained within 2 to 3 days with the optimized condition of 100 mM ammonium citrate pH 4.5, 50 mM Tris-Cl pH 7.2 and 22-28% PEG400 (Figure 2.11). The crystallization was done by the hanging drop vapour diffusion method mixing 1.5  $\mu$ l of TOM core complex with a concentration of 14 mg/ml and 1  $\mu$ l of the crystallization solution. The crystals contained all components of TOM core complex, but crystallization appeared to be selective for unprocessed Tom6, as the degradation product of Tom6 was absent in the crystals<sup>238</sup>. The concentration of PEG400 was critical for crystallization and had to be determined independently for each TOM preparation. Additive screens were used to increase the crystal quality, but almost none of the additives could improve the diffraction quality, although some could change the morphology of the crystals. Initially received TOM core complex crystals diffracted only to a resolution of less than 15  $\text{\AA}$  showing only weak individual reflection spots, however, the optimized purification strategy significantly improved resolution to 7  $\text{\AA}$ . The crystallization condition turned out to be already a cryo-protectant, thus crystals were directly frozen in a stream of liquid nitrogen from their mother liquor. This procedure did not cause any damage to the crystals or any loss of resolution which was confirmed by measuring crystals grown in a capillary by counter diffusion (Figure 2.11c). Further improvement of the diffraction quality to 6  $\text{\AA}$  resolution was achieved by using specific detergent additives, allowing to determine the space group as well as the unit cell parameters (see below).



**Figure 2.11:** Crystallization of the TOM core complex (a) Crystals in the standard condition plus 100  $\mu$ M  $\text{CuCl}_2$  as additive. (b) Crystal caught in a cryo loop. (c) Crystals in capillary.

## 2.3.5.2 Optimization of the diffraction quality of the crystals

### 2.3.5.2.1 Experiments to improve crystallization

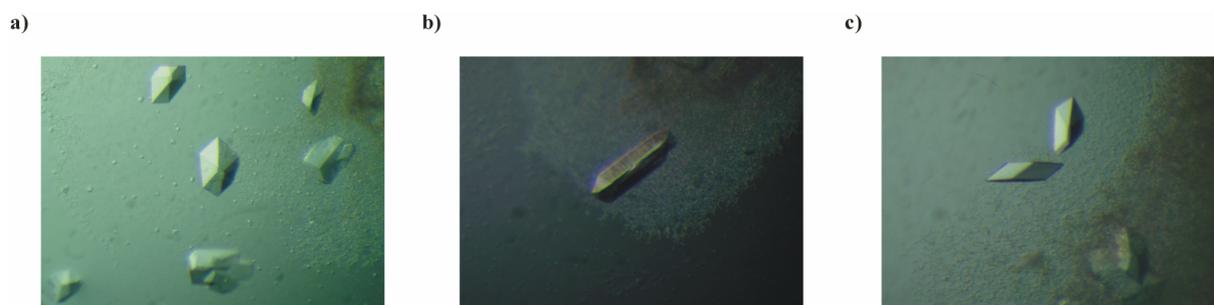
Since the resolution limit of 7-8 Å abolished elucidation of the three-dimensional protein structure of TOM core complex, the diffraction properties of the crystals had to be improved considerably. All additives from the detergent screens from Hampton Research were tested in the standard TOM crystallization condition and used at CMC (critical micelle concentration). For all additives, where crystal growth occurred (Table 2.2), fine screens were performed to find the optimal concentration of PEG400. Optimized crystals, which were large enough for measuring, were examined using synchrotron radiation. For most of the additives, the diffraction quality of the crystals became worse and the resolution limit was significantly higher than 8 Å (Table 2.2). For maltosides with shorter carbon chain length than n-Dodecyl-β-D-maltoside, such as n-Undecyl-β-D-maltoside and n-Nonyl-β-D-maltoside (Figure 2.12), the resolution limit was comparable to DDM, although crystals in n-Undecyl-β-D-maltoside were smaller.

Detergent	MM (g/mol)	CMC (mM)	Type	Synchrotron	max. Resolution
C <sub>12</sub> E <sub>9</sub>	583.1	0.08	N	Y	>7 Å
C <sub>12</sub> E <sub>8</sub>	539.1	0.11	N	Y	>7 Å
n-Dodecyl-β-D-maltoside	510.6	0.17	N	Y	7 Å
Sucrose monolaurate	524.6	0.2	N	N	
CYMAL®-6	508.5	0.56	N	Y	>7 Å
n-Decyl-β-D-maltoside	482.6	1.8	N	N	
ZWITTERGENT® 3-12	335.6	4.0	Z	Y	>7 Å
n-Hexadecyl-β-D-maltoside	566.6	0.0006	N	N	
n-Tetradecyl-β-D-maltoside	538.6	0.01	N	N	
n-Tridecyl-β-D-maltoside	524.6	0.033	N	N	
n-Undecyl-β-D-maltoside	496.6	0.59	N	Y	7 Å
n-Decyl-β-D-thiomaltoside	498.6	0.9	N	Y	>7 Å
FOS-Choline®-12	315.5	1.5	Z	Y	<7 Å
1-s-Nonyl-β-D-thioglucoside	322.4	2.9	N	Y	>7 Å
n-Nonyl-β-D-thiomaltoside	484.6	3.2	N	N	
n-Nonyl-β-D-maltoside	468.4	6.0	N	Y	7 Å

**Table 2.2:** List of detergent additives in which crystals grew still under the typical TOM core complex crystallization condition. N=nonionic, Z=zwiterionic. Yes and No in the column “Synchrotron” means, if the crystals were measured or not.

In case of the detergents sucrose monolaurate, n-Decyl-β-D-maltoside and the maltosides with longer carbon chains than DDM, crystals were too small for measuring. FOS-Choline®-12

seemed to be the only detergent of the additive screen, which slightly improved the diffraction quality of the TOM core complex crystals to a maximal resolution of  $\sim 6.5$  Å (Table 2.2, Figure 2.12).



**Figure 1.12:** Crystals of TOM core complex with different additives (a)  $C_{12}E_8$ . (b) FOS-Choline-12. (c) *n*-Nonyl- $\beta$ -D-maltoside.

Since it was shown that copper ions can block channel activity of reconstituted TOM complex in conductivity measurements<sup>221</sup>,  $CuCl_2$  was tried as an additive for crystallization. A fraction of TOM core complex (10-14 mg/ml) mixed with  $CuCl_2$  is precipitating and if the mixture is used for crystallization, crystal grew to bigger size, but without any significant alteration of diffraction properties.

All crystallized membrane protein complexes show defined lipid-protein contacts and lipid requirement may also be necessary for stabilization of membrane proteins. Therefore, different lipids were used as additives to enhance ordered crystallization of TOM core complex, but addition of the lipids L- $\alpha$ -phosphatidylcholine, L- $\alpha$ -phosphatidylethanolamine, L- $\alpha$ -phosphatidylinositol and cardiolipin prevented crystal growth of the protein under the original crystallization condition.

Due to the thermostability of TOM core complex, another trial to improve crystallization was performed by incubation of the protein at  $45^\circ C$  for 20 min and subsequent crystallization, but no amelioration of the crystal order could be achieved by this experiment.

#### 2.3.5.2.2 Manipulation of the crystals

The direct manipulation of already grown crystals represents an alternative option to the optimization of the crystallization conditions. It is known, that glutaraldehyde, a frequently used amine-reactive homobifunctional crosslinker reagent, can link protein molecules in a crystal and thereby improve the order of a crystal (Prof. Cramer, LMU Munich, personal communication). Such a protein crystal, in which the molecules are sufficiently crosslinked, will not dissolve in water anymore. Crosslinking of the protein molecules in the TOM core

complex crystal with this reagent was possible, but this treatment appeared to disturb the crystal order as the diffraction became even worse.

As typical for membrane proteins, it is expected that TOM core complex crystals possess a high solvent content, since crystals turned out to be very soft. An alternative to improve crystal order might be the stepwise transfer of TOM core complex crystals to increasing PEG400 concentration, thereby reducing the solvent content and possibly transforming the crystals to higher order. However, the crystals could not be transformed using this method.

### 2.3.5.2.3 Heavy metal atom soaks and cocrystallization

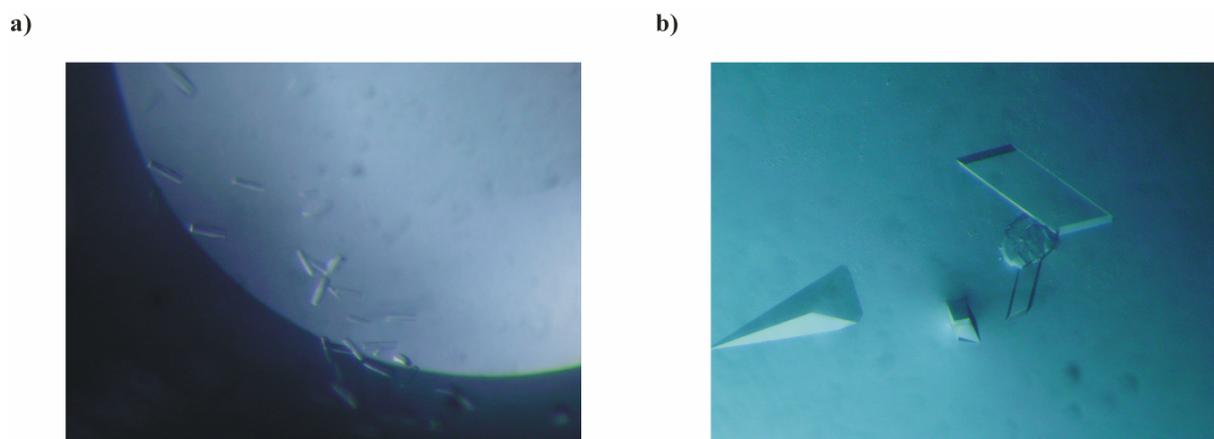
Owing to the high solvent content of protein crystals, it is possible to soak chemicals and in some cases even small proteins into crystals which bind at defined positions of the protein. Heavy metal atoms or heavy metal atom clusters are able to bind specifically to sites of a protein and are therefore essential for determining the crystallographic phases to solve the three-dimensional structure. Soaking of crystals with heavy metal atoms can not only deliver initial phases, but sometimes also improve the diffraction properties of the crystals by occupation of specific symmetry planes or alteration of crystal contacts. For this experiments, the Heavy Atom Screen kits (Hampton Research) and the collection of heavy metal atom clusters like  $W_6Br_{12}$ ,  $W_{18}$ ,  $K_2Ta_6Br_{12}$ ,  $Mo_6Cl_{14}$  and  $K_6Mo_7O_{22}(O_2)_2$  (kindly provided by Prof. Patrick Cramer, Gene Center, Munich and Prof. Huber, Max-Planck-Institute of Biochemistry, Munich) served as the platform to try various different heavy metal compounds. Additionally, heavy metal atoms were used for co-crystallization trials with TOM core complex. In case of  $Ta_6Br_{12}^{2-}$ , which shows a green colour in solution, the soaked crystals adopted the colour, indicating the presence of the cluster in the crystals (Figure 2.13). However, the diffraction quality of the crystals suffered for each employed heavy metal and consequently, determination of initial phases was not possible.



**Figure 2.13:** *TOM core complex crystals, soaked with the TaBr heavy metal atom cluster.*

### 2.3.5.3 Crystallization of the TOM core:P1C10 antibody complex

As reported in Section 2.3.4.2, the complex of TOM core protein with the Fv fragment P1C10 was copurified and subjected to crystallization. Therefore, the initial crystallization conditions as well as conditions for robot screening (200 nl + 200 nl drops) were tested to grow crystals.



**Figure 2.14:** Crystals of TOM core:P1C10 complex. (a) Initial crystals from the robot screening sitting drop plate (200 nl + 200 nl drops). (b) Optimized crystals using the hanging drop technique.

Crystals occurred in the original TOM core complex condition and in a similar condition from the Nextal Classics Screen 0.2 M  $\text{CaCl}_2$ , 0.1 M HEPES pH 7.5 and 28% PEG 400 (Figure 2.14). After optimization in both cases, crystals were large enough for measuring using synchrotron radiation. Although crystal morphology was different, space group, cell dimensions and diffraction quality were the same as for crystals of the TOM core complex. SDS gel analysis of washed and dissolved crystals revealed that all components of the TOM core complex and both chains of the antibody were present in the crystals.

#### 2.3.5.4 Detergent exchange during purification

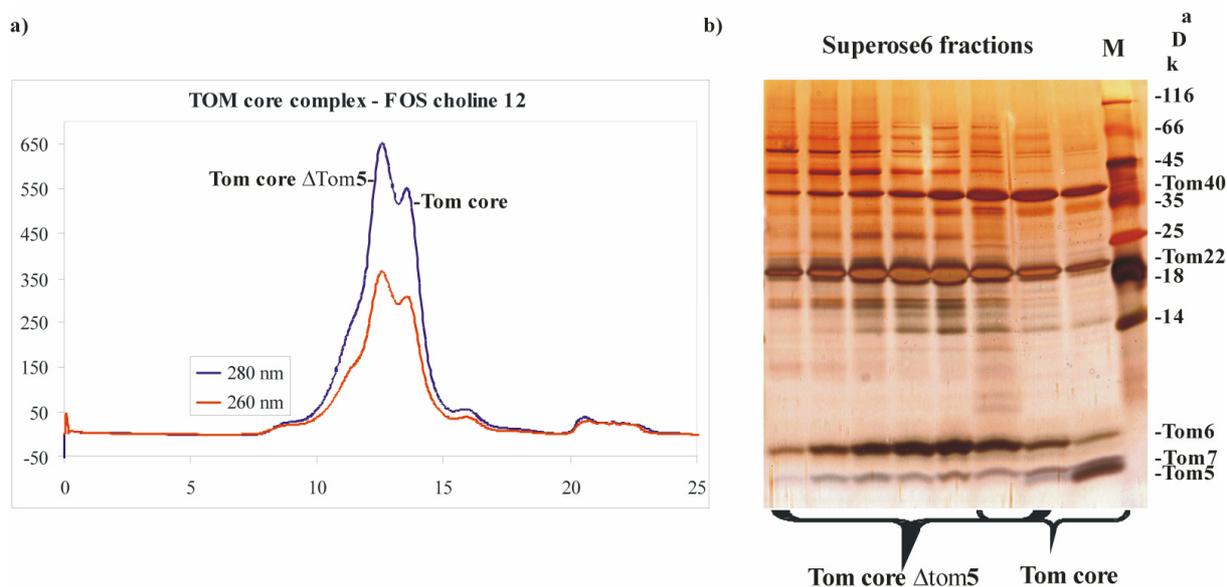
Stability and the crystallization characteristics of membrane proteins are highly dependent of the choice of the detergent. Since the optimized solubilization and purification of TOM core complex yielded highly pure and sufficient amounts of protein, it would not be suitable to switch to another detergent during the first steps of the isolation. However, the detergent exchange upon gel filtration allowed monitoring of the complex stability as its disassembly would significantly alter the peak distribution. Detergents which did not negatively influence the crystallization in the detergents additive screen from Section 2.3.5.2.1 were used in concentrations above their CMC in the gel filtration experiment. TOM core complex was

purified according to the standard procedure, but in the final step was run on a Superose 6 column with the new detergent in the mobile phase, respectively. As it is reported, that the use of n-Dodecyl- $\alpha$ -D-maltoside instead of n-Dodecyl- $\beta$ -D-maltoside can improve the diffraction quality markedly<sup>239</sup>, this was also tried.

Detergent	CMC (%)	Concentration (%)
n-Dodecyl- $\alpha$ -D-maltoside	0.009	0.03
n-Undecyl- $\beta$ -D-maltoside	0.029	0.10
n-Decyl- $\beta$ -D-maltoside	0.087	0.26
n-Nonyl- $\beta$ -D-maltoside	0.280	0.84
FOS-choline 12	0.047	0.14

**Table 2.3:** Detergents of the exchange experiment with CMC and applied concentration.

Crystallisation trials were performed by hanging drop vapour diffusion using the original TOM core complex crystallization condition in DDM and by sitting drop screening using the Nextal system as well as a crystallization-robot to perform drops on nanolitre scale. Fractions of the size exclusion were analyzed by urea SDS-PAGE using silver staining. For all detergents of the maltoside class the TOM core complex was fully assembled and showed standard running behaviour on the Superose 6 column.



**Figure 2.15:** Detergent exchange with FOS choline 12 (a) Peak distribution upon gel filtration shows two peaks. (b) Silver-stained urea SDS-PAGE of the fractions. TOM core complex of the first peaks lacked the subunit Tom5 (named Tom core  $\Delta$ Tom5). M = marker.

In case of the detergent FOS choline 12, the size exclusion chromatography exhibited a distribution of two peaks, which corresponds to TOM core complex with and without the

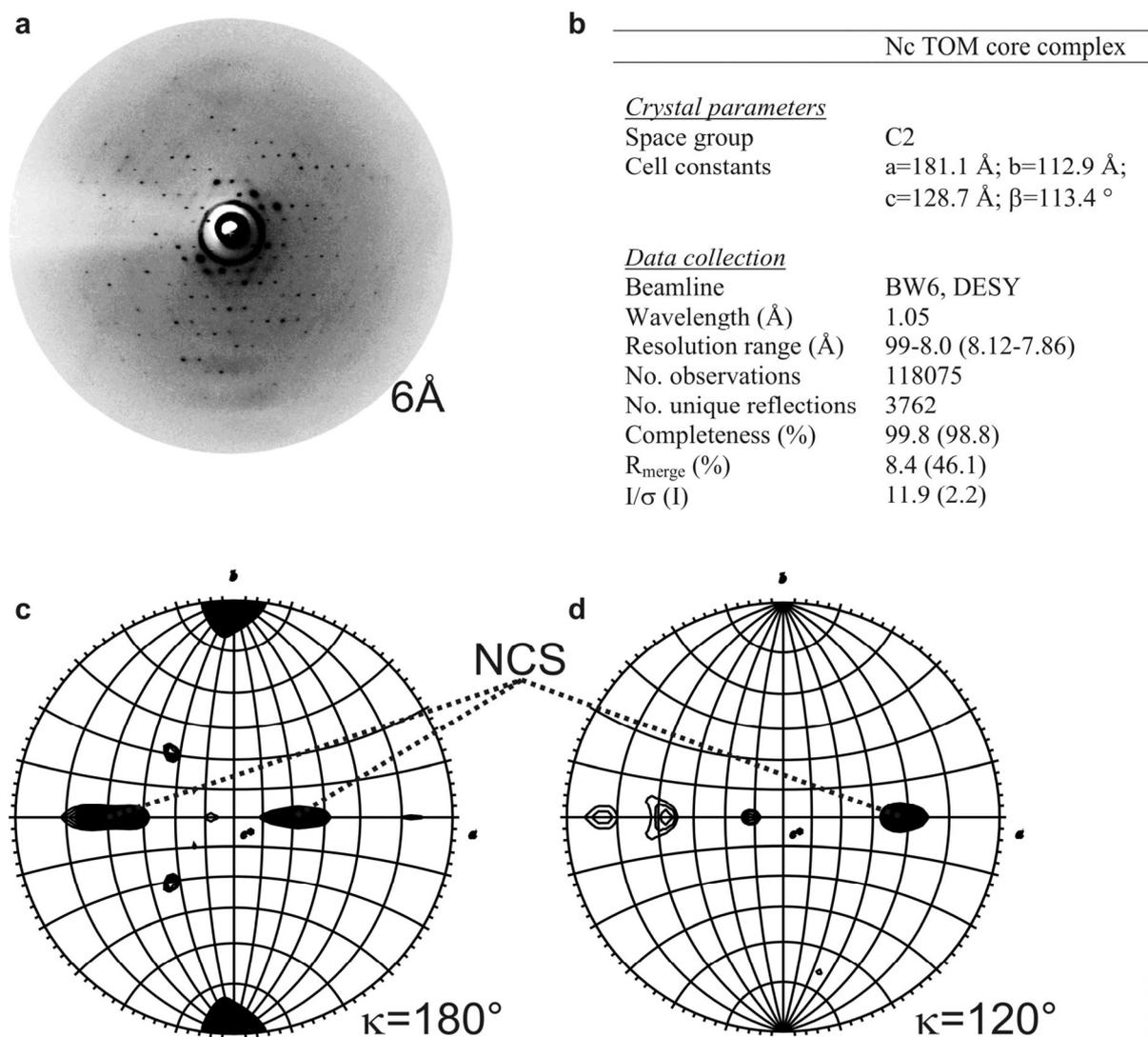
subunit Tom5 (Figure 2.15). In all these crystallization trials TOM core complex crystallized only in the detergent n-Undecyl- $\beta$ -D-maltoside so far. However, these crystals diffracted comparable to the crystals obtained in DDM despite their smaller size.

### 2.3.5.5 Crystallographic data of TOM core complex

The purified TOM core complex was subjected to crystallization and after optimization large crystals of different morphologies were obtained (Section 2.3.5.1). The change of the purification strategy and the use of additives improved the diffraction quality and allowed for the first time determination of crystallographic data such as the space group and the unit cell parameters (Figure 2.16b). The measurement of the crystals took place at the synchrotron radiation sources BW6 at Deutschen Elektronen Synchrotron (DESY) in Hamburg, Germany and at the X06SA Swiss Light Source (SLS) at Paul Scherrer Institut in Villingen, Switzerland. The crystals of the TOM core complex were exposed to X-rays both at RT in capillary and flash-cooled at 100 K in a nitrogen-gas stream (for fast-freezing crystals were taken directly out of the mother liquor which served as a cryoprotectant). Crystals of the Tom:antibody complex were only measured frozen. The frozen crystals were stable in the beam, which allowed collection of the first isomorphous complete dataset at 8 Å resolution (Figure 2.16).

Indexing, processing and reduction of the data were possible, but low resolution did not allow determination of phases. Data were processed with DENZO and data reduction was performed with SCALEPACK<sup>240</sup>. The space group was determined to be C2 with unit cell dimensions of  $a=181$  Å,  $b=113$  Å,  $c=129$  Å,  $\beta=113^\circ$  and these parameters remained constant for all measured crystals (Figure 2.16).

Only one out of 20 crystals showed diffraction in the range of 6-8 Å and morphology of crystals did not correlate with diffraction quality, but resolution appeared to be dependent on the protein batch. No radiation damage was observed. With the estimative molecular mass (gel filtration) of about 400 kDa, the relatively high solvent content of 85% was calculated, which is consistent with the resolution limit and the softness of the crystals. Based on the analyzed structure factors, the autocorrelation of the squared structure factors (Patterson selfrotation function) at different kappa angles of  $180^\circ$ ,  $120^\circ$ ,  $90^\circ$ ,  $72^\circ$  and  $60^\circ$  resulted in noncrystallographic (ncs) two- and three-fold molecule symmetry of the TOM core complex (Figure 2.16).



**Figure 2.16:** Crystallographic data on TOM core complex (**a**) Diffraction image. The plate frame reflects 6 Å resolution. (**b**) Result table of the processed data at 8 Å. (**c**) Selfrotation with  $\kappa = 180^\circ$  (twofold axis) and (**d**)  $\kappa = 120^\circ$  (threefold axis). Owing to the crystallographic axis, an additional twofold axis, which was orthogonal to the twofold ncs axis, was generated

The three-fold ncs symmetry axis was not along the two-fold ncs and crystallographic axis. The four-, five- and six-fold molecule symmetry could be excluded. The analysis of the Patterson showed no pseudo origin peak and therefore two-fold ncs symmetry along the crystallographic axis also could be excluded.

## 2.3.6 Investigation of the pore-forming component Tom40

The main component of the TOM complex is Tom40, which constitutes the pore of the translocase. The stoichiometry of the complex and how many Tom40 molecules are present in one TOM core complex molecule are not known so far. Tom40 was predicted to adopt  $\beta$ -barrel topology like many translocases of the bacterial outer membrane (outer membrane proteins = OMPs) such as porins. Due to the size of the complex, more than one Tom40 molecule would be necessary to build the protein conducting channel. The recombinant expression and refolding of Tom40 should lead to crystallization-grade protein and its structure determination would provide essential information about the topology and the working mechanism of the TOM complex.

### 2.3.6.1 Secondary structure prediction of Tom40

An important step in investigating the architecture and topology of Tom40 is primary sequence analysis. For sequence analysis and the prediction of protein structure and function PredictProtein was used (Figure 2.17), which returns multiple sequence alignments, PROSITE sequence motifs, low-complexity regions (SEG), prediction of secondary structure, transmembrane helices, coiled-coil regions, disulfide-bonds (DISULFIND), sub-cellular localization, and functional annotations.

```

.....10.....20.....30.....40.....50.....60
AA      MASFSTESPLAMLRDNAIYSSLSDAFNFAFERRKQFGLSNPGTIE TIAREVQRDTLLTNY
PROF_sec          HHHHHHHHHHHH          HHHHHHHHHH
Rel_sec          966676564212245765220011231000100026776767466785462024410232
SUB_sec          LLLLLLLI.....LLLI.....LLLLLLI.HHHHH.H.....

.....70.....80.....90.....100.....110.....120
AA      MFSGLRADVTKAFSLAPLFQVSHQFAMGERLNPYAFAALYGTNQIFAQGNLDNEGALSTR
PROF_sec          EEEEE          EEEEEEEEE          EEEEE          EEEEEEE          EEEE
Rel_sec          001023431034444541345444553354564035566415405786454366403336
SUB_sec          .....L.....E.....EE.....L.LI.....EEEE.....L.....EEEE.E.....LI.....E

.....130.....140.....150.....160.....170.....180
AA      FNYRWGDRITIKTQFSIGGGQDMAQFEHEHLGDDFSASLKA INPSFLDGGLTGIFVGDYL
PROF_sec          EEEEE          EEEEEEE          EEEEEEE          EEEEE          EEEEEHHHH
Rel_sec          654303430446566741563356654210254022432411553223540112201132
SUB_sec          EE.....EEEE.....LI.....EEEE.....L.....LI.....L.....

```

```

.....190.....200.....210.....220.....230.....240
AA      QAVTPRLGLGLQAVWQRQGLTQGPDTAISYFARYKAGDWWVASAQLQAQALNTSFWKLLT
PROF_sec H      EEEEEEEEE   EEEEEEEEE   EEEEEEE   EEEEEEEEE
Rel_sec  012433013014565045565565103443334125422677785244315666655305
SUB_sec  .....EEE..LLLLLL.....L..EEEEEE.....EEEEEE..L

.....250.....260.....270.....280.....290.....300
AA      DRVQAGVDMTSLVAPSQSMGGLTKEGITTGAKYDFRMSTFRAQIDSKGKLSCLLEKRL
PROF_sec EEEEEEEEE   EEEEEEEEE   EEEEEEE   EEEEEEEEE
Rel_sec  520344566531135432223445430456522453132035767547550211111120
SUB_sec  L.....EEEE..L.....L.....EEE..E.....EEEE.LLL.....

.....310.....320.....330.....340.....
AA      GAAPVTLTFAADVVDHVTQQAKLGMSVSI EASDVDLQEQQEGAQSLNIPF
PROF_sec EEEEEEEEE   EEEEEEE   HHHHHH
Rel_sec  467605788755303346530013677516860055430147777778
SUB_sec  .LLL.EEEEEEE.....LL.....EEEE.LLL.HH.....LLLLLLL

```

**Figure 2.17:** Secondary structure prediction of Tom40 from *Neurospora crassa*. PROFsec: predicted secondary structure in three states: helix H, sheet E, rest L. RELsec: reliability index for secondary structure prediction (0=low to 9=high). SUBsec: subset of the prediction, for all residues with an expected average accuracy >82%.

Based on its functional and structural similarity with other  $\beta$ -barrel proteins such as Tob55 and porin, the secondary structure of Tom40 should mainly reveal  $\beta$ -strands. A beta barrel is a large beta-sheet that twists and coils to form a closed structure in which the first strand is hydrogen bonded to the last. Beta-strands in beta-barrels are typically arranged in an antiparallel fashion. Barrel structures are commonly found in porins and other proteins that span cell membranes. According to the program SUBsec, representing the subset of all residues, for which secondary structure elements are predicted with an average accuracy above 82%, Tom40 from *Neurospora crassa* (ncTom40) contains 14  $\beta$ -strands flanked by one  $\alpha$ -helix at the N- and C-terminus, respectively. None of the helices appeared to be a transmembrane helix as calculated by the program PHDsec. Presumably, the 14  $\beta$ -strands of each pore-constituting Tom40 molecule form the  $\beta$ -barrel, which is located in the membrane, whereas the soluble flanking helices are exposed to the cytosol and the intermembrane space (IMS). For the first approximately 40 amino acid residues of ncTom40, no pronounced secondary structure could be observed, whereas in the human homologue same is true for the first 80 amino acid residues. Interestingly, the human Tom40 lacks the C-terminal  $\alpha$ -helix. Tom40 from *N. crassa* contains only one cysteine and according to DISULFIND, a program calculating disulfide bonding state, this cysteine is not disulfide-bonded. The human homologue for example possesses six cysteines, which were also predicted to be non-bonded.

The predicted secondary structure was composed of 9.17%, 43.84% and 46.99% of the types helix, strand and loop, respectively.

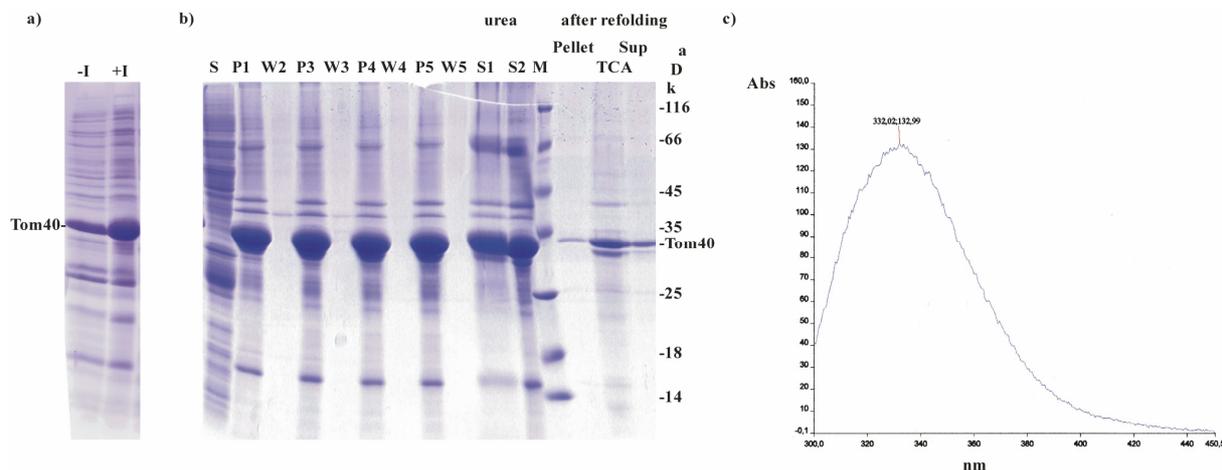
### 2.3.6.2 Sequence conservation of Tom40

The investigation of the sequence conservation among homologues together with the secondary structure prediction could provide important information about the domain architecture of a protein. Multiple sequence alignment of Tom40 family proteins revealed the high sequence conservation among higher eukaryotes with 27% sequence identity and 47% sequence similarity between the Tom40 homologues from *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster* and *Caenorhabditis elegans*. The protein sequence of Tom40 is even well conserved between higher and lower eukaryotes with 44% similarity between human and yeast. The N-terminal region (27 to 76 amino acid residues in length, depending on the species), which lacks secondary structure elements (Section 2.3.7.1) is not conserved in length and charge among species. In contrast to higher eukaryotes fungi contain a C-terminal extension of 20-27 amino acid residues. As shown by the secondary structure prediction, this C-terminal flanking residues seem to form an  $\alpha$ -helix in *Neurospora crassa*, whereas the human Tom40 for example lacks this extension and the  $\alpha$ -helix at its C-terminus.

### 2.3.6.3 Recombinant expression and refolding of Tom40

For some membrane proteins of the  $\beta$ -barrel class, recombinant expression in *E. coli* and subsequent refolding in the presence of detergent was successful and crystal structures of these proteins were determined<sup>241</sup>. The expression, refolding and crystallization of the pore component Tom40 might represent an efficient alternative to the crystallization and structure determination of TOM core complex.

The coding sequences of Tom40 from the species *H. sapiens*, *C. elegans* and *N. crassa* were cloned into the pRSET vectors PL0 and PL1 and proteins were expressed in *E. coli*. The vector PL0 does not contain any tag, whereas PL1 introduces an N-terminal His-tag to the expressed protein. Solely Tom40 from *N. crassa* was expressed in these vectors (Figure 2.18a) and also usage of the pET-Duet<sup>TM</sup>-1 system for hsTom40 and ceTom40 did not result in noticeable expression. All variants, ncTom40 (aa 38-329), hsTom40 (aa 77-361) and ceTom40 (aa 17-301) were truncated to obtain the conserved core domain of Tom40.



**Figure 2.18:** Expression and refolding of Tom40 from *Neurospora crassa*. (a) Coomassie stained SDS-gel of *ncTom40* expression.  $-I$ =before induction;  $+I$ =after induction. (b) Coomassie stained SDS gel of the isolation of inclusion bodies with subsequent refolding.  $S$ =cell lysat;  $P1$ ,  $P3$ - $P5$ =inclusion bodies after different washing steps;  $W2$ - $W5$ =different washing steps;  $S1$  and  $S2$ =inclusion bodies dissolved in 8M urea before and after centrifugation;  $M$ =marker; Pellet and Sup represent the pellet and the supernatant after refolding and centrifugation. Most of Tom40 is in the supernatant. (c) Fluorescence spectroscopy of refolded Tom40. The maximum peak of the emission is at 332 nm for folded proteins.

Tom40 from *N. crassa* was highly expressed in inclusion bodies (Figure 2.18a), which were isolated in different washing and centrifugation steps (Figure 2.18b). Purified inclusion bodies were treated with 8M urea to completely unfold the Tom40 molecules, centrifuged and refolded. For the refolding two methods were tried: stepwise dialysis of the unfolded protein against decreasing concentrations of urea as well as sudden dilution of Tom40 in urea-free buffer. The first option of dialysis failed as the protein precipitated during the process, whereas the dilution assay succeeded in diluted, but folded Tom40. Most of Tom40 was soluble after refolding as depicted in Figure 1.18b.

The folding state of the protein was analyzed by fluorescence spectroscopy. The protein was excited at 280 nm and the maximum of fluorescence emission is located at around 332 nm, which indicates folding of Tom40. Unfolded proteins show maximas at 350 nm. Concentration and further purification of Tom40 should result in crystallization-grade protein and will be subjected to crystallization.

## 2.4 Discussion

### 2.4.1 Membrane protein crystallization

#### 2.4.1.1 Detergents in crystallization of TOM core complex

Twenty to thirty-five percent of the proteins encoded by an organism's genome are integral membrane proteins. Membrane proteins such as channels, transporters, and receptors are critical components of many fundamental biological processes<sup>242</sup>. Most of the successful experimental methods for obtaining membrane proteins crystals suitable for structure determination by X-ray crystallography are identical to those used for soluble proteins. The major difference is the necessity for inclusion of detergents above the critical micelle concentration (CMC) in the membrane protein solution. The most typical use of a detergent is to maintain a target membrane protein in a functional, folded state in the absence of a membrane. Eukaryotic multisubunit membrane proteins such as the TOM core complex cannot be overexpressed in *E. coli* and have to be isolated directly from its native environment, for example the outer mitochondrial membrane. When the membrane is removed during the solubilization of the protein, it must be replaced with a detergent as solvent. Despite the large number of detergents that are commercially available, no single "universal detergent" is ideally suited to all biochemical applications. As a result, the choice of detergent is one of the most fundamental decisions in isolation and crystallization procedures of membrane proteins. With essentially no exception, membrane proteins purified for structural studies are isolated in alkyl-chain detergents with generally 7-12 carbon in length and with headgroups that are typically uncharged or zwitterionic, but conceivably could also be charged. However, most membrane proteins have been crystallized from maltosides, glucosides, dimethyl *N*-amine oxides (*e.g.* LDAO) and  $C_nEm$  polyoxyethylene detergents<sup>242</sup>.

The detergent *n*-dodecyl- $\beta$ -D-maltoside (DDM, the alkyl-chain is 12 carbons in length) is used for the isolation of TOM core complex from mitochondria which results in highly pure and stable protein. The TOM core complex in DDM is functional: it has the characteristics of the general insertion pore, contains high-conductance channels and binds preprotein in a targeting sequence-dependent manner. In contrast to the holo complex, it forms a homogeneous double ring structure in electron microscopy<sup>231</sup>. DDM is one of the gentler detergents, and has very favourable properties for maintaining the functionality of more aggregation-prone membrane proteins in solution. Its main drawback is the formation of large

micelles making crystallization difficult. This may be responsible for the limited diffraction quality of TOM core complex crystals. The crystal lattice of a membrane protein crystal has to accommodate three components: protein, detergent and aqueous solution. Protein-protein contacts from the principal scaffolding of the lattice, and proteins with large extra-membranous domains are often favoured in crystallization trials. Small-micelle detergents such as octyl glycoside (OG) and lauryldimethylamine oxide (LDAO) form smaller belts around the transmembrane region of a protein. Potentially, this allows more contacts between exposed polar surface of the protein, and hence stronger lattices and better diffracting crystals<sup>243</sup>. Unfortunately, these detergents are often destabilizing and their application results in partial dissociation of TOM complex<sup>233</sup>. The use of different detergents can yield different crystal forms, as in the case of the bacterial photosynthetic reaction centre from *Rhodobacter sphaeroides*<sup>244</sup> and often, the best-quality crystals may be obtained only in one or a small number of detergents. Due to the limited quality of TOM core complex crystals in DDM, detergent exchange was performed in this study by size exclusion chromatography to screen for more suitable detergents. The low CMC of DDM makes it considerably easier to remove DDM by dialysis or gel filtration. Membrane protein detergent exchange can be accomplished by a variety of methods. However, the size exclusion approach allows monitoring the stability of TOM core complex in various detergents by observation of the peak shape and the retention time. Thereby, aggregation and oligomerization are easily detectable. A very successful strategy for improving poorly-diffracting membrane protein crystals is to first optimize a lead condition by testing closely related detergents (for example, undecylmaltoside versus dodecylmaltoside). In general, the head group has a strong influence on the interactions of detergents with proteins, while the length of the alkyl chain affects the detergent CMC and aggregation number. Upon variation of alkyl chain-length the physical properties remain very similar, whereas the slight but significant differences in micelle size can result in improved packing within the crystal lattice<sup>243</sup>. For example, good-diffracting crystals of bovine mitochondrial cytochrome *c* oxidase were obtained with the detergent decylmaltoside and varying of the detergent's chain-length degraded crystal quality<sup>245</sup>. In case of TOM core complex maltosides with longer alkyl chain length like tridecylmaltoside<sup>238</sup> and shortened maltosides like undecyl- and decylmaltoside were exchanged but could not improve crystallization. The  $\alpha$ -anomer of DDM (*n*-dodecyl- $\alpha$ -D-maltoside) is generally less soluble and not commonly used. Although it was essential for the crystallization of the NhaA Na<sup>+</sup>/H<sup>+</sup> antiporter<sup>239</sup>, no crystals of TOM core complex were obtained in  $\alpha$ -DDM. A next level of optimization might involve the use of detergent mixtures as in case of the crystallization of bacterial outer membrane protein TolC<sup>246</sup>.

In addition to the choice of a suitable detergent its concentration in the protein-detergent complex (PDC) may be critical. Too much detergent can denature the protein or impede crystallization. Too little and the protein can become insoluble. Detergents self-assemble into relatively small, well-defined micelles that typically contain several hundred molecules<sup>243</sup>. The number of molecules in the micellar particle is termed the aggregation number, a characteristic of each detergent. Every detergent possesses a critical micelle concentration (CMC). At concentrations above the CMC, there is equilibrium between monomers and an increasing concentration of micelles.

After the final purification step, the protein is often concentrated many-fold by centrifugation which can increase the detergent concentration, with possible negative consequences. This is especially true in case of a final gel filtration polishing step, which dilutes the protein. Excess detergent can be denaturing and can lead to a large amount of unwanted phase separation in crystallization experiments. Existing methods to decrease detergent concentration include dialysis, the use of absorbing materials such as BioBeads<sup>247,248</sup> or other detergent removing gels (e.g., Extracti-Gel D, Pierce). However, for low CMC detergents possessing large micelles (such as dodecylmaltoside) dialysis fails. The application of the largest possible molecular mass cut-off filter will reduce the amount of detergent that is concentrated along with the protein. Due to the large size of TOM core complex (~400 kDa), centrifugal devices with 100 kDa cut-off can be used for concentration and even the large DDM micelles (~60 kDa) should pass the filter.

Controlling the detergent concentration may be important for crystal reproducibility. A variety of methods exist for determination of detergent concentration. For example, for detergents that contain a sugar headgroup (such as maltosides or glucosides), phenol/sulphuric acid hydrolysis and reaction with molybdate is a valuable colorimetric assay. Furthermore attenuated total reflection Fourier transformation infrared spectroscopy (ATR-FTIR) and quantitative thin-layer chromatography (TLC) can be used to determine the amount of detergent. It might be important for crystallization of TOM core complex to monitor the detergent concentration during the purification and crystallization procedures by one of these methods.

#### **2.4.1.2 Lipid requirements of TOM core complex**

Membrane proteins are isolated as protein-lipid-detergent particles and in addition to the detergent influence the lipid composition is often a critical aspect for preparing samples suitable for structural analysis. Indeed, specific lipids are sometimes required to maintain the

structural stability of membrane proteins. The detergent choice correlates with the presence of a particular amount of lipids in the PDCs as detergents can directly remove tightly bound lipids that may be important for the native structure of the protein. Ordered lipids have been observed in several high-resolution membrane protein crystal structures<sup>249,250</sup>, and excessive detergent concentrations may lead to protein aggregation by stripping away these essential bound lipids. The presence of excess lipids in the PDC can also reduce the chances for crystallization since they reduce monodispersity<sup>251</sup>, but lipid/detergent mixtures may be favourable for maintaining protein function. For example, the first successful functional reconstitution of the lactose permease required a mixture of OG and lipids to preserve protein function during purification<sup>252</sup>. Without added phospholipids the protein did not survive OG solubilization. Twenty-five years later, the addition of a mixture of *E. coli* phospholipids was found to improve crystals of lac permease purified in DDM<sup>253</sup>. It can be very difficult to adjust the appropriate lipid levels in a protein-detergent complex. For instance, suitable crystals of the glycerol 3-phosphate transporter (Glp3T) were not obtained if too much lipid was removed during purification<sup>254</sup>, whereas optimal crystals of the bacterial cytochrome *b<sub>6</sub>f* complex were only obtained when lipids were added to the purified protein<sup>255</sup>.

Characterization of PDCs in respect to their lipid content is critical for the optimization of downstream sample preparation and crystallization. FTIR and TLC can both determine the amounts and specific types of lipids present. From experiments where lipids were added to the crystallization trials of TOM core complex no crystals were obtained, but determination of the lipid composition in the isolated complex might give hints about its lipid requirements and for optimization of its crystallization.

### 2.4.1.3 Additive approach in TOM core complex crystallization

The use of additives has played a significant role in membrane protein crystallization. Due to the high solvent content of protein crystals, soaking of chemicals into the crystal is possible. Soaking or co-crystallization with additives can improve the diffraction quality of crystals. Therefore, crystals were subjected to crystallisation additives such as detergents, heavy metal compounds and lipids. For instance, the use of amphiphilic additives (such as heptane-triol) has been shown to assist in the formation of highly ordered crystals<sup>256</sup>. Small amphiphiles can act to effectively reduce the size of the detergent region with subsequent improvement of crystal quality<sup>257,258</sup>. Heavy metal additives can deliver initial phases, form and stabilize crystal contacts, and have led to some spectacular improvements in resolution<sup>259,260</sup>.

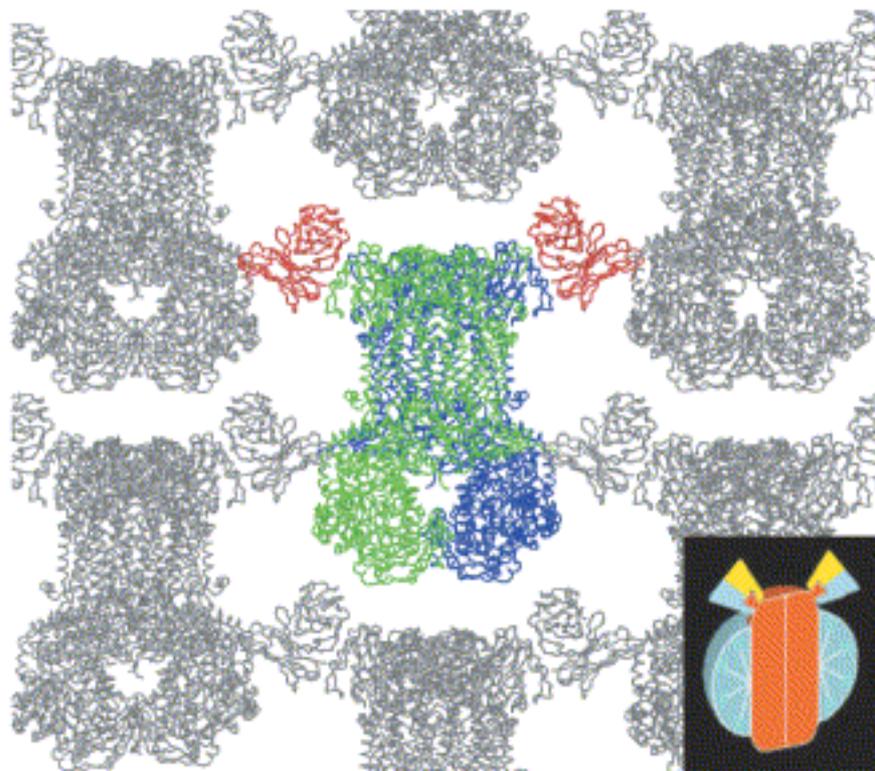
In case of TOM core complex mainly FOS-choline12 and  $\text{CuCl}_2$  showed an effect on the crystallization. Addition of the detergent FOS-choline12 shifted the crystallisation condition to higher PEG concentrations and slightly improved crystal quality, whereas addition of  $\text{CuCl}_2$ , which was reported to block the TOM channel<sup>221</sup>, increased the crystal size.

## 2.4.2 Antibody-fragment mediated crystallization of TOM core complex

Protein-detergent complexes form three-dimensional crystals in which contacts between adjacent protein molecules are made by the polar surfaces of the protein protruding from the detergent micelle<sup>257</sup>. Although attractive interactions between the micelles might stabilize the crystal packing<sup>261,262</sup>, these interactions do not lead to rigid crystal contacts. A strategy to increase the probability of getting well-ordered crystals is to attach polar domains to the membrane protein and thereby enlarge the polar surface.

Although the antibody-fragment mediated crystallization of membrane proteins represents a time-consuming and expensive approach, it is a valuable tool for their structural investigation. X-ray structures of several different membrane proteins in complex with antibody fragments have been published<sup>263-265</sup>. The binding of Fv or Fab fragments increases the hydrophilic part of integral membrane proteins and provides additional surface for crystal contacts. In addition, antigen-antibody complexes can be used to stabilize specific conformations of a protein to aid crystallization attempts. In all reported cases, antibody binding was either essential for the crystallization of the membrane protein or it substantially improved the diffraction quality of the crystals. Up to now all antibody fragments successfully used for co-crystallization were derived from hybridoma cell lines and these antibodies recognize native, nonlinear epitopes of their respective antigen<sup>236,266</sup>. Two prominent examples are the cytochrome *c* oxidase<sup>263</sup> and the cytochrome *bc*<sub>1</sub> complex<sup>264</sup> (depicted in Figure 2.19).

The availability of suitable amounts of highly purified TOM core complex is the basis for three-dimensional crystallization, but only poor-diffracting crystals were obtained. Disorder in the crystals may be caused by the protein existing in different conformations, substoichiometry of its subunits or flexibility. In addition, TOM core complex may exhibit only small hydrophilic domains which are essential for stable crystal contacts. Thus, enlarging the hydrophilic surface and stabilizing defined conformations by binding antibody fragments is a promising approach to obtain well-ordered crystals.



**Figure 2.19:** *Crystal structure of cytochrome  $bc_1$  complex. The two monomers of the central molecule of the crystal are coloured in blue and green, with bound antibody in red<sup>264</sup>.*

Native antibodies are not suitable for co-crystallization attempts. They possess flexible linker regions connecting the variable and constant domains<sup>267</sup>, and their bivalent binding mode is undesirable. Monovalent antibody fragments can be generated by proteolytic cleavage of the whole antibody, producing two Fab fragments per antibody molecule. These proteolytic fragments are easy to obtain, but care has to be taken to produce homogeneous fragment preparations suitable for crystallization. Residual parts of the flexible linker and, in some cases, glycosylation can hinder crystallization attempts. The crystallization and structure determination of the KcsA  $K^+$  channel provides the first example of an integral membrane protein co-crystallized with a proteolytically derived Fab fragment<sup>265</sup>. Recombinant antibody fragments are more versatile: they can be used in either Fab or Fv format. The different size might be critical for crystallization success, as the bound fragment should outreach the detergent micelle. Antibody-fragments suitable for co-crystallization attempts should bind the protein in its native conformation, have a high binding affinity and form a stable and rigid complex with their native antigen. The last implies the use of antibodies that bind to a discontinuous epitope and therefore, one critical aspect appears to be obtaining antibodies that are native-state specific, i.e., ELISA-positive and Western-negative. In addition, antibodies can help in trapping a flexible protein in a fixed conformation. A variety of structures have been solved in complex with Fv or Fab fragments<sup>264,265,268-270</sup>. Strikingly, in these cases most or all of the crystal contacts are between the antibodies.

In this study, TOM-directed monoclonal Fv fragments were manufactured by the hybridoma technique and their subsequent bicistronic expression in the periplasma of *E. coli*<sup>237</sup>. Although eight different IgG antibodies were obtained, which showed increased affinity in ELISA experiments, the Fv fragment P1C10 showed a much higher affinity upon characterization and was the only fragment which specifically bound to TOM complex. Addition of this antibody fragment to the purified TOM complex in the molar ratio of 5:1 has not been sufficient to saturate the complex, indicating the presence of more than one antibody binding sites. Typically, selection of specific antibodies starts from several thousands of hybridoma clones<sup>271</sup>. In case of TOM core complex, just a few hundred clones were found, which might be related to the fact that the fusions were frozen for several years before their cultivation took place. The high excess of antibodies from the IgM subclass may result from a biased immunization procedure. Although copurification and cocrystallization of the Fv fragment P1C10 with TOM core complex was successful, no alteration of the crystal quality was observed. Since only one suitable antibody fragment was found, this reduces the probability to achieve an antibody-fragment mediated improvement of the diffraction quality. There is no comprehensive overview of binding constants for all antibodies, which have been co-crystallized with their cognate antigen. Therefore it is not possible to specify yet which affinity threshold is required to successfully obtain well-diffracting crystals of antibody complexes, nor can one rule out the possibility that lower affinities may result in reduced chances of crystallization. In case of the Na<sup>+</sup>/H<sup>+</sup> antiporter NhaA all specific antibodies possess affinities in the nanomolar range, but extensive co-crystallization did not lead to the desired structure determination. This was achieved through the optimization of the purification strategy and other approaches<sup>272</sup>.

A promising alternative to raising monoclonal antibodies in mice is the construction of large combinatorial antibody libraries and the adaption of phage display antibody technology for *in vitro* selection. Another possibility in crystallization of TOM core complex might be to shift to other adaptable binding scaffolds such as the ankyrin repeat using ribosome display for selection<sup>141,273,274</sup>. For both techniques the selection and production of high-affinity binders are less expensive and time-consuming than for the standard hybridoma technology. The cytochrome *c* oxidase (COX) from *Paracoccus denitrificans* was the first successful example to demonstrate the feasibility of the antibody-fragment mediated crystallization approach<sup>263</sup>. In this case, the co-complex was purified by indirect immunoaffinity chromatography. This strategy could be also pursued for the purification of TOM core complex and might alter its crystallization state.

### 2.4.3 Expression of $\beta$ -barrel protein Tom40

Although the first membrane protein structure was solved in 1985<sup>275</sup>, the first structure of a recombinant membrane protein was not determined until 1998<sup>276</sup>. This fact speaks to the challenges of integral membrane protein overexpression. As a basic rule for any crystallization attempt, the protein of interest should be chemically and conformationally homogeneous and large amounts of the produced protein favour this homogeneity. Accordingly, in case of bacterial outer membrane proteins, the initial isolation procedures starting from the naturally expressed protein are often abandoned for methods that increased the yield by homologous overexpression into the outer membrane. For instance, sufficient pure material of bacterial TolC was obtained by functional expression in *E. coli* BL21 cells and the purified protein was crystallized in two and in three dimensions<sup>246,277</sup>. As an alternative, protein yields were also enhanced by non-functional overexpression. For bacterial outer membrane proteins, this expression into cytoplasmic inclusion bodies with subsequent solubilization in denaturing buffer and *in vitro* refolding was surprisingly successful. For example, the refolded outer membrane phospholipase A (OmpLA) yielded better crystals than the protein isolated from membranes<sup>278,279</sup>. In most cases of outer membrane protein expression into inclusion bodies, the T7-promoter system together with the strain *E. coli* BL21(DE3) was used<sup>241</sup>. In contrast to the quite detailed knowledge about the structural properties of bacterial outer membrane proteins<sup>280</sup>, no high-resolution structures of similar mitochondrial or plastid outer membrane proteins are known. The low-resolution projection of the human voltage-dependent anion-selective channel (VDAC) from two-dimensional crystals<sup>281,282</sup> and X-ray analysis of the human liver monoamine oxidase B (MAO-B)<sup>283</sup> are examples of rarely described structures of mitochondrial outer membrane proteins so far. However, MAO-B was classified as an anchored cytosolic rather than a membrane protein. Until now the only possibility to express eukaryotic organellar outer membrane protein in *E. coli* is using inclusion bodies. Recombinant expression and crystallization of the main component Tom40 of the TOM core complex may represent a convenient alternative to the crystallization of the whole complex. The yeast protein Tom40 was refolded by dilution and incorporated into liposomes<sup>220</sup>. As shown in this study, expression and refolding of the homologue from *N. crassa* was achieved. Since the first approx. 40 amino acid residues of ncTom40 exhibit no pronounced secondary structure and sequence conservation, a truncated version was used for this experiment to increase chance for crystallization. Further analytical measurement like CD spectroscopy should give more detailed information about the folding state of this construct and guide to initial crystallization attempts. The main problem with non-functional expression into inclusion bodies is the *in vitro* refolding of the inactive

polypeptide, an obstacle which is already surmounted. As Tom40 homologues are also present in thermophilic fungi as *Chaetomium thermophilum*, their recombinant expression, refolding and crystallization might increase the chances for structure determination.

## 2.5 Materials and Methods

### 2.5.1 Materials

#### 2.5.1.1 Oligonucleotides

The following oligonucleotides were used as primers to amplify different constructs from the cDNA of the hybridoma cells and to clone the Fv fragments into the pASK68 vector. Additionally, primers were used to amplify the Tom40 genes from *Neurospora crassa*, *Homo sapiens* and *Caenorhabditis elegans*. For the ligation of the amplified genes into the vectors pASK68, pRSET-PL0 and pRSET-PL1, all primers carried specific restriction sites. These sites were BamHI (forward and reverse) in case of the pRSET vectors and BamHI/EcoRI in case of pET-Duet<sup>TM</sup>-1 vector. Primers used for cloning into bicistronic pASK68 vector carried the restriction sites PstI (NsiI)/BstEII for the Fv heavy chain (V<sub>H</sub>) and SacI/XhoI for the Fv light chain (V<sub>L</sub>). Forward primers for cloning into pET-Duet<sup>TM</sup>-1 vector carry additionally a PreScission Protease cleavage site. For the amplification of genes from the cDNA of hybridoma clones a set of 10 different forward primer for V<sub>H</sub> and 11 forward primer for V<sub>L</sub> were used, whereas for the backward primers one primer was enough. All these primers were kindly provided by Carola Hunte (MPI of Biophysics, Frankfurt). All other primers were designed to have an overhang to ensure high cleavage efficiency and were synthesized and purified from Metabion international AG.

<u>Name</u>		<u>Enzyme</u>	<u>Organism</u>
<u>VH-P1C10</u>			
VH_P1C10_fo	5'-GAA GTT AAA <b>ATG CAT</b> CAG TCT GGA GCT GAG CTG-3'	NsiI	<i>Mus musculus</i>
VH_P1C10_re	5'-GGA GAC <b>GGT GAC</b> CAG AGT CCC TTG GCC CCA G-3'	BstEII	<i>Mus musculus</i>
<u>VL-P1C10</u>			
VL_P1C10_fo	5'-GAC ATC <b>GAG CTC</b> ACC CAG TCT CCA TCC TCC-3'	SacI	<i>Mus musculus</i>
VL_P1C10_re	5'- CCG TTT GAT <b>CTC GAG</b> CTT GGT GCC TCC ACC GAA C-3'	XhoI	<i>Mus musculus</i>

VH-P9E9

VH_P9E9_fo	5'- GAA GTT AAA <b>ATG CAT</b> CAG TCT GGG CCT GAG CTG -3'	NsiI	<i>Mus musculus</i>
VH_P9E9_re	5'- GGA GAC <b>GGT GAC</b> CGT GGT GCC TTG GCC CCA G -3'	BstEII	<i>Mus musculus</i>

VL-P9E9

VL_P9E9_fo	5'- GAC ATC <b>GAG CTC</b> ACC CAG TCT CCA GCT TCT TTG -3'	SacI	<i>Mus musculus</i>
VL_P9E9_re	5'- CCG TTT GAT <b>CTC GAG</b> CTT GGT GCC TCC ACC GAA C -3'	XhoI	<i>Mus musculus</i>

VH-P7C1

VH_P7C1_fo	5'- GAA GTT AAA <b>ATG CAT</b> CAG TCT GGG ACT GTG CTG G-3'	NsiI	<i>Mus musculus</i>
VH_P7C1_re	5'- GGA GAC <b>GGT GAC</b> CGT GGT GCC TTG GCC CCA ATA G-3'	BstEII	<i>Mus musculus</i>

Tom40

hsTom40_77_BamHI_fo	5'-AAA AAA <b>GGA TCC</b> CTG CCC AAC CCG GGC AC-3'	BamHI	<i>Homo sapiens</i>
hsTom40_361_BamHI_re	5'-AAA AAA <b>GGA TCC</b> TTA GCC GAT GGT GAG GCC AAA G-3'	BamHI	<i>Homo sapiens</i>
hsTom40_Duet_77_BamHI_fo	5'-AAA AAA <b>GGA TCC</b> GCT GGA AGT BamHI TCT GTT CCA GCT GCC CAA CCC `' GGG CAC-3	BamHI	<i>Homo sapiens</i>
hsTom40_Duet_361_EcoRI_re	5'-AAA AAA <b>GAA TTC</b> TTA GCC GAT EcoRI GGT GAG GCC AAA G-3'	EcoRI	<i>Homo sapiens</i>
ceTom40_17_BamHI_fo	5'-AAA AAA <b>GGA TCC</b> AAT CCG GGT BamHI TCA TAC GAA GAA C-3'	BamHI	<i>C. elegans.</i>
ceTom40_301_BamHI_re	5'-AAA AAA <b>GGA TCC</b> TTA ACC AAT BamHI GAT AAG TCC GAT TCC-3'	BamHI	<i>C. elegans.</i>
ceTom40_Duet_17_BamHI_fo	5'-AAA AAA <b>GGA TCC</b> GCT GGA AGT BamHI TCT GTT CCA GAA TCC GGG TTC ATA CGA AGA AC-3'	BamHI	<i>C. elegans.</i>
ceTom40_Duet_301_EcoRI_re	5'-AAA AAA <b>GAA TTC</b> TTA ACC AAT EcoRI GAT AAG TCC GAT TCC	EcoRI	<i>C. elegans.</i>
ncTom40_38_BamHI_fo	5'-AAA AAA <b>GGA TCC</b> CTT TCC AAC BamHI CCC GGC ACG-3'	BamHI	<i>N. crassa</i>
ncTom40_329_BamHI_re	5'-AAA AAA <b>GGA TCC</b> TTA CTC AAT BamHI GGA GAC GGA CAT GC-3'	BamHI	<i>N. crassa</i>

## 2.5.1.2 Vectors

### 2.5.1.2.1 pRSET-PL0 and pRSET-PL1

The expression vectors pRSET-PL0 and PL1 were derived from the pRSET vector (Invitrogen). pRSET-PL1 was constructed by cutting out the MCS with restriction enzymes NdeI and BamHI and insertion of the DNA encoding a PreScission Protease cleavage site and a His-tag. Tom40 constructs were ligated with BamHI restriction sites into the pRSET-PL0 and PL1 vectors, which carry a T7 promoter, ampicillin resistance ( $Amp^r$ ) and a lac operator, allowing gene expression to be induced by IPTG (isopropyl-b-D-thiogalactopyranosid). A protein expressed with the pRSET-PL1 vector could be purified by affinity chromatography via his-tag, which can be cleaved off with PreScission Protease.

### 2.5.1.2.2 pET-Duet<sup>TM</sup>-1

The pET-Duet<sup>TM</sup>-1 vector (Novagen) was applied for expression of hsTom40 and ceTom40 variants, which did not express in the pRSET system. The Duet vectors are based on T7 promoter expression and each designed to coexpress two target proteins in *E. coli*, but can also be used for expression of only one gene, which is then inserted into the first open reading frame. The vector encodes two multiple cloning sites (MCS), each of which is preceded by a T7 promoter, lac operator and ribosome binding site (rbs). The vector also carries the pBR322-derived ColE1 replicon, lacI gene and ampicillin resistance gene as well as the N-terminal fusions to his-tag. A PreScission Protease cleavage site was introduced by the forward primer.

### 2.5.1.2.3 pCR2.1-TOPO and pDrive

The PCR products from the amplified genes of the Fv fragments, which sequences were unknown, were cloned directly from the PCR reaction into the pCR2.1-TOPO (Invitrogen) or the pDrive cloning (Qiagen) vector. Both vectors carried ampicillin resistance for selection. The cloning was performed according to the instructions of the protocols of the TOPO and the pDrive cloning kits as well as the Qiagen A-Addition kit. The insertions could then be sequenced and were used as templates for the PCR using specific primers with restriction sites, followed by cloning into the pASK68 vector (Section 2.5.1.2.4).

### 2.5.1.2.4 pASK68

The pASK68 vector is designed for periplasmic production of Fv fragments and was kindly provided by Carola Hunte (MPI of Biophysics, Frankfurt). It represents a bicistronic system, where heavy ( $V_H$ ) and light chain ( $V_L$ ) of the Fv fragment are inserted into the open reading frames one and two. The vector provides  $V_H$  with a C-terminal strep-tag I and the  $V_L$  with a C-terminal myc-tag. Both chains are fused to the OmpA and PhoA signals, respectively, for periplasmic targeting, ensuring proper folding of the disulfide bonds of the antibody fragment. The  $V_H$  and  $V_L$  genes were inserted with the restriction enzymes PstI (NsiI)/BstEIII and SacI/XhoI. The vector carries ampicillin resistance ( $Amp^r$ ) for selection.

### 2.5.1.3 Bacteria

The following *E. coli* strains were used for molecular biological working and for the expression. The strains were ordered from Stratagene, Heidelberg or provided by Carola Hunte (in case of JM 83, from MPI of Biophysics, Frankfurt).

Strain	Genotype	Application
XL-1 blue strain a	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac{F' proAB lacI <sub>q</sub> ZM15Tn10(Tet <sup>r</sup> )} non inducible promoter	Preparation of high-quality plasmid DNA.
BL21-CodonPlus(DE3)-RIL	B F-ompT hsdS(rB- mB-) dcm+ Tet <sup>r</sup> galλ (DE3) endA Hte {argU ileY leuW Cam <sup>r</sup> ), extra copies of argU, ileY and leuW tRNA genes on a ColE1- compatible plasmid with canamycin resistance marker, protease deficiency, chromosomal T7-polymerase gene	Induced high-yield protein expression with, especially with T7 promoter.
JM83	rpsL ara Δ(lac-proAB) Φ80dlacZΔM15, streptomycin resistance	Periplasmic expression of Fv fragments

**Table 2.4:** Different *E. coli* strains which were used in this study.

### 2.5.1.4 Enzymes

Name	Company
Taq Polymerase	New England Biolabs, Schwalbach
Pfu Polymerase	Promega
Shrimp Alkaline Phosphatase	Roche
T4 DNA Ligase	New England Biolabs, Schwalbach
Restriction Enzymes (BamHI, EcoRI.....)	New England Biolabs, Schwalbach

**Table 2.5:** List of applied enzymes.

### 2.5.1.5 Equipment and chemicals

ÄKTA Explorer, ÄKTA Purifier and ÄKTA Prime Chromatography Systems	Amersham Biosciences, Freiburg
Amicon Ultra 10K, 30K device	Millipore, Bedford, MA, USA
Autoklav Bioclav	Schütt
Autoklav Varioklav	H+P Labortechnik GmbH
Balance MC1 AC210P	Sartorius, Göttingen
Balance CPG22	Sartorius, Göttingen
Branson Sonifier 250	G. Heinemann
Centrifuges: Avanti J-25, Avanti J-20 XP	Beckman Instruments
Centrifuge 5810 R	Eppendorf
22 mm x 0.96 mm thick Plain circle cover slides	Hampton R., Laguna Niguel, CA,US
22 mm Circular Cover Slides (siliconized)	Jena Bioscience
Cups (1.5 ml and 2.0 ml)	Sarstedt
ThermoTube™ PCR Tubes 0.2 ml	Peqlab Biotechnology GmbH
Digital Camera	Sony
Econo-Column Chromatography Column	Biorad
Electrophoresis Power Supply	Consort
Electrophoresis Power Supply EPS 601	Amersham Biosciences
Freezer –20 °C	Liebherr
Freezer –80 °C	Forma Scientifie, Life Sciences International GmbH
Fridge 4 °C	Liebherr
Fraction collector Frac 100	Amersham Pharmacia, Freiburg
GDS-7900 System Transilluminator	UVP

(Agarose gel documentation)

Gel chambers	workshop, Institut für Physiologische Chemie, München
Gel Drying Film	Promega Corp., Madison, WI, USA
Glas ware	Schott, Duran
Grease	Hampton R., Laguna Niguel, CA, USA
Greiner 96well plates	Greiner und Söhne, Frickenhausen
Incubators for molecular biology	Heraeus Christ
Incubator for crystallisation	RUMED
Microscope	Carl Zeiss AG, Oberkochen, Germany
Mikro 22 R centrifuge	Hettich
Mikrowave	Bosch
Mono Q HR 10/10	Amersham Pharmacia, Freiburg
Multichannel Pipette Research Pro 10 and 300	Eppendorf
Multitron II shaker	Infors AG, Bottmingen
Nanosep Centrifugal Devices (3 and 10K)	Pall Life Sciences (New York, USA)
NMR spectrometer	Varian
Overhead shaker	workshop, Institut für Physiologische Chemie, München
Mastercycler gradient (PCR)	Eppendorf
Petriplates	Greiner und Söhne, Frickenhausen
Peristaltik-Pumpe P-1	Amersham Biosciences
pH-Meter CG 822	Schott
Photometer Gene Quant pro	Amersham Biosciences
Photometer Ultrospec 10	Amersham Biosciences
Pipettes 2, 20, 100, 200, 1000 µl	Gisco
Primus (PCR)	MWG Biotech
PP-Test tubes, 15 and 50 ml (falcons)	Greiner
Quarz precision cuvettes 105.201-QS	Hellma Müllheim
Röntgen films NewRX	Fuji
Sterile benches	BDK, Genkingen
Superdex 200 HR 10/30	Amersham Pharmacia, Freiburg
Table centrifuge 5415 D	Eppendorf-Netheler, Hamburg
Thermomixer compact	Eppendorf
VDX™ 24well plate (greased)	Hampton R., Laguna Niguel, CA, US
Vortex Mixer	Bender&Hobein, Zürich

***Unless stated otherwise all chemicals were from the following companies in p.A. quality.***

- Amersham Biosciences, Freiburg
- Fluka, Neu-Ulm
- Fermentas, St.Leon-Rot
- Merck, Darmstadt
- New England Biolabs, Schwalbach/Taunus
- Qiagen, Hilden
- Promega
- Roth, Karlsruhe
- Serva, Heidelberg
- Sigma Aldrich, Deisenhofen

### 2.5.1.6 Media and Buffers

All solutions and medium were prepared using bidestillized water (ddH<sub>2</sub>O, Millipore) and were autoclaved or sterile filtered if required. All media and buffers are listed in Table 2.6 except cell culture media which is shown below.

NAME	PRESCRIPTION
<b><i>E. coli</i> Medium</b>	
LB	10 g Trypton, 5 g Yeast Extract, 10 g NaCl, ad 1 l ddH <sub>2</sub> O (for culture plates ad 1.5% w/v Agar)
<b>Competent Cells</b>	
electrocompetent	10% (v/v) Glycerol
chemical competent	0.1 M CaCl <sub>2</sub>
<b>Isolation of plasmid DNA</b>	
Resuspension buffer	50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNase
Lysis buffer	200 mM NaOH, 1.0% (w/v) SDS
Neutralization buffer	3 M potassium acetate, pH adjusted to 5.5 with acetic acid
<b>Agarose gels</b>	
5xSample Buffer	50 mM EDTA, 0.05% (w/v) Bromphenolblue, 0.05% (w/v) Xylencyanol, 30% (v/v) Glycerol
1xTAE - Electrophoresis-Buffer	40 mM Tris/Acetate, 20 mM Natriumacetate, 1 mM EDTA, pH 7.5
<b>Antibiotics (1000xstocks)</b>	
Ampicillin	100 mg/ml H <sub>2</sub> O

Chloramphenicol	30 mg/ml Ethanol
Tetracyclin	12.5 mg/ml Ethanol
Streptomycin	30 mg/ml H <sub>2</sub> O
<b>ELISA</b>	
Coating buffer	0.2 M Na <sub>2</sub> CO <sub>3</sub> , 1.243% (w/v) NaHCO <sub>3</sub> , 0.551% (w/v) Na <sub>2</sub> CO <sub>3</sub> or 1.487% (w/v) Na <sub>2</sub> CO <sub>3</sub> * 10 H <sub>2</sub> O, pH without titration ~ 9.55
Wash buffer	0.9 % (w/v) NaCl, 0.1% (v/v) Tween-20 (NaCl can be prepared as 10x stock, Tween 20 was added fresh to 1x)
Blocking buffer	50 mM Tris-Cl pH 7.5, 150 mM NaCl (can be prepared as 10 X stock), 2 % BSA
TOM buffer	50 mM KAc/MOPS pH 7.2, 0.05 % (w/v) DDM
Antibody solution	50 mM KAc, 0.05% (v/v) Tween-20, 0.5 % (w/v) BSA
Development solution	10% (v/v) Diethanolamine pH 9.8, 0.5 mM MgCl <sub>2</sub> (2.5 µl from 2 M stock for 10 ml) 1 mg/ml p-NPP (p-nitrophenylphosphate, 10 mg/ml stock in aliquots at -20 °C)
<b>Protein A affinity purification</b>	
Binding buffer	100 mM Tris-Cl pH 9.0
Elution buffer	100 mM sodium citrate pH 3.0
<b>Gene expression</b>	
Lysis buffer	20 mM Tris-Cl pH 8.0, 100 mM NaCl
<b>Periplasmic extraction</b>	
Extraction buffer	100 mM Tris-Cl pH 8.0, 500 mM sucrose, 1 mM EDTA
<b>Isolation of inclusion bodies</b>	
Washing buffer 1	20 mM Tris-Cl pH 8.0, 200 mM NaCl, 0.5% (w/v) LDAO
Washing buffer 2	20 mM Tris-Cl pH 8.0, 200 mM NaCl
<b>Protein refolding</b>	
Urea buffer	20 mM Tris-Cl pH 8.0, 100 mM NaCl, 8 M urea
Dilution buffer	20 mM Tris-Cl pH 8.0, 100 mM NaCl, 0.03% (w/v) DDM, 0.2 M L-arginine
<b>Streptavidin affinity chromatography</b>	
Loading buffer	100 mM Tris-Cl pH 8.0, 150 mM NaCl, 1 mM EDTA
Elution buffer	100 mM Tris-Cl pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin
<b>Purification of TOM core complex</b>	
Solubilization buffer	20 mM Tris-Cl pH 8.5, 300 mM NaCl, 20% (v/v) glycerol, 20 mM imidazole, 1% (w/v) <i>n</i> -dodecyl-β-D-maltoside (β-DDM; Glycon Biochemicals), 1 mM PMSF (freshly added)
Equilibration buffer	50 mM Tris-Cl pH 8.5, 20% (v/v) glycerol, 0.2% (w/v) β-DDM, 1mM PMSF
TOM washing buffer 1	20 mM Tris-Cl pH 8.5, 300 mM NaCl, 10% (v/v) glycerol, 20 mM imidazole, 0.1% (w/v) β-DDM, 1 mM PMSF
TOM washing buffer 2	50 mM KAc/MOPS pH 7.2, 10% (v/v) glycerol, 40 mM imidazole, 0.1% (w/v) β-DDM, 1 mM PMSF
TOM elution buffer	50 mM KAc/MOPS pH 7.2, 10% (v/v) glycerol, 300 mM imidazole, 0.1% (w/v) β-DDM, 1 mM PMSF
ResQ low salt buffer	50 mM KAc/MOPS pH 7.2, 10% (v/v) glycerol, 0.03% (w/v) β-DDM

ResQ high salt buffer	50 mM KAc/MOPS pH 7.2, 10% (v/v) glycerol, 7.457% (w/v) KCl, 0.03% (w/v) $\beta$ -DDM
Gel filtration buffer	10 mM KAc/MOPS pH 7.2, 0.03% (w/v) $\beta$ -DDM
<b>SDS-PAGE</b>	
4xSample buffer	40% (v/v) glycerol, 240 mM Tris-Cl pH 6.8, 0.02% (w/v) Bromphenolblue, 20% (w/v) $\beta$ -mercaptoethanol, 8% (w/v) SDS
10xSDS runningbuffer	248 mM Tris-Cl, 1.9 M glycine, 10% (w/v) SDS
5xUrea-SDS running buffer	3.26% (w/v) Tris-Cl, 7.214% (w/v) glycine, 0.5% (w/v) SDS
<b>Coomassie staining of Protein gels</b>	
Coomassie stain	40% (v/v) ethanol, 10% (v/v) acetic acid, 0.2% (w/v) Coomassie Brilliant Blue R-250
Coomassie destain	30% (v/v) methanol, 10% (v/v) acetic acid
<b>Silver staining of SDS gels</b>	
TCA solution	30 g trichloroacetic acid (TCA), 125 ml methanol, ddH <sub>2</sub> O ad 250 ml
Fixation solution	30% (v/v) ethanol, 10% (v/v) acetic acid
Reducing solution	0.1 % (w/v) sodium thiosulfate, 2% (v/v) glutaraldehyde
Silver solution	1.89 ml NaOH (1 M), 1.4 ml NH <sub>3</sub> (25%), ddH <sub>2</sub> O ad 96 ml, drop wise addition of AgNO <sub>3</sub> (20% (w/v) solution) under stirring till saturation is reached.
Development solution	36.5 $\mu$ l citric acid (34% (w/v)), 150 $\mu$ l formaldehyde, ddH <sub>2</sub> O ad 250 ml
<b>Western Blot</b>	
Transfer buffer	20 mM Tris-Cl, 150 mM glycine, 20% (v/v) methanol, 0.08% (w/v) SDS
Ponceau S solution	0.2% (w/v) Ponceau S in 3% (w/v) trichloroacetic acid
<b>Immunodecoration</b>	
TBS	10 mM Tris-Cl pH 7.5, 135 mM NaCl, blocking with either 3% (w/v) BSA or 5% (w/v) milk powder
ECL reagent	0.1 M Tris-Cl pH 8.5, 2.5 mM 3-aminophthalhydrazide, 0.4 mM p-cumaric acid, 0.0113% (v/v) H <sub>2</sub> O <sub>2</sub>
10xNET/gelatine	500 mM Tris-Cl pH 7.5, 1.5 M NaCl, 0.05 M EDTA, 0.5% (v/v) Triton-X-100, 8.3% (w/v) gelatine (Serva, heated in buffer)
<b>TCA precipitation</b>	
Precipitant	72% trichloroacetic acid
TCA wash buffer	100 mM Tris-Cl pH 8.8, 70% (v/v) Ethanol, small amount (spatula tip) Bromphenolblue
<b>TOM core complex crystallisation</b>	
TOM core crystallisation condition	100 mM ammonium citrate, 50 mM Tris-Cl pH 7.2, 20-30% (v/v) PEG 400
TOM:P1C10 crystallisation condition	0.2 M calcium chloride, 0.1M HEPES pH 7.5, 25-28% (v/v) PEG 400

**Table 2.6:** Media and buffers except for cell culture.

**Media of hybridoma cell culture:**

Order numbers if not otherwise noted are from SIGMA.

RPMI 1640		500 ml (supplements added to 500 ml medium)	<i>R-0883</i>
FCS	10/20%	50 ml/100 ml	
Glutamin	2 mM	5 ml (200 mM)	<i>G-7513 or G-2150/Hybrimax</i>
Hepes	10 mM	5 ml (1 M)	<i>H-0887</i>
Na-Pyruvat	1 mM	5 ml (50 mM)	<i>I1360039 Boehringer</i>
2-ME	50 µM	0.5 ml (50mM)	<i>M-7522</i>
IL-6	50 U/ml	625 µl (40000) U/ml	<i>1444581</i>

**Optional supplements:**

OPI	1 X	5 ml 100 X	<i>O-5003</i>
Haz	1 X	10 ml (1 X is 100 µM Hypoxanthin and 1 µg/ml Azaserin)	<i>A-9666</i>
Hypoxanthin	2 X	5 ml 100 X	<i>I1364-015 GIBCO</i>

**Stock solutions:**

Glutamin	200 mM stock, store in 5 ml aliquots, -20°C
Hepes	1 M, 4°C
Na-Pyruvat	100 mM, 4°C
OPI (Sigma O-5003)	1 vial + 10 ml cell culture water (100 X), store at -20°C
Haz	1 vial + 10 ml cell culture water (50 X), store at -20°C
2-ME (78.13 MW, d=1.114 g/ml, p.a.)	88 µl to 25 ml with cell culture water (50 mM), store 4°C
IL-6 (Interleukin-6, mouse, <i>Roche</i> )	1 ml (200000 U/ml) + 4 ml PBS (cell culture grade), store in aliquots at -20°C
6- Thioguanin is 2-Amino-6-Mercaptopurine ( <i>Sigma A-4660</i> )	2.5 mg + 10 ml cell culture water ( <i>W-3500</i> ), 50 X, store -20°C

**Supplements for freezing**

DMSO (*Sigma D-8418*) 10 % in basis medium, precooled medium before addition to the cells

**Alternative medium:**

**Pan401 Medium** Panserin401 (*PanBiotech, P04-710401*), supplemented Iscove's medium  
 Glutamin  
 IL-6

Can be used for fusion with 10 % FCS plus 10 % HES (*Sigma H-8142*, enhancing supplement). Cells can be quickly adapted to 2 FCS or even FCS free in this medium

**Selection Medium A**

RPMI 1640		500 ml (supplements are added to 500 ml medium)
FCS	20 %	100 ml (Myocclone)
Glutamin	2 mM	5 ml
Hepes	10 mM	5 ml
Na-Pyruvat	1 mM	5 ml (100 mM)
2-ME	50 µM	0.5 ml (50 mM)
IL-6	50 U/ml	625 µl (40000 U/ml)
Hypoxanthin	100 µM	10 ml 50 X Haz
Azaserin	1 µg/ml	
1 X OPI		5 ml 100 X OPI

*no feeder cells, use conditioned medium*

*as before but 10 % FCS 10% Hybridoma enhancing (Sigma H-8142)*

*maybe 2 X OPI*

**Selection Medium B**

<b>Panserin401</b>		500 ml (supplements are added to 500 ml)
FCS	10 %	50 ml
Glutamin	2 mM	5 ml
IL-6	50 U/ml	625 µl (40000 U/ml)
Hypoxanthin	100 µM	10 ml 50 X Haz
Azaserin	1 µg/ml	
1 X OPI		5 ml 100 X OPI

10% hybridoma enhancing supplement (HES, Sigma H-8142)

*no feeder cells, use conditioned medium*

**2.5.1.7 DNA/RNA Isolation and Preparation Kits**

<b>Name</b>	<b>Application</b>
Pure Yield™ Plasmid MidiPrep System	Isolation of plasmid DNA from <i>E. coli</i> .
Wizard® SV Gel and PCR Clean Up System	Purification of PCR-products and extraction of DNA from agarose gel

Company: Promega Corp., Madison, WI, USA

RNeasy Kit

Isolation of mRNA from tissue culture

Company: QIAGEN GmbH, Hilden, Germany

### 2.5.1.8 Columns

NAME	VOLUME	V <sub>0</sub> = void volume	LOADING RANGE	MMW-Range	TECHNIQUE
HiTrap Protein A HP	1 ml		~20 mg		Affinity chromatography
Resource Q	6 ml		~25 mg		Anion exchange
Superose 6 10/300 GL	24 ml	7.77 ml		5 kDa - 5 MDa	Size exclusion
Superose 6 PC 3.2/30	2.4 ml	0.796 ml		5 kDa - 5 MDa	Size exclusion

**Table 2.7:** *Columns.*

### 2.5.1.9 Bioinformatics

DNA and protein sequences were obtained from the „National centre for Biotechnology Information” (<http://www.ncbi.nlm.nih.gov>) and SWISS-PROT (<http://www.expasy.ch>). Primary sequence predictions were obtained from PredictProtein. PredictProtein is an Internet service (accessible via SWISS-PROT portal) for sequence analysis and the prediction of protein structure and function. Users submit protein sequences and PredictProtein returns multiple sequence alignments, PROSITE sequence motifs, low-complexity regions (SEG), nuclear localisation signals, regions lacking regular structure (NORS) and predictions of secondary structure, solvent accessibility, globular regions, transmembrane helices, coiled-coil regions, structural switch regions, disulfide-bonds (DISULFIND), sub-cellular localization, and functional annotations. The secondary structure prediction was made by PROFsec, an advanced method of PHDsec.

The computation of various physical and chemical parameters for a given protein by entering its sequence was performed by the ProtParam tool (SWISS-PROT portal). The computed parameters include the molecular mass, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY).

## 2.5.2 Methods

### 2.5.2.1 Primer design

The length of the primers varied from 29 to 53 bases and the primers had at least one G or C at the 3'-end to obtain better binding and elongation. The annealing parts were designed to have a melting temperature of approximately 60°C to allow sufficiently high annealing temperatures.

In case of the pRSET-PL0, -PL1, pET-Duet<sup>TM</sup>-1 and pASK68, no start codon had to be introduced because it is already included in the vector. The forward primers for cloning into the pET-Duet<sup>TM</sup>-1, additional bases were introduced to shift the BamHI restriction site into frame. For all constructs the reverse primers had to contain a stop codon (TGA) before the restriction site (see Section 2.5.1.1).

### 2.5.2.2 Methods in molecular biology

#### 2.5.2.2.1 Polymerase chain reaction (PCR)

Since the beginnings of the PCR new applications were constantly found for this method, which allows amplifying DNA sequences of any sequence. For the PCR two oligonucleotide primers complementary to sequences at the ends of the required DNA, a mixture of all four dNTP's and a for high temperature stable polymerase with the suitable buffer and the right concentrations of salt are required. After the denaturizing of the dsDNA (denaturizing), the primers hybridize with the complementary DNA (annealing) and the polymerase elongates the ssDNA construct starting with the 3'-end of the primer (elongation). Thus, the number of dsDNA constructs increases exponentially with each cycle. Either the enzymes Taq or Pfu were used as polymerases.

Reaction mixture:

	Substance	Volume
1)	DNA template (approx. 0.5 µg/µl)	1 µl
2)	Forward primer (100 pmol/µl)	0.5 µl
3)	Reverse primer (100 pmol/µl)	0.5 µl
4)	dNTP's (5 mM per nucleotide)	2 µl
5)	Taq or Pfu buffer (10x)	5 µl

6)	H <sub>2</sub> O steril	40 µl
7)	Taq (5 U/µl) or Pfu polymerase (3 U/µl)	1 µl

The reaction mixture was pipetted into 500 µl PCR tubes and PCR reaction was carried out in a thermocycler.

PCR program:

	Temperature	Time
1)	95 °C	10 min
2) a)	92 °C	1 min
b)	58 °C	1 -2 min
c)	72 °C (Taq)/ 75 °C (Pfu)	1 min
Step 2 was repeated thirty times.		
3)	72 °C (Taq)/ 75 °C (Pfu)	10 min

Annealing temperature (2b) and elongation time (2c) of the reaction can be varied and had to be adjusted to the respective reaction. PCR products were analyzed by agarose gel electrophoresis and purified using the Wizard<sup>®</sup> SV Gel and PCR Clean Up System. To check for positive clones by PCR on *E. coli* colonies, colonies were picked with a sterile tip and resuspended in 10 µl of sterile ddH<sub>2</sub>O. The reaction volume was then only half of the above mixture, Taq was used as polymerase and the resuspended colony as template. The forward primer of the insert and a reverse primer annealing in the vector served as primers. For all PCR reactions concerning Tom40 constructs, commercially available cDNAs of the organisms were used as templates, respectively, whereas for cloning of the Fv fragments, cDNA of the hybridoma clones served as template.

#### 2.5.2.2.2 RNA isolation from hybridoma cell culture and Reverse Transcription

Isolated cells from positive hybridoma clones were harvested and mRNA was isolated using the RNeasy kit (Section 2.5.1.7). Reverse transcription of the mRNA was performed with the Omniscript RT Kit (Qiagen), following the instructions of the protocol. The resulting cDNA served as template for the cloning of the Fv fragments into TOPO or pDrive vectors (Section 2.5.1.2.3).

### 2.5.2.2.3 Digestion of DNA with restriction endonucleases

For analytical and preparative purposes plasmid DNA and PCR products were digested with particular restriction endonucleases, which digest the DNA at specific, mostly palindromal sequences. All used enzymes produced an overhang, thus only ligation of sticky ends occurred. The incubation time, temperature and the buffer used in different reactions were applied according to the manufacturer's recommendations. The obtained fragments were analyzed by agarose gel electrophoresis. Before the final sequencing of the resulting plasmid DNA, an additional restriction digestion was carried out to see whether the clone contained the required insert.

PCR product	Vector	Control Restriction
5 µl PCR construct	3 µg vector	2 µl vector
5 µl BSA (10x)	5 µl BSA (10x)	2 µl BSA (10x)
0.5 µl Enzyme 1	0.5 µl Enzyme 1	0.5 µl Enzyme 1
0.5 µl Enzyme 2	0.5 µl Enzyme 2	0.5 µl Enzyme 2
5 µl 10xBuffer	5 µl 10xBuffer	2 µl 10xBuffer
ad 50µl sterile ddH <sub>2</sub> O	ad 50µl sterile ddH <sub>2</sub> O	13 µl sterile ddH <sub>2</sub> O

**Table 2.8:** Restriction digestion of PCR product or vector.

All constructs for insertion into the pRSET-PL0 and -PL1 vectors were digested with BamHI, whereas all constructs for ligation into pET-Duet<sup>TM</sup>-1 vector were digested with BamHI/EcoRI. Constructs for ORF1 and ORF2 of the pASK68 vector have to be digested with PstI/BstEII and SacI/XhoI, respectively. Inserts that contained an internal PstI restriction site, were cloned with NsiI (both endonucleases produce the same overhang). Analysis of the desired DNA fragments and cleaved vectors was done by gel electrophoresis (Section 2.5.2.2.5). Self ligation of the linearized plasmid was suppressed by dephosphorylating the 5'-end using shrimp alkaline phosphatase (SAP). Therefore 1 µl of SAP (10 U/µl) and the appropriate amount of 10 fold SAP buffer were added to the purified, digested vector and incubated for one hour at 37°C. The reaction was stopped through gel electrophoresis or purification with Wizard<sup>®</sup> SV Gel and PCR Clean Up System. This reaction was not used in all cases because the dephosphorylation could reduce the number of positive clones after transformation.

### 2.5.2.2.4 Ligation

If PCR product and vector are restricted with the same enzymes the desired DNA construct can be inserted into the open plasmid and ligated with T4 DNA ligase. Linearized DNA vector and an approximately 10 fold molar excess of DNA fragment to be inserted were incubated with T4 DNA ligase and buffer. The ligation mixture was prepared on ice and incubation was performed either at RT for 1 h or at 4°C over night. The self ligation of the vector served as negative control. 0.5 to 1 µl of this mixture were used for transformation into *E. coli* cells.

Ligation		Negative-Control	
1 µl	vector	1 µl	vector
10 x molar excess	PCR-construct	16 µl	H <sub>2</sub> O sterile
2 µl	10xLigase-buffer	2 µl	10xLigase-buffer
1 µl	T4-DNA-Ligase	1 µl	T4-DNA-Ligase

**Table 2.9:** Ligation mixture.

### 2.5.2.2.5 DNA agarose gel electrophoresis

The agarose gel electrophoresis was used to separate and analyze DNA fragments between 0.1 and 8 kb. Voltages were in the range of 80 to 120 V and the concentration of agarose was 1%. Detection was done by the DNA intercalating dye ethidium bromide which can be observed in the ultraviolet. For preparation of the gel, agarose was solved under heat in 1xTAE-buffer and supplemented with 1:10000 fold dilution of ethidium bromide (Sigma). The rigid gel was placed in a gel chamber and covered with TAE-buffer. Samples were mixed with 5 x sample buffer and loaded onto the gel. For the marker a 1kb or a 100bp DNA ladder (New England Biolabs) was used. DNA constructs and linearized vectors that were utilized for ligation were excised from the gel under UV light ( $\lambda = 254$  nm) and purified with the Wizard<sup>®</sup> SV Gel and PCR Clean Up System.

### 2.5.2.2.6 Competent *E. coli* cells

#### 2.5.2.2.6.1 Chemical competent *E. coli* cells

500 ml LB medium supplemented with tetracycline (XL1blue), chloramphenicol (BL21-(DE3)-RIL) or streptomycin (JM83) were inoculated with 5 ml of over-night culture of the desired *E. coli* strain. The cells were grown at 37°C to OD<sub>600</sub> of 0.4. After incubation on ice for 15 minutes cells were centrifuged for 15 minutes at 2580 x g (4000 rpm, JA10 rotor) and 4°C. All following steps were done at ≤4°C. The pellet was resuspended in 40 ml 0.1 M CaCl<sub>2</sub>, incubated on ice for 30 min and harvested by centrifugation at 2580 x g and 4°C for 15 min. This pellet was resuspended in 20 ml 0.1 M CaCl<sub>2</sub> supplemented with 4 ml of glycerol. After an additional incubation on ice for 2 h, the suspension was dispensed in 200 µl aliquots and after shock-freezing in liquid nitrogen stored at -80°C.

#### 2.5.2.2.6.2 Electro-competent *E. coli* cells

500 ml LB medium supplemented with tetracycline (XL1blue), chloramphenicol (BL21-(DE3)-RIL) or streptomycin (JM83) were inoculated with 5 ml of over-night culture of the desired *E. coli* strain. The cells were grown at 37°C to OD<sub>600</sub> of 0.5 and then harvested by centrifugation for 15 minutes at 2580 x g (4000 rpm, JA10 rotor) and 4°C. All following steps were done at ≤4°C. The pellet was washed three times with 500, 250 and 50 ml of 10% (v/v) glycerol. After the washing steps, cell pellet was resuspended in 500 µl of 10% (v/v) glycerol, dispensed in 40 µl aliquots and after shock-freezing in liquid nitrogen stored at -80°C.

#### 2.5.2.2.7 Transformation

Transformation of plasmid DNA into chemical competent *E. coli* cells was mostly done with the heat shock method. Therefore 100 µl of the competent cells were thawed on ice, mixed with 4-10 µl of the ligation suspension (or 1 µl of purified plasmid) and incubated on ice for 10-30 min. Afterwards cells were heated at 42°C for 1 min. After a short recovery on ice for 2 min, 500 µl of fresh LB medium were added and the mixture was shaken at 37°C for 45 min. Finally, the bacteria were harvested by centrifugation, resuspended in 200 µl of fresh LB medium and plated on a selective medium (LB/Amp<sup>r</sup> for all used vectors). The plates were incubated at 37°C over night.

For the transformation of plasmid DNA into electrocompetent cells 0.5-1  $\mu$ l of ligation suspension (or purified plasmid) were added to 40  $\mu$ l of thawed competent cells. The suspension was transferred to a precooled cuvette and the cuvette was introduced in an electroporation apparatus (settings: 2.0 kV, 400  $\Omega$ , 25  $\mu$ F, time constant 8-9 ms). After a brief application of a high electric voltage, the suspension was diluted with 800  $\mu$ l LB-medium, and shaken for 30-60 min at 37°C. The experiment proceeded according to the transformation into chemical competent cells (see above). XL-1 blue cells were used for preparation of plasmid DNA, whereas BL21-(DE3)-RIL and JM83 cells served for gene expression.

#### 2.5.2.2.8 Isolation of plasmid DNA from *E. coli*

To analyse whether the PCR constructs have been properly inserted into the vector, the plasmid DNA was isolated from XL1blue cells and digested with restriction enzymes. Alternatively, *E. coli* colonies were tested for the right insert by PCR as described in Section 2.5.2.2.1. To isolate plasmid DNA, colonies were picked with pipet tips from the selective medium plate, transferred into 2 ml LB medium containing the suitable antibiotic (at a dilution of 1:1000) and placed into a 37°C shaker over night.

The cultures were transferred to a 2 ml reaction tube, harvested by centrifugation for 1 min at 16,000 g. Plasmid DNA was isolated according to the following procedure:

- 10) Resuspension of cell pellet in 300  $\mu$ l resuspension buffer.
- 11) Addition of 300  $\mu$ l of lysis buffer and incubation at RT for 5 min.
- 12) Neutralization with 300  $\mu$ l of neutralization buffer
- 13) Centrifugation at 21900 x g for 10 min.
- 14) Precipitation of DNA by mixing supernatant of centrifugation with 600  $\mu$ l isopropanol.
- 15) Centrifugation at 21900 x g for 10 min and 4°C.
- 16) Washing with 500  $\mu$ l of ice-cold 70% ethanol.
- 17) Drying of DNA pellet at 50°C for at least 10 min.
- 18) Dissolving of DNA in 30  $\mu$ l of ddH<sub>2</sub>O.

For large scale plasmid DNA isolation from 50 ml of culture medium the Pure Yield™ Plasmid MidiPrep System (Promega) was used according to the manufacturer's instructions. After the restriction digestion and DNA sequencing the plasmid DNA was transformed into the expression strains BL21-(DE3)-RIL or JM83.

### 2.5.2.2.9 Mutagenesis

For the cloning of the heavy and the light chains of the Fv fragment P9E9 as the first step the light chain V<sub>L</sub> was cloned into pASK68. Since this gene contained an internal PstI restriction site and PstI is needed to insert the heavy chain of P9E9, this restriction site in the V<sub>L</sub> gene was mutated (amino acid sequence was preserved). This experiment was performed by *in vitro* mutagenesis using the QuickChange Site-directed Mutagenesis Kit (Stratagene) and verified by restriction analysis and sequencing.

### 2.5.2.2.10 DNA sequencing

The concentration of isolated plasmid DNA was determined at 260 nm using a photometer and send to the company Eurofins Medigenomix GmbH for sequencing.

## 2.5.2.3 Production of monoclonal antibodies

For the production of monoclonal antibodies against TOM core complex from *Neurospora crassa*, fusion and single cell cloning were performed according to the standard procedure<sup>284</sup>. All experiments in cell culture to derive murine monoclonal antibodies and to analyze and test these antibodies were done in the lab of Carola Hunte in the MPI of Biophysics in Frankfurt.

### 2.5.2.3.1 Immunization of mice and cell fusion

The highly purified, detergent solubilized TOM core complex was utilized to produce murine monoclonal antibodies. The initial immunization of mice was followed by four injections with protein suspension in 4-week intervals and immunization was tested with ELISA (Section 2.5.2.3.5). The myeloma cell line P3x63Ag4.653 was used for fusion with the spleen cells isolated from an immunized mouse. The fusions were dispensed to 1 ml aliquots, frozen and stored in liquid nitrogen. See the following timetable:

1. Immunization	28.11.2000	
2. Immunization	09.01.2001	
ELISA	24.01.2001	(mouse3 showed the best signal)

3. Immunization		16.02.2001
4. Immunization		19.03.2001
ELISA		06.04.2001
5. Immunization	mouse2	23.04.2001
	mouse3	24.04.2001
Fusion1	mouse2	26.04.2001
Fusion2	mouse3	27.04.2001

Fusions were performed in Pan401 medium (5.1.6) and frozen.

Fusion1:  $1.26 \cdot 10^8$  spleen cells

Freezing medium:	PANserin401	39 ml
	DMSO	5 ml
	Glutamin	500 $\mu$ l
	OPI (100x)	500 $\mu$ l
	<u>FCS</u>	<u>5 ml</u>
		50 ml

Fusion 2:  $1.5 \cdot 10^8$  spleen cells

Freezing medium:	RPMI	38 ml
	FCS	5 ml
	DMSO	5 ml
	Glutamin	500 $\mu$ l
	Hepes	500 $\mu$ l
	Na-pyruvate	500 $\mu$ l
	$\beta$ -mercaptoethanol	50 $\mu$ l
	<u>OPI (100x)</u>	<u>500 <math>\mu</math>l</u>
		50 ml

Immunization and fusion was not done in this study, but was carried out by Simone Schmitt.

### 2.5.2.3.2 Hybridoma cell culture

While working with cell culture, it is necessary to check for contamination and for good growth rate every day by inspection under the microscope. The risk of mycoplasma contamination can be reduced by using medium without supplemented antibiotics. Handling, storing by freezing and thawing of cells was done according to standard procedures<sup>285</sup>.

For cultivation and selection of hybridoma clones, 3 x 1 ml aliquots of the frozen cells (fusion1 or fusion2) were thawed carefully in a 37°C waterbath and transferred each into 5 ml (1:5 dilution), 10 ml (1:10) dilution and 20 ml (1:20 dilution) of precooled selection medium B (Section 2.5.1.6). Cell suspensions were distributed in 96well plates with 100 µl per well and plates incubated at 37°C. After 1 ½ weeks, grown hybridoma clones could be detected and 100 µl of fresh medium was added. If about one half of the well is covered by the clone and the medium start to turn yellow, the culture (200 µl) was transferred into 1 ml of fresh medium in 24well plates. In case of well grown cells from 24well plates, 1 ml of cell suspension was mixed with 2 ml of fresh selections medium B without azaserin and incubated in 6well plates. Cells were then grown logarithmic on the 6well plates with splitting every two days. Before the transfer steps cell culture supernatants were checked by ELISA (Section 2.5.2.3.5). In case of analysis by Western Blot and subtyping, only supernatants from 6well plates were used due to the necessity of larger volumes of supernatant.

OPI represents a mixture of oxalacetate, pyruvate and insulin which improves and sustains cell growth under conditions of stress. Haz is a mixture of hypoxanthine and azaserine, which is important for selective growth for the following reasons:

-non-fused murine spleen cells die *in vitro*.

-Ag4 myeloma cells are HGPRT<sup>-</sup> (namely hypoxanthine-guanine-phosphoribosyl transferase). Azaserine interferes with the cellular purine synthesis allowing growth of cells only when these are supplemented with hypoxanthine (alternative way of purine synthesis). The alternative pathway includes formation of inosin-monophosphate as an intermedium via the HGPRT enzyme. The mutant myeloma cell lacking this enzyme will die, if it is not fused to a spleen cell.

-fused clones (hybridomas) survive since they are HGPRT<sup>+</sup> and have been immortalized via conjugation with the myeloma cells.

### 2.5.2.3.3 Single cell cloning

Hybridoma clones, which supernatants showed positive signals in ELISA test, were diluted in a single cell cloning step. This procedure has to be repeated a minimum of three times to ensure single cell cloning essential for the characterization of the antibody and the cloning of Fv fragments. Therefore, in each well of a 96well plate, 100 µl of cloning medium were pipetted and 100 µl of cell suspension of a nicely grown well were mixed with 900 µl of cloning medium (Pan401 Medium with 20% FCS and 1xOPI). 100 µl of this mixture were added to well A1 of the prefilled plate and carefully mixed. In the next step, 1:1 dilution were prepared by mixing 100 µl suspension of A1 with medium in A2, then 100 µl of this mix with well A3 and so forth. 1:1 dilutions for all lines were performed in a similar way using an 8-channel multi-channel pipet by pipetting from row 1 to row 12. Clones should be grown after 8 days and growth of the cells was checked regularly under the microscope. Supernatants were tested as above.

### 2.5.2.3.4 Freezing and cryo-storage of hybridoma cells

If cells are nicely grown in 6well plates for 2-3 cycles, cells were centrifuged at 1200 rpm for 10 min. Ice cold cryo medium (Section 2.5.1.6) was added drop wise (1 ml for one 6-well) and cells were resuspended. The suspension was filled into a precooled cryo-tube, slowly frozen in a closed Styrofoam box (2 h at -20°C, o/n at -80°C) and transferred to the liquid nitrogen storage device.

### 2.5.2.3.5 Enzyme Linked ImmunoSorbent Assays (ELISA)

After testing different ELISA techniques, the sandwich ELISA method was used for clone selection. His-tag ELISA worked with the same efficiency, but Ni<sup>2+</sup>-NTA pre-coated plates (Qiagen, Hilden, Germany) are much more expensive.

The principle of a Sandwich ELISA (or two antibodies assay) is to coat an antibody to the plate and thereafter binding of the antigen to the coated antibody. The method prevents denaturation of the antigen, which is convenient to find antibodies, which are directed against discontinuous epitopes. The assay was done with 96well plates of Greiner, because of their better signal to noise ratio in comparison to Polysorb or Maxisorb plates:

- 1) Coating of rabbit anti-TOM40 (polyclonal serum, 1:1000 diluted) in coating buffer for 1h at RT, 100  $\mu$ l per well. Empty wells by draining over sink.
- 2) 3 x washing (20 sec) with 400  $\mu$ l of wash buffer per well.
- 3) Blocking of the free sites with blocking buffer, 200  $\mu$ l per well, 15 min at RT.
- 4) 2 x washing with washing buffer, 1 x with TOM buffer.
- 5) Incubation of the TOM complex antigen (10  $\mu$ g/ml) in TOM buffer for 1h at RT, 100  $\mu$ l per well.
- 6) 3 x washing with TOM buffer.
- 7) Incubation with undiluted cell culture supernatant or polyclonal mouse serum (diluted 1:10000) for 1h at RT, 100  $\mu$ l per well.
- 8) washing step as 6).
- 9) Incubation with secondary antibody (anti-mouse IgG; Sigma) conjugated to alkaline phosphatase, diluted 1:1000 in antibody solution, for 1 h at RT.
- 10) washing step as 6).
- 11) Developing of the plates in freshly prepared development solution for 15 – 30 min, 100  $\mu$ l per well.
- 12) Measurement of the optical density at 405 nm (450 nm as reference) with an ELISA reader.

Composition of the buffers is shown in Table 2.6 in Section 2.5.1.6.

### **2.5.2.3.6 Determination of the antibody subtype**

Immunological subtyping of the monoclonal antibodies was done using the kit from Boehringer (1183117) according to the manufacturer's instructions. This kit is also based on ELISA.

### **2.5.2.3.7 Large-scale production of monoclonal antibodies**

For the purification of monoclonal antibodies the production in cell culture was scaled up and the cells were adapted to FCS-free medium. Cells were grown in 6well plates and adapted stepwise to cloning medium (Section 2.5.1.6) supplemented only with 5% FCS. Hybridoma cells from three fully grown 6well plates of each antibody were harvested by centrifugation (1200 rpm, 10 min) and resuspended in 30 ml of FCS-free cloning medium. The suspension

was transferred in a 260 ml cell culture flask, which was kept in the incubator in upright position. After ~4 days, further 30 ml of medium was added. Cells were grown for another three days, harvested by centrifugation, supplemented with 1/10 volume of 1 M Tris-Cl pH 8.0 and stored at -20°C. This method will result in hybridoma supernatants with very low FCS content, which is important for subsequent purification of the antibodies. For each antibody, three flasks with the total volume of 180 ml were cultivated.

#### **2.5.2.3.8 Monoclonal antibody purification by protein A affinity chromatography**

Affinity purification is the most suitable method for monoclonal antibody purification. It is based on the binding of the antibody molecule to immobilized staphylococcal protein A or streptococcal protein G. These are bacterial F<sub>c</sub> receptors that mainly bind to the F<sub>c</sub> part of the antibody molecule. For protein A affinity purification of monoclonal antibodies from tissue culture supernatant, the 1 ml HiTrap Protein A HP column (Section 2.5.1.8) was used together with the Äkta Chromatography System (Amersham Bioscience). The ~200 ml of the cell culture supernatant was loaded on the equilibrated column (binding buffer, Section 2.5.1.6) using a sample pump (flow rate: 1 ml/min) and the column was washed with 10 ml of binding buffer. Bound antibodies were eluted at low pH with 10 ml of elution buffer. The pH was adjusted to higher pH directly after fractionation. Typical monoclonal antibody yields are 2-20 µg/ml. The protein concentration was measured and the fractions containing protein were immediately concentrated using membrane filters with a 50 kDa molecular mass cut-off (Millipore, Pall). Purified proteins were used for binding assays via analytical gel filtration.

#### **2.5.2.3.9 Analytical gel filtration binding assay**

In this experiment, the purified antibodies were adjusted to TOM buffer (Section 2.5.1.6, ELISA) and the purified TOM core complex was mixed with the respective antibody in the molar ratio of 1:4 and the total volume of 50 µl (approx. 50 µg of TOM core and 70 µg of the antibody). After incubation of 1 h at RT, the mixture was sterile filtered and run on a Superose 6 PC 3.2/30 column (Section 2.5.1.8) using a SMART system. Fractions of 50 µl were collected and analyzed by SDS-PAGE.

### 2.5.2.3.10 Biochemical pull-down assay

His-tagged TOM core complex antigen was bound to Ni-NTA Magnetic Agarose beads (Qiagen) to pull down monoclonal antibodies from cell culture supernatants. Therefore, 1 ml of cleared cell culture supernatant was adjusted to TOM buffer conditions (Section 2.5.1.6), mixed with 25  $\mu$ l (50  $\mu$ g) of TOM core complex and incubated for 1 h at 4°C on an end-over-end shaker. After addition of 10  $\mu$ l of Ni-NTA beads suspension the mixture was incubated for an additional hour. The supernatant was removed and the beads were washed three times with TOM buffer supplemented with 20 mM imidazole. The bound proteins were eluted with 25  $\mu$ l of elution buffer (TOM buffer + 300 mM imidazole) and analyzed by SDS-PAGE and Western Blot.

## 2.5.2.4 Methods in protein biochemistry

### 2.5.2.4.1 Gene expression

After transformation of the Tom40 plasmid DNA in *E. coli* BL21-(DE3)-RIL cells, 100 ml of LB medium containing ampicillin (100  $\mu$ g/ml) and chloramphenicol (30  $\mu$ g/ml) were inoculated with the transformed cells (500  $\mu$ l suspension, see Section 2.5.2.2.6) and incubated in a 37°C shaker over night. Three litres of LB<sub>Amp,Cam</sub> medium were inoculated with this preculture at a ratio of 1:40 and incubated at 37°C. After cell density reached OD<sub>600</sub> 0.6 to 0.8, expression was induced with 0.5 mM IPTG (end concentration). The cells were incubated at 24°C over night and harvested by centrifugation at 4°C and 2580 x g (4000 rpm, JLA 8.1000 rotor) for 10 min. Finally the supernatant was discarded and the resulting pellet was gently resuspended in cold lysis buffer (Section 2.5.1.6) at 1/100 of the culture volume. For storage, the suspension was frozen and stored at -20°C. For test expressions 100 ml of LB medium as well as different expression times and temperatures were applied (normally 3 h at 37°C, O/N at 37°C and O/N at 24°C).

The monoclonal antibody fragments were expressed bicistronically in the plasmid pASK68. The plasmid was transformed into *E. coli* strain JM83 for periplasmic protein expression. 100 ml LB medium containing 100  $\mu$ g/ml Ampicillin and 30  $\mu$ g/ml streptomycin were inoculated using the transformed *E. coli* stock and cells grew over night at 37°C. 75 ml of the preculture were used to inoculate 3 l LB<sub>Amp</sub> medium in 3 l flasks (Fernbach flasks, Corning) preincubated at 22.5°C. After cells were grown at 22.5°C and 140 rpm to an OD of 0.5, expression was induced by adding IPTG to a final concentration of 1 mM, and incubation was

continued under the same condition for 3 hrs. Cells were harvested as described in Section 2.5.2.4.3

#### **2.5.2.4.2 Cell lysis**

The thawed suspended cells were disrupted on ice using an appropriately equipped sonicator for the suspended volume (parameters: big tip, pulse maximum and duty cycle 60) for 10 min. After centrifugation at 24000 x g (20000 rpm in Beckman JA20 rotor) and 4°C for 20 min, the cell pellet including the Tom40 proteins was used for isolation of inclusion bodies.

#### **2.5.2.4.3 Periplasmic extraction**

For periplasmic extraction cells were harvested at 4000 x g at 4°C in precooled centrifugation buckets and gently resuspended in ice-cold extraction buffer (Section 2.5.1.6) at 1/100 of the culture volume. After 30 min incubation on ice, cells and cell debris were spun down at 5000 x g for 45 min at 4°C. The periplasmic extract was poured off and stored at -20°C.

#### **2.5.2.4.4 Isolation of inclusion bodies**

Tom40 from *Neurospora crassa* was expressed in inclusion bodies which were isolated as follows: The cell pellet (see Section 2.5.2.4.2) from 1 l of culture was washed three times with washing buffer 1 and twice with washing buffer 2 (all centrifugations at 24000 x g for 15 min). Sonifying could facilitate resuspension of the inclusion bodies. Between the washing steps, samples of the resuspended inclusion bodies and of the washing fractions were taken for SDS-PAGE analysis.

#### **2.5.2.4.5 Protein refolding**

The washed inclusion bodies (Section 2.5.2.4.4) were dissolved in urea buffer (Section 2.5.1.6) and stirred for 1 h at 70°C. The mixture was cleared by centrifugation at 24000 x g for 15 min and contained the unfolded Tom40. The protein concentration was measured spectrometrically and diluted to the end concentration of 100-200 µg/ml. Therefore, the protein solution was added drop wise (under stirring) to 500 ml of cold dilution buffer (Section

2.5.1.6) using a peristaltic pump and incubated over night at 4°C. The diluted solution was centrifuged resulting in refolded Tom40 in the supernatant.

#### 2.5.2.4.6 Streptavidin affinity purification

Since the vector pASK68 provides a C-terminal fusion of strep-tag I to the heavy chain of the fragment, the Fv fragments P1C10 and P9E9 were purified by streptavidin affinity chromatography. The periplasmic extract (from 48 and 6 l of culture volumes for P1C10 and P9E9, respectively) was loaded onto a 1 ml Strep-Tactin Sepharose column (IBA, Göttingen) equilibrated with loading buffer (Section 2.5.1.6). The column was washed with 10 column volumes of loading buffer and specifically bound protein was eluted with elution buffer (Section 2.5.1.6). The eluted protein was collected and concentrated using spin filtration devices (10 kDa cut-off) to an appropriate concentration for binding experiments. The purification was monitored by SDS-PAGE analysis and Western blot was performed to detect the light chain which carries a C-terminal myc-tag.

#### 2.5.2.4.7 Purification of TOM core complex

TOM core complex was isolated from a *Neurospora crassa* strain (GR-107) that carried a version of Tom22 with a hexahistidinyl tag at its C terminus. Growth of the cells and preparation of mitochondria were performed as described previously<sup>230,235,238</sup>. The TOM core complex was isolated according to Athing et al., 1999<sup>231</sup>, with some modifications in the protocol. Isolated mitochondria were solubilized in solubilization buffer (Section 2.5.1.6) at a protein concentration of 10 mg/ml for 30 min at 4°C. Insoluble material was removed by centrifugation, and the clarified extract was loaded onto the equilibrated (equilibration buffer, Section 2.5.1.6) nickel nitrilotriacetic acid agarose column (Ni-NTA; Qiagen) using 4 ml resin per g of total mitochondrial protein. The column was washed with 5 column volumes of TOM washing buffer 1 and with 10 vol of TOM washing buffer 2. The bound protein was eluted with 4 column volumes of TOM elution buffer (Section 2.5.1.6). The eluted fraction of the Ni-NTA column was loaded onto an equilibrated Resource Q anion exchange column. The complex was eluted with a linear potassium chloride gradient (buffers in Section 2.5.1.6). For further purification, the fractions containing TOM core complex were pooled, concentrated to 500 µl (spin filtration devices from Pall Corporation, 100 kDa cut-off) and loaded onto a Superose 6 size exclusion column. Proteins were eluted in gel filtration buffer.

Stock solutions of purified TOM core complex were stored at a protein concentration of 10-20 mg/ml at 4°C. An average preparation of the protein started with ~1.5 kg of *Neurospora* cells (wet wt) which yielded about 5 g of mitochondrial proteins and 1-2 mg of purified Tom core complex.

#### **2.5.2.4.8 Co-complex formation of TOM core complex and Fv fragments**

To obtain the co-complex, purified TOM core complex was mixed with the Fv fragment at a molar ratio of 1:10 and incubated at RT for 30 min in gel filtration buffer (Section 2.5.1.6). Due to their much smaller size, the unbound Fv fragments were separated from the co-complex by size exclusion chromatography using the Superose 6 column (Section 2.5.1.8). The co-complex containing fractions were pooled, analyzed by SDS-PAGE and concentrated.

#### **2.5.2.4.9 Detergent exchange**

For the detergent exchange, TOM core complex was purified in DDM according to the standard procedure (Section 2.5.2.4.7) except the gel filtration, which was performed in the new detergent. After anion exchange chromatography TOM core complex was concentrated to approx. 500 µl of volume and its buffer was adjusted to the concentration of the new detergent as shown in Table 2.3. After the clarifying spin, the protein was loaded onto the Superose 6 column (Section 2.5.1.8) and run in gel filtration buffer (Section 2.5.1.6) supplemented with the new detergent instead of β-DDM. This experiment was monitored by urea SDS-PAGE (Section 2.5.2.5.1).

### **2.5.2.5 Detection and quantitative determination of the protein**

#### **2.5.2.5.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)**

Proteins can be separated according to their molecular mass by discontinuous sodiumdodecylsulfat polyacrylamide gelelectrophoresis (SDS-PAGE) as described<sup>286</sup>. Smaller proteins migrate faster through the gel towards the negative pole than larger ones. Although the electrophoretic mobility depends not only on molecular mass but also on the shape and the charge of the protein, the molecular mass could be estimated through

comparison to a marker because the protein solution is denaturated under reducing conditions ( $\beta$ -mercaptoethanol) by adding SDS (it binds to the unfolded peptide and has a high negative charge).

In general, the standard SDS gels (Table 2.10) served for separation of proteins, but in case of TOM core complex urea-SDS gels were used (Table 2.10), which show much higher resolution for proteins in the low molecular mass range (necessary to separate the small Tom components). The concentration of acrylamide and bis-acrylamide in the standard separating gel was chosen according to the molecular size of the proteins to be separated (mostly 16 or 17.5%). Glass plates of 160x140 mm and spacers of 1 mm thickness were utilized. Before loading, the samples were dissolved in appropriate volumes of 4 x sample buffer and incubated at 95°C for 5 min. The electrophoresis for standard gels was performed at 30-45 mA for 1-2 h in 1 x SDS running buffer and in case of urea gels at 35 mA for 3-4 h in urea gel running buffer. Protein molecular mass standards (#SM0431, Fermentas) were used in each gel.

<b>Standard SDS gel</b>	
Separating gel	Separating gel 16% (w/v) acrylamide, 0.16-0.33% (w/v) bis-acrylamide, 380 mM Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.035% (v/v) TEMED.
Stacking gel	Stacking gel 5% (w/v) acrylamide , 0.1% (w/v) bis-acrylamide, 60 mM Tris-HCl (pH 6.8), 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.035% (v/v) TEMED
<b>Urea-SDS gel</b>	
Separating gel	Separating gel 20% (w/v) acrylamide, 0.25% (w/v) bis-acrylamide, 6 M urea, 750 mM Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.07% (w/v) APS, 0.13% (v/v) TEMED.
Stacking gel	Stacking gel 5% (w/v) acrylamide , 0.07% (w/v) bis-acrylamide, 6 M urea, 125 mM Tris-HCl (pH 6.8), 0.1% (w/v) SDS, 0.05% (w/v) APS, 0.17% (v/v) TEMED

**Table 2.10:** *Composition of SDS and urea-SDS gels.*

All SDS-PAGE experiments, which were done in the MPI of Biophysics in Frankfurt were based on commercial NuPAGE® Novex 4-12% Bis-Tris gels (Invitrogen).

#### **2.5.2.5.2 Staining SDS-PA gels with Coomassie brilliant blue**

To visualize the protein on the gel, Coomassie Brilliant Blue dye was used, which allows detection of protein amounts down to 0.1 µg. After SDS-PAGE and removal of the bottom and stacking gels, the separating gel was heated in Coomassie stain for 1 min using a microwave, incubated for 20-40 min at room temperature, subsequently washed with water and destained in Coomassie destain until distinct bands became visible against a clear background. The gel was scanned and then dried overnight between two gel-drying films.

#### **2.5.2.5.3 Silver staining of SDS-PA gels**

Silver staining of SDS gels represents a staining method that shows much higher sensitivity than staining with Coomassie brilliant blue and is particular convenient to detect small amounts of proteins or proteins of small size. To analyze small TOM components, silver staining is necessary and was performed by the following incubation steps (under shaking):

- 30 min incubation in TCA solution
- 30 min fixation in fixation solution
- 20 min reduction in reducing solution
- 5 x 5 min wash in ddH<sub>2</sub>O
- 15 min incubation in silver solution
- 5 x 5 min wash in ddH<sub>2</sub>O
- development in development solution till bands become visible
- stop the reaction by washing in ddH<sub>2</sub>O or 0.5 M EDTA

All solutions were prepared freshly and are shown in Table 2.6 (Section 2.5.1.6).

#### **2.5.2.5.4 Transfer of proteins to nitrocellulose membrane (Western-blot)**

Proteins separated via SDS-PAGE were transferred onto nitrocellulose membrane using the semi-dry blotting method<sup>149,150</sup>. The gel, the membrane, and six sheets of Whatman filter paper (3MM) were incubated in transfer buffer. Three sheets of filter paper were placed on the anode electrode followed by the membrane and on top of it the gel. This was covered with other three filter papers and with the cathode electrode. The transfer was performed at 200 mA for 1.5 h. To verify transfer efficiency, the nitrocellulose membranes were reversibly stained with Ponceau S solution. Afterwards, the membrane was immunodecorated.

#### **2.5.2.5.5 Immunodecoration**

To visualize the immobilized proteins on nitrocellulose membrane, immunodecoration with specific antibodies was carried out. Membranes were first incubated for 1 h in 5% (w/v) milk powder/3% (w/v) BSA in TBS (Section 2.5.1.6) to block all non-specific binding sites. The immunodecoration was done for 1 h at RT or over night at 4°C, with specific antiserum (1:200 to 1:1000 dilutions in milk/BSA in TBS). The membrane was then washed 3 times (each wash lasts 5 min), with TBS, TBS/0.05% (w/v) Triton X-100 and again with TBS, and incubated for 1 h with horseradish peroxidase coupled to secondary goat anti-rabbit-IgG (Sigma, diluted 1:10.000 dilutions in milk/BSA in TBS). The membrane was again washed (as above) and treated with ECL reagents (H<sub>2</sub>O<sub>2</sub> was added fresh). The luminescence reactions were detected with Röntgen films. In case of the decoration with mouse  $\alpha$ -myc antibody, all steps were done in NET/gelatine buffer. The mouse  $\alpha$ -myc antibody was used undiluted (kindly provided from the group of Kai Hell, LMU Munich) and the conjugated goat anti-mouse-IgG (Sigma) served as secondary antibody.

#### **2.5.2.5.6 Determination of the protein concentration by Bradford assay**

The Bradford assay is based on the fact that the absorption of Coomassie-Brilliant Blue dye shifts from 465 nm to 595 nm when it binds to protein. Its sensitivity allows to measure protein concentrations between 0.1 mg/ml and 30 mg/ml per ml Bradford solution. Protein concentrations were determined using the “Bio-Rad-Protein assay” reagent (Bio-Rad)<sup>152</sup>. The protein solution was mixed with 1 ml reagent (1:5 dilution) and incubated for at least 5 min at RT. To increase the accuracy of the assay six samples were examined (twice 2.5  $\mu$ l, 5  $\mu$ l and

7.5 µl of the protein solution). The absorbance was measured at 595 nm using a 10 mm path length microcuvette. Finally, the protein concentration was calculated according to a standard curve (bovine serum albumin standard from Biorad Laboratories was used as protein standard).

#### **2.5.2.5.7 Determination of the protein concentration by UV spectroscopy**

Protein concentration could be also determined by measuring the absorbance at 280 nm using a photometer. Therefore, the extinction of an appropriate dilution of 150-200 µl of the protein solution was measured in a quartz cuvette with the buffer solution as blank value. The theoretical extinction coefficient was calculated by the ProtParam tool (Section 2.5.1.9) and used for the determination of the protein concentration.

#### **2.5.2.5.8 Concentration of proteins**

For crystallization high protein concentrations were required. Protein was concentrated by centrifugation using Amicon Ultra 10K or 30K devices (Millipore) or Nanosep Centrifugal 3 and 10K Devices (Pall Corporation) according to the manufacturer's instructions.

#### **2.5.2.5.9 Protein precipitation**

Sometimes the concentration of a protein fraction was less than is required for SDS-PAGE. Therefore, the protein was precipitated with 1/5 volume of 72% trichloroacetic acid (TCA), incubated on ice for 10 min and centrifuged at 14000 rpm for 1 min. The precipitate was washed with 0.5 ml of TCA wash buffer (Section 2.5.1.6), and dried at 56°C for 10 min. The precipitated protein was dissolved in 1 x sample buffer and analysed by gel electrophoresis.

## **2.5.2.6 Crystallization**

### **2.5.2.6.1 Crystallization of TOM core complex**

The TOM core complex alone and in complex with the Fv fragment were concentrated to 10-14 mg/ml and subjected to crystallization trials. The TOM core complex was initially crystallized by the vapour diffusion method using a hanging drop setup. 1  $\mu$ l of protein solution was mixed with 1  $\mu$ l mother liquor to constitute the hanging drop in twenty-four well Linbro-plates. Initially Crystal Screens I, II and the MembFac Screen from Hampton Research were used. Small crystals appeared after two days in the solution containing 0.2 M magnesium chloride, 0.1 M HEPES, pH 7.5 and 30% (v/v) PEG 400. Large crystals were obtained after setting up fine screen (mixing 1.5  $\mu$ l of protein solution with 1  $\mu$ l of well solution) with solutions containing 100 mM ammonium citrate, 50 mM Tris-Cl pH 7.2, 20-30% PEG 400. The SDS gel analysis of washed and dissolved crystals showed that all components of the TOM core complex were present in the crystals. This work was done by Simone Schmitt<sup>238</sup>.

In case of the TOM core complex with bound antibody initial crystallisation screens were set up using a pipetting robot. 200 nl of the concentrated protein solution were mixed with 200 nl of reservoir solution in 96well plates. The initial screening was performed with Classics, pHClear I, pHClear II and the PACT suites (Qiagen). Small crystals appeared in sitting drop and were reproducible in hanging drop with the optimized solution containing 0.2 M calcium chloride, 0.1M HEPES, pH 7.5 and 25-28% (v/v) PEG 400. Large crystals of the complex also grew with the original reservoir solution.

### **2.5.2.6.2 Crystallization in capillary**

The crystallization in capillary was carried out similar to the protocol of Simone Schmitt<sup>238</sup>.

### **2.5.2.6.3 Crystallization with additives**

The detergent additive screens 1, 2 and 3 from Hampton Research were used according to the manufacturer's instructions. After determination of the optimal PEG400 concentration in the crystallization trials for each TOM core protein batch, the hanging drop vapour diffusion method was performed as described above by mixing 1.5  $\mu$ l of the protein and 1  $\mu$ l of the

reservoir solution. In the next step, 0.28  $\mu\text{l}$  of the additive were added each to one drop (1:10 dilution of the additive). If crystals occurred fine screens were applied to optimize crystal growth and crystals were measured at synchrotron (Section 2.5.2.7). In case of copper chloride as additive, TOM core complex and copper chloride (10 mM) were mixed in a 10:1 ratio and subjected to crystallisation attempts.

#### **2.5.2.6.4 Treatment of crystals with glutaraldehyde**

Incubation of the crystals to glutaraldehyde (crystallization-grade, Patrick Cramer, Gene Centre, Munich) was performed in the 24well plates by hanging drop vapour diffusion. 5  $\mu\text{l}$  of glutaraldehyde were pipetted into the deep reservoir well and the cover slide containing the hanging crystal drop with the crystals was positioned on the well. Different incubation times from 0.5 h to 5 h in 30 min steps (at 20°C) were applied and the crosslink efficiency was tested by transfer of the crystals in ddH<sub>2</sub>O. Manipulated crystals were measured at synchrotron.

#### **2.5.2.6.5 Crystallisation with heavy metal atoms**

All heavy metals except tantalbromid (Proteros biostructures) and all other heavy metal atom clusters (Patrick Cramer, Gene Center, Munich) were provided by the Heavy Metal Atom Screens from Hampton Research. Only a very small spatula tip of the heavy metal was solved in 100  $\mu\text{l}$  of the reservoir solution to prepare stock solutions and 0.1 to 0.5  $\mu\text{l}$  of the stock solution was added to the crystallization drop. Either crystals were transferred to the heavy metal atom containing drop or the heavy metal solution was added to the crystals. In case of co-crystallization the heavy metal and the crystallisation solution were mixed and crystals should grow within the next few days.

#### **2.5.2.7 Data collection and processing**

The measurement of the crystals took place at the synchrotron radiation sources BW6 at Deutsches Elektronen Synchrotron (DESY) in Hamburg and at the X06SA Swiss Light Source (SLS) at Paul Scherrer Institut in Villingen. The crystals of the TOM core complex were exposed to X-rays both at RT in capillary and frozen, whereas crystals of the

Tom:antibody complex were only measured frozen. The crystals were flash-frozen in a stream of cold nitrogen gas at 100 K (Oxford Cryosystems) directly from the crystal solution, which represents a cryo-protectant. The wavelength was set to 1.05 Å. Data was recorded using a MarCCD detector and the stability of the crystals in the beam allowed collection of the first isomorphous complete dataset at 8 Å resolution. Data were processed with DENZO and data reduction was performed with SCALEPACK<sup>240</sup>. The space group was determined to be C2 with unit cell dimensions of a=181 Å, b=113 Å, c=129 Å,  $\beta=113^\circ$ .

### 3 Abbreviations

3'-end	three prime (hydroxyl) end of the oligonucleotide	GST	Glutathione-S-Transferase
		h	hours
5'-end	five prime (phosphate) end of the oligonucleotide	HADDOCK	High Ambiguity Driven protein–protein Docking
Å	Ångstrom	hs/h	<i>Homo sapiens</i> /human
aa	amino acid residue(s)	His (H)	Histidine
AAA	ATPases associated with various cellular activities	IPTG	Isopropyl-β-D-Thiogalactopyranosid
		K	Kelvin
Adrm1/Arm1	Adhesion regulating molecule 1	kDa	kilodaltons
Amp <sub>r</sub>	Ampicillin resistance	kV	kilovolt
APS	Ammonium peroxy disulphate	LB	Luria-Bertani
Asp	Aspartate	mm/m	<i>Mus musculus</i> /murine
at	<i>Arabidopsis thaliana</i>	MCS	Multiple cloning site
ATP	Adenosintriphosphate	MIU	motif interacting with ubiquitin
A20 ZnF	A20 zinc finger	mRNA	messenger ribonucleic acid
BSA	Bovine Serum Albumin	Nc/N.c.	<i>Neurospora crassa</i>
Cam <sub>r</sub>	Chloramphenicol resistance	NMM	New Minimal Medium
ce	<i>Caenorhabditis elegans</i>	NMR	Nuclear magnetic resonance
CP	core particle of proteasome	NOE	Nuclear Overhauser Effect
C-terminus	Carboxy terminus	N-terminus	amino terminus
CUE	coupling of ubiquitin conjugation to endoplasmatic reticulum degradation	OD	optical density
		O/N	over night
CV	Column volumes	ORF	Open reading frame
Da	Dalton	PAGE	polyacrylamid Gel Electrophoresis
ddH <sub>2</sub> O	Millipore filtrated, demineralised water	PACE	proteasome–Associated Control Element
DNA	Deoxyribonucleic Acid	PAN	proteasome–activating nucleotidase
dm	<i>Drosophila melanogaster</i>	PBS	Phosphate buffered saline
dNTP	Desoxynucleosidtriphosphate	PCR	Polymerase chain reaction
DTT	Dithiothreitol	PDB	Protein Data Bank
DUB	Deubiquitinating enzyme	PEG	Polyethylenglycol
DUIM	double-sided UIM	PH	Pleckstrin Homology
<i>E. coli</i>	<i>Escherichia coli</i>	PIP	phosphatidylinositolphosphate
EDTA	Ethylendiamintetraacetate	PP	PreScission Protease
ERAD	Endoplasmic Reticulum Associated Protein Degradation	ppm	Parts per million
		Pru	Pleckstrin–like receptor for ubiquitin
GAT	Golgi-localized, gamma-ear-containing, ADP-ribosylation-factor-binding protein and target of myb	PVDF	polyvinylidene difluoride
		RBS	Ribosomal binding site
		RP	regulatory particle of proteasome
GFP	Green Fluorescent Protein	rpm	Rounds per minute

## Abbreviations

---

GLUE	GRAM-like ubiquitin binding in ELL-associated protein 45	Rpn	regulatory particle non-ATPase
Rpt	regulatory particle triple-A protein	RT	room temperature
SAP	Shrimp Alkaline Phosphatase	Ub	Ubiquitin
sc	<i>Saccharomyces cerevisiae</i>	UBA	ubiquitin-associated
SDS	Sodiumdodecylsulfate	Ubc	ubiquitin-interacting motif
Ser	Serine	UBL	ubiquitin-like
sp	<i>Schizosaccharomyces pombe</i>	UCH	Ubiquitin C-terminal hydrolase
TBS	Tris-buffered saline	UEV	Ubc E2 variant
TCA	Tri-chloroacetate	UIM	ubiquitin-interacting motif
TEMED	N,N,N',N'-Tetramethylethylenediamine	ZnF UBP	<u>z</u> inc <u>f</u> inger <u>u</u> biquitin-specific <u>p</u> rocessing protease
TIM	Translocase of the inner mitochondrial membrane		
TOM	Translocase of the outer mitochondrial membrane		
Tris	Tris(hydroxymethyl)-aminomethan		
U	Units		

## 4 References

- 1 P. Schreiner, X. Chen, K. Husnjak et al., *Nature* **453** (7194), 548-52 (2008).
- 2 A. Hershko and A. Ciechanover, *Annu Rev Biochem* **67** (10), 425-79. (1998).
- 3 M. Groll, M. Bochtler, H. Brandstetter et al., *Chembiochem* **6** (2), 222-56 (2005).
- 4 D. M. Blow, *Trends in biochemical sciences* **22** (10), 405-8 (1997).
- 5 P. E. Petrides G. Löffler, P. H. Heinrich, *Springer Lehrbuch* (2006).
- 6 M. Schmidt, J. Hanna, S. Elsasser et al., *Biological chemistry* **386** (8), 725-37 (2005).
- 7 C. M. Pickart and R. E. Cohen, *Nature reviews* **5** (3), 177-87 (2004).
- 8 D. Voges, P. Zwickl, and W. Baumeister, *Annu Rev Biochem* **68** (1), 1015-68. (1999).
- 9 P. Zwickl, D. Voges, and W. Baumeister, *Philos Trans R Soc Lond B Biol Sci* **354** (1389), 1501-11. (1999).
- 10 A. L. Schwartz and A. Ciechanover, *Annu Rev Med* **50** (4), 57-74. (1999).
- 11 B. Tomkinson, *Trends Biochem Sci.* **24** (9), 355-9. (1999).
- 12 B. Rockel, J. Peters, B. Kuhlorgen et al., *EMBO J.* **21** (22), 5979-84. (2002).
- 13 J. Löwe, D. Stock, B. Jap et al., *Science* **268** (5210), 533-9. (1995).
- 14 M. Groll, L. Ditzel, J. Lowe et al., *Nature* **386** (6624), 463-71 (1997).
- 15 M. Unno, T. Mizushima, Y. Morimoto et al., *Structure* **10**(5), 609-18 (2002).
- 16 M. Groll and T. Clausen, *Curr Opin Struct Biol* **13** (6), 665-73 (2003).
- 17 A. Köhler, P. Cascio, D.S. Leggett et al., *Molecular Cell* **7** (6), 1143-52 (2001).
- 18 M. Groll, M. Bajorek, A. Kohler et al., *Nat Struct Biol* **7** (11), 1062-7 (2000).
- 19 F. G. Whitby, E. I. Masters, L. Kramer et al., *Nature* **408** (6808), 115-20. (2000).
- 20 D. Nandi, P. Tahiliani, A. Kumar et al., *Journal of biosciences* **31** (1), 137-55 (2006).
- 21 W. Heinemeyer, M. Fischer, T. Krimmer et al., *J Biol Chem* **272** (40), 25200-9. (1997).
- 22 L. Borissenko and M. Groll, *Chem Rev* **107** (3), 687-717 (2007).
- 23 O. Cux, K. Tanaka, and A. L. Goldberg, *Annu Rev Biochem* **65** (1), 801-47. (1996).
- 24 M. Hochstrasser, *Annu Rev Genet* **30** (1), 405-39. (1996).
- 25 M. Groll and R. Huber, *Int J Biochem Cell Biol* **35** (5), 606-16 (2003).
- 26 M. H. Glickman, D. M. Rubin, V. A. Fried et al., *Mol Cell Biol* **18** (6), 3149-62. (1998).
- 27 W. Dubiel, K. Ferrell, and M. Rechsteiner, *Molecular biology reports* **21** (1), 27-34 (1995).
- 28 D. Finley, K. Tanaka, C. Mann et al., *Trends Biochem Sci* **23** (7), 244-5. (1998).
- 29 P. Zwickl, D. Ng, K. M. Woo et al., *J Biol Chem* **274** (37), 26008-14. (1999).
- 30 M. Schmidt, D. Zantopf, R. Kraft et al., *J Mol Biol* **288** (1), 117-28. (1999).

- 31 T. Sone, Y. Saeki, A. Toh-e et al., *J Biol Chem* **279** (27), 28807-16 (2004).
- 32 R. Verma and R. J. Deshaies, *Cell* **101** (4), 341-4. (2000).
- 33 D. M. Smith, N. Benaroudj, and A. Goldberg, *Journal of structural biology* **156** (1), 72-83 (2006).
- 34 R. Hartmann-Petersen, K. Tanaka, and K. B. Hendil, *Archives of biochemistry and biophysics* **386** (1), 89-94 (2001).
- 35 U. M. Gerlinger, R. Guckel, M. Hoffmann et al., *Molecular biology of the cell* **8** (12), 2487-99 (1997).
- 36 C. N. Larsen and D. Finley, *Cell* **91** (4), 431-4. (1997).
- 37 B. C. Braun, M. Glickman, R. Kraft et al., *Nat Cell Biol* **1** (4), 221-6. (1999).
- 38 E. Strickland, K. Hakala, P. J. Thomas et al., *The Journal of biological chemistry* **275** (8), 5565-72 (2000).
- 39 M. H. Glickman, D. M. Rubin, H. Fu et al., *Mol Biol Rep* **26** (1-2), 21-8. (1999).
- 40 A. V. Kajava, *The Journal of biological chemistry* **277** (51), 49791-8 (2002).
- 41 S. Elsasser, R. R. Gali, M. Schwickart et al., *Nat Cell Biol* **4** (9), 725-30 (2002).
- 42 Y. Saeki, T. Sone, A. Toh-e et al., *Biochemical and biophysical research communications* **296** (4), 813-9 (2002).
- 43 C. R. Wilkinson, M. Seeger, R. Hartmann-Petersen et al., *Nature cell biology* **3** (10), 939-43 (2001).
- 44 T. G. Ortolan, P. Tongaonkar, D. Lambertson et al., *Nature cell biology* **2** (9), 601-8 (2000).
- 45 B. L. Bertolaet, D. J. Clarke, M. Wolff et al., *Nature structural biology* **8** (5), 417-22 (2001).
- 46 Q. Deveraux, V Ustrell, C. Pickart et al., *J. Biol. Chem.* **272**, 182-8 (1994).
- 47 S. Elsasser, D. Chandler-Militello, B. Muller et al., *The Journal of biological chemistry* **279** (26), 26817-22 (2004).
- 48 Y. A. Lam, T. G. Lawson, M. Velayutham et al., *Nature* **416** (6882), 763-7 (2002).
- 49 K. Asano, H. P. Vornlocher, N. J. Richter-Cook et al., *J Biol Chem* **272** (43), 27042-52. (1997).
- 50 M. Seeger, R. Kraft, K. Ferrell et al., *Faseb J* **12** (6), 469-78. (1998).
- 51 N. Wei, T. Tsuge, G. Serino et al., *Curr Biol* **8** (16), 919-22. (1998).
- 52 K. Hofmann and P. Bucher, *Trends Biochem Sci* **23** (6), 204-5. (1998).
- 53 L. Aravind and C. P. Ponting, *Protein Sci* **7** (5), 1250-4. (1998).
- 54 R. Verma, L. Aravind, R. Oania et al., *Science* **298** (5593), 611-5 (2002).
- 55 T. Yao and R. E. Cohen, *Nature* **419** (6905), 403-7 (2002).
- 56 Y. Dong, M. A. Hakimi, X. Chen et al., *Molecular cell* **12** (5), 1087-99 (2003).

- 57 M. Fujimuro, K. Tanaka, H. Yokosawa et al., *FEBS Lett* **423** (2), 149-54. (1998).
- 58 X. Wang, C. F. Chen, P. R. Baker et al., *Biochemistry* **46** (11), 3553-65 (2007).
- 59 D. M. Rubin, S. van Nocker, M. Glickman et al., *Mol Biol Rep* **24** (1-2), 17-26. (1997).
- 60 C. Schaubert, L. Chen, P. Tongaonkar et al., *Nature* **391** (6668), 715-8. (1998).
- 61 H. Hiyama, M. Yokoi, C. Masutani et al., *J Biol Chem* **274** (39), 28019-25. (1999).
- 62 H. Rao and A. Sastry, *The Journal of biological chemistry* **277** (14), 11691-5 (2002).
- 63 F. Hofmann, F. Martelli, D. M. Livingston et al., *Genes Dev* **10** (23), 2949-59. (1996).
- 64 I. Kim, K. Mi, and H. Rao, *Molecular biology of the cell* **15** (7), 3357-65 (2004).
- 65 L. Kaplun, R. Tzirkin, A. Bakhrat et al., *Molecular and cellular biology* **25** (13), 5355-62 (2005).
- 66 H. Richly, M. Rape, S. Braun et al., *Cell* **120** (1), 73-84 (2005).
- 67 M. Seeger, R. Hartmann-Petersen, C. R. Wilkinson et al., *The Journal of biological chemistry* **278** (19), 16791-6 (2003).
- 68 DS. Leggett, J. Hanna, A. Borodovsky et al., *Mol Cell*. **10** (3), 495-507. (2002).
- 69 V. Maytal-Kivity, N. Reis, K. Hofmann et al., *BMC biochemistry* **3**, 28 (2002).
- 70 A. Guterman and M. H. Glickman, *The Journal of biological chemistry* **279** (3), 1729-38 (2004).
- 71 R. Verma, S. Chen, R. Feldman et al., *Mol Biol Cell* **11** (10), 3425-39. (2000).
- 72 T. A. Chernova, K. D. Allen, L. M. Wesoloski et al., *The Journal of biological chemistry* **278** (52), 52102-15 (2003).
- 73 J. Hanna, D. S. Leggett, and D. Finley, *Molecular and cellular biology* **23** (24), 9251-61 (2003).
- 74 Y. A. Lam, W. Xu, G. N. DeMartino et al., *Nature* **385** (6618), 737-40 (1997).
- 75 H. Holzl, B. Kapelari, J. Kellermann et al., *The Journal of cell biology* **150** (1), 119-30 (2000).
- 76 T. Yao, L. Song, W. Xu et al., *Nat Cell Biol* **8** (9), 994-1002 (2006).
- 77 X. B. Qiu, S. Y. Ouyang, C. J. Li et al., *Embo J* **25** (24), 5742-53 (2006).
- 78 J. Hamazaki, S. Iemura, T. Natsume et al., *Embo J* **25** (19), 4524-36 (2006).
- 79 T. Li, N. I. Naqvi, H. Yang et al., *Biochem Biophys Res Commun* **272** (1), 270-5. (2000).
- 80 M. Stone, R. Hartmann-Petersen, M. Seeger et al., *Journal of molecular biology* **344** (3), 697-706 (2004).
- 81 J. You and C. M. Pickart, *The Journal of biological chemistry* **276** (23), 19871-8 (2001).
- 82 Y. Xie and A. Varshavsky, *Proc Natl Acad Sci U S A* **97** (6), 2497-502. (2000).

- 83 Y. Xie and A. Varshavsky, *Nature cell biology* **4** (12), 1003-7 (2002).
- 84 P. Tongaonkar, L. Chen, D. Lambertson et al., *Mol Cell Biol* **20** (13), 4691-8. (2000).
- 85 M. Funakoshi, X. Li, I. Velichutina et al., *Journal of cell science* **117** (Pt 26), 6447-54 (2004).
- 86 J. P. Jorgensen, A. M. Lauridsen, P. Kristensen et al., *J Mol Biol* **360** (5), 1043-52 (2006).
- 87 S. Shimada, M. Ogawa, M. Takahashi et al., *Cancer research* **54** (14), 3831-6 (1994).
- 88 A. Palmer, A. J. Rivett, S. Thomson et al., *Biochem J* **316** (Pt 2)(8 Suppl), 401-7. (1996).
- 89 P. Brooks, G. Fuertes, R. Z. Murray et al., *Biochem J* **346 Pt 1** (4), 155-61. (2000).
- 90 T. Ito, T. Chiba, R. Ozawa et al., *Proceedings of the National Academy of Sciences of the United States of America* **98** (8), 4569-74 (2001).
- 91 T. K. Gandhi, J. Zhong, S. Mathivanan et al., *Nature genetics* **38** (3), 285-93 (2006).
- 92 K. Hasegawa, N. Sakurai, and T. Kinoshita, *Development, growth & differentiation* **43** (1), 25-31 (2001).
- 93 A. B. Simins, H. Weighardt, K. M. Weidner et al., *Clinical & experimental metastasis* **17** (8), 641-8 (1999).
- 94 C. Pilarsky, M. Wenzig, T. Specht et al., *Neoplasia (New York, N.Y)* **6** (6), 744-50 (2004).
- 95 X. Pan, P. Ye, D. S. Yuan et al., *Cell* **124** (5), 1069-81 (2006).
- 96 K. M. Seong, J. H. Baek, M. H. Yu et al., *FEBS letters* **581** (13), 2567-73 (2007).
- 97 D. Kornitzer, B. Raboy, R. G. Kulka et al., *The EMBO journal* **13** (24), 6021-30 (1994).
- 98 S. Irniger and G. H. Braus, *Current genetics* **44** (1), 8-18 (2003).
- 99 K. Husnjak, S. Elsassner, N. Zhang et al., *Nature* **453** (7194), 481-8 (2008).
- 100 TR. Schneider and GM. Sheldrick, *Acta Crystallogr D* **58**, 1772-9. (2002).
- 101 K. D. Cowtan and P. Main, *Acta Crystallogr D Biol Crystallogr* **52** (Pt 1), 43-8 (1996).
- 102 A. Perrakis, RM. Morris, and VS. Lamzin, *Nature Struct. Biol.* **6**, 458-63 (1999).
- 103 D. Turk, *Thesis*, Technische Universitaet Muenchen (1992).
- 104 G. N. Murshudov, A. A. Vagin, and E. J. Dodson, *Acta Crystallogr D Biol Crystallogr* **53** (Pt 3), 240-55 (1997).
- 105 A. T. Brünger, *Acta Crystallogr D Biol Crystallogr* **49** (Pt 1), 24-36 (1993).
- 106 H. S. Yoon, P. J. Hajduk, A. M. Petros et al., *Nature* **369** (6482), 672-5 (1994).
- 107 B. J. Mayer, R. Ren, K. L. Clark et al., *Cell* **73** (4), 629-30 (1993).
- 108 R. J. Haslam, H. B. Koide, and B. A. Hemmings, *Nature* **363** (6427), 309-10 (1993).

- 109 M. J. Macias, A. Musacchio, H. Ponstingl et al., *Nature* **369** (6482), 675-7 (1994).
- 110 M. A. Lemmon, *Biochem Soc Trans* **32** (Pt 5), 707-11 (2004).
- 111 K. E. Prehoda, D. J. Lee, and W. A. Lim, *Cell* **97** (4), 471-80 (1999).
- 112 K. M. Ferguson, J. M. Kavran, V. G. Sankaran et al., *Mol Cell* **6** (2), 373-84 (2000).
- 113 V. Gervais, V. Lamour, A. Jawhari et al., *Nat Struct Mol Biol* **11** (7), 616-22 (2004).
- 114 I. R. Vetter, C. Nowak, T. Nishimoto et al., *Nature* **398** (6722), 39-46 (1999).
- 115 K. J. Walters, M. F. Kleijnen, A. M. Goh et al., *Biochemistry* **41** (6), 1767-77 (2002).
- 116 R. S. Kang, C. M. Daniels, S. A. Francis et al., *Cell* **113** (5), 621-30 (2003).
- 117 Q. Wang, A. M. Goh, P. M. Howley et al., *Biochemistry* **42** (46), 13529-35 (2003).
- 118 Q. Wang, P. Young, and K. J. Walters, *J Mol Biol* **348** (3), 727-39 (2005).
- 119 V. Chau, J. W. Tobias, A. Bachmair et al., *Science (New York, N.Y)* **243** (4898), 1576-83 (1989).
- 120 D. Finley, S. Sadis, B. P. Monia et al., *Molecular and cellular biology* **14** (8), 5501-9 (1994).
- 121 C. Dominguez, R. Boelens, and A. M. Bonvin, *J Am Chem Soc* **125** (7), 1731-7 (2003).
- 122 AT. Brünger, PD. Adams, GM. Clore et al., *Acta Crystallogr D Biol Crystallogr.* **1** (54), 905-21 (1998).
- 123 G. Cornilescu, J.L. Marquardt, M. Ottiger et al., *J.Am.Chem.Soc.* **120**, 6836-7 (1998).
- 124 I. K. McDonald and J. M. Thornton, *J Mol Biol* **238** (5), 777-93 (1994).
- 125 P. Young, Q. Deveraux, R. E. Beal et al., *J Biol Chem* **273** (10), 5461-7 (1998).
- 126 H. Teo, D. J. Gill, J. Sun et al., *Cell* **125** (1), 99-111 (2006).
- 127 T. Slagsvold, R. Aasland, S. Hirano et al., *J Biol Chem* **280** (20), 19600-6 (2005).
- 128 K. Fujiwara, T. Tenno, K. Sugawara et al., *J Biol Chem* **279** (6), 4760-7 (2004).
- 129 T. D. Mueller and J. Feigon, *Journal of molecular biology* **319** (5), 1243-55 (2002).
- 130 K. J. Walters, P. J. Lech, A. M. Goh et al., *Proceedings of the National Academy of Sciences of the United States of America* **100** (22), 12694-9 (2003).
- 131 A. Ohno, J. Jee, K. Fujiwara et al., *Structure* **13** (4), 521-32 (2005).
- 132 R. Varadan, M. Assfalg, S. Raasi et al., *Molecular cell* **18** (6), 687-98 (2005).
- 133 J. H. Hurley, S. Lee, and G. Prag, *Biochem J* **399** (3), 361-72 (2006).
- 134 R. D. Fisher, B. Wang, S. L. Alam et al., *The Journal of biological chemistry* **278** (31), 28976-84 (2003).
- 135 K. A. Swanson, R. S. Kang, S. D. Stamenova et al., *The EMBO journal* **22** (18), 4597-606 (2003).
- 136 S. Lee, Y. C. Tsai, R. Mattera et al., *Nature structural & molecular biology* **13** (3), 264-71 (2006).

- 137 L. Penengo, M. Mapelli, A. G. Murachelli et al., *Cell* **124** (6), 1183-95 (2006).
- 138 M. G. Bomar, M. T. Pai, S. R. Tzeng et al., *EMBO reports* **8** (3), 247-51 (2007).
- 139 S. L. Alam, C. Langelier, F. G. Whitby et al., *Nat Struct Mol Biol* **13** (11), 1029-30 (2006).
- 140 S. Hirano, N. Suzuki, T. Slagsvold et al., *Nat Struct Mol Biol* **13** (11), 1031-2 (2006).
- 141 A. Kohl, H. K. Binz, P. Forrer et al., *Proc Natl Acad Sci U S A* **100** (4), 1700-5 (2003).
- 142 K. Narayan and M. A. Lemmon, *Methods (San Diego, Calif)* **39** (2), 122-33 (2006).
- 143 J. Hanna, N. A. Hathaway, Y. Tone et al., *Cell* **127** (1), 99-111 (2006).
- 144 B. Crosas, J. Hanna, D. S. Kirkpatrick et al., *Cell* **127** (7), 1401-13 (2006).
- 145 A. Stanhill, C. M. Haynes, Y. Zhang et al., *Molecular cell* **23** (6), 875-85 (2006).
- 146 D. Reverter and C. D. Lima, *Nature* **435** (7042), 687-92 (2005).
- 147 N. Lamerant and C. Kieda, *Febs J* **272** (8), 1833-44 (2005).
- 148 N. Budisa, B. Steipe, P. Demange et al., *Eur J Biochem* **230** (2), 788-96 (1995).
- 149 J. Kyhse-Andersen, *Journal of biochemical and biophysical methods* **10** (3-4), 203-9 (1984).
- 150 H. Towbin, T. Staehelin, and J. Gordon, *Proceedings of the National Academy of Sciences of the United States of America* **76** (9), 4350-4 (1979).
- 151 U. K. Laemmli, *Nature* **227** (5259), 680-5 (1970).
- 152 M. M. Bradford, *Analytical biochemistry* **72**, 248-54 (1976).
- 153 Z. Otwinowski, W. Minor, and Volume F International Tables for Crystallography, *International Tables for Crystallography, Volume F*.
- 154 F. Delaglio, S. Grzesiek, G. W. Vuister et al., *J Biomol NMR* **6** (3), 277-93 (1995).
- 155 C. Bartels, T.-H. Xia, M. Billeter et al., *J Biomol NMR* **6**, 1-10 (1995).
- 156 K. Henze and W. Martin, *Nature* **426** (6963), 127-8 (2003).
- 157 H. M. McBride, M. Neuspiel, and S. Wasiak, *Curr Biol* **16** (14), R551-60 (2006).
- 158 A. S. Reichert and W. Neupert, *Trends Genet* **20** (11), 555-62 (2004).
- 159 M. W. Gray, G. Burger, and B. F. Lang, *Science* **283** (5407), 1476-81. (1999).
- 160 P. Dolezal, V. Likic, J. Tachezy et al., *Science* **313** (5785), 314-8 (2006).
- 161 D. C. Chan, *Cell* **125** (7), 1241-52 (2006).
- 162 R. J. Wiesner, J. C. Ruegg, and I. Morano, *Biochem Biophys Res Commun* **183** (2), 553-9 (1992).
- 163 B. Alberts, A. Johnson, J. Lewis et al., *Molecular Biology of the Cell*. (Garland Publishing Inc., New York, 1994).
- 164 D. Voet, J. G. Voet, and C. W. Pratt, *Fundamentals of Biochemistry*, 2nd Edition ed. (John Wiley and Sons, Inc., 547, 2006).
- 165 D. C. Chan, *Annu Rev Cell Dev Biol* **22**, 79-99 (2006).

- 166 C. A. Mannella, *Biochim Biophys Acta* **1763** (5-6), 542-8 (2006).
- 167 C. A. Mannella, *Biochim Biophys Acta* **1762** (2), 140-7 (2006).
- 168 J. B. McMillin and W. Dowhan, *Biochim Biophys Acta* **1585** (2-3), 97-107 (2002).
- 169 D. C. Wallace, *Science* **283** (5407), 1482-8 (1999).
- 170 D. Mokranjac and W. Neupert, *Biochem Soc Trans* **33** (Pt 5), 1019-23 (2005).
- 171 W. Neupert and J. M. Herrmann, *Annu Rev Biochem* **76**, 723-49 (2007).
- 172 N. Bolender, A. Sickmann, R. Wagner et al., *EMBO Rep* **9** (1), 42-9 (2008).
- 173 K. N. Truscott, P. Kovermann, A. Geissler et al., *Nat Struct Biol* **8** (12), 1074-82 (2001).
- 174 A. Chacinska, M. Lind, A. E. Frazier et al., *Cell* **120** (6), 817-29 (2005).
- 175 A. Geissler, A. Chacinska, K. N. Truscott et al., *Cell* **111** (4), 507-18 (2002).
- 176 H. Yamamoto, M. Esaki, T. Kanamori et al., *Cell* **111** (4), 519-28 (2002).
- 177 D. Mokranjac, S. A. Paschen, C. Kozany et al., *Embo J* **22** (4), 816-25 (2003).
- 178 M. Meinecke, R. Wagner, P. Kovermann et al., *Science* **312** (5779), 1523-6 (2006).
- 179 A. Geissler, J. Rassow, N. Pfanner et al., *Mol Cell Biol* **21** (20), 7097-104 (2001).
- 180 K. Okamoto, A. Brinker, S. A. Paschen et al., *Embo J* **21** (14), 3659-71 (2002).
- 181 K. Shariff, S. Ghosal, and A. Matouschek, *Biophys J* **86** (6), 3647-52 (2004).
- 182 P. R. D'Silva, B. Schilke, W. Walter et al., *Proc Natl Acad Sci U S A* **102** (35), 12419-24 (2005).
- 183 T. Sato, M. Esaki, J. M. Fernandez et al., *Proc Natl Acad Sci U S A* **102** (50), 17999-8004 (2005).
- 184 M. Krayl, J. H. Lim, F. Martin et al., *Mol Cell Biol* **27** (2), 411-25 (2007).
- 185 Y. Li, J. Dudek, B. Guiard et al., *J Biol Chem* (2004).
- 186 D. Mokranjac, G. Bourenkov, K. Hell et al., *Embo J* **25** (19), 4675-85 (2006).
- 187 A. P. van Loon, A. W. Brandli, and G. Schatz, *Cell* **44** (5), 801-12. (1986).
- 188 N. Wiedemann, N. Pfanner, and M. T. Ryan, *EMBO J.* **20** (5), 951-60. (2001).
- 189 S. P. Curran, D. Leuenberger, E. Schmidt et al., *J Cell Biol* **158** (6), 1017-27 (2002).
- 190 S. P. Curran, D. Leuenberger, W. Oppliger et al., *Embo J* **21** (5), 942-53 (2002).
- 191 C. Sirrenberg, M. Endres, H. Folsch et al., *Nature* **391** (6670), 912-5. (1998).
- 192 K. N. Truscott, N. Wiedemann, P. Rehling et al., *Mol Cell Biol* **22** (22), 7780-9 (2002).
- 193 S. Vial, H. Lu, S. Allen et al., *J Biol Chem* **277** (39), 36100-8 (2002).
- 194 C. T. Webb, M. A. Gorman, M. Lazarou et al., *Mol Cell* **21** (1), 123-33 (2006).
- 195 P. Rehling, K. Model, K. Brandner et al., *Science* **299** (5613), 1747-51 (2003).
- 196 A. Chacinska, S. Pfannschmidt, N. Wiedemann et al., *Embo J* (2004).
- 197 N. Terziyska, T. Lutz, C. Kozany et al., *FEBS Lett* **579** (1), 179-84 (2005).

- 198 M. Naoe, Y. Ohwa, D. Ishikawa et al., *J Biol Chem* (2004).
- 199 D. Milenkovic, K. Gabriel, B. Guiard et al., *J Biol Chem* **282** (31), 22472-80 (2007).
- 200 D. P. Sideris and K. Tokatlidis, *Mol Microbiol* **65** (5), 1360-73 (2007).
- 201 N. Mesecke, N. Terziyska, C. Kozany et al., *Cell* **121** (7), 1059-69 (2005).
- 202 S. Allen, V. Balabanidou, D. P. Sideris et al., *J Mol Biol* **353** (5), 937-44 (2005).
- 203 M. Rissler, N. Wiedemann, S. Pfannschmidt et al., *J Mol Biol* **353** (3), 485-92 (2005).
- 204 K. Diekert, A. I. de Kroon, U. Ahting et al., *Embo J* **20** (20), 5626-35 (2001).
- 205 M. E. Dumont, J. F. Ernst, and F. Sherman, *J. Biol. Chem.* **263**, 15928-37 (1988).
- 206 F. E. Nargang, M. E. Drygas, P. L. Kwong et al., *J. Biol. Chem.* **263** (19), 9388-94. (1988).
- 207 E. Schleiff and J. Soll, *EMBO Rep* **6** (11), 1023-7 (2005).
- 208 S. A. Paschen, T. Waizenegger, T. Stan et al., *Nature* **426** (6968), 862-6 (2003).
- 209 N. Wiedemann, V. Kozjak, A. Chacinska et al., *Nature* **424** (6948), 565-71 (2003).
- 210 S. C. Hoppins and F. E. Nargang, *J Biol Chem* **279** (13), 12396-405 (2004).
- 211 N. Wiedemann, K. N. Truscott, S. Pfannschmidt et al., *J Biol Chem* **279** (18), 18188-94 (2004).
- 212 V. Kozjak, N. Wiedemann, D. Milenkovic et al., *J Biol Chem* **278** (49), 48520-3 (2003).
- 213 C. Meisinger, S. Pfannschmidt, M. Rissler et al., *Embo J* **26** (9), 2229-39 (2007).
- 214 L. Sanchez-Pulido, D. Devos, S. Genevrois et al., *Trends Biochem Sci* **28** (10), 523-6 (2003).
- 215 S. J. Habib, T. Waizenegger, A. Niewianda et al., *J Cell Biol* **176** (1), 77-88 (2007).
- 216 Y. Wu and B. Sha, *Nat Struct Mol Biol* **13** (7), 589-93 (2006).
- 217 N. C. Chan, V. A. Likic, R. F. Waller et al., *J Mol Biol* **358** (4), 1010-22 (2006).
- 218 Y. Abe, T. Shodai, T. Muto et al., *Cell* **100** (5), 551-60 (2000).
- 219 K. P. Kunkele, P. Juin, C. Pompa et al., *J. Biol. Chem.* **273** (47), 31032-9. (1998).
- 220 K. Hill, K. Model, M. T. Ryan et al., *Nature* **395** (6701), 516-21. (1998).
- 221 L. Becker, M. Bannwarth, C. Meisinger et al., *J Mol Biol* **353** (5), 1011-20 (2005).
- 222 C. Meisinger, M. T. Ryan, K. Hill et al., *Mol. Cell. Biol.* **21** (7), 2337-48. (2001).
- 223 P. Rehling, K. Brandner, and N. Pfanner, *Nat Rev Mol Cell Biol* **5** (7), 519-30 (2004).
- 224 C. M. Koehler, *Annu Rev Cell Dev Biol* **20**, 309-35 (2004).
- 225 S. van Wilpe, M. T. Ryan, K. Hill et al., *Nature* **401** (6752), 485-9. (1999).
- 226 A. Mayer, F. E. Nargang, W. Neupert et al., *Embo J* **14** (17), 4204-11. (1995).
- 227 P. J. Dekker, M. T. Ryan, J. Brix et al., *Mol. Cell. Biol.* **18** (11), 6515-24. (1998).
- 228 K. Dietmeier, A. Honlinger, U. Bomer et al., *Nature* **388** (6638), 195-200. (1997).
- 229 E. L. Sherman, N. E. Go, and F. E. Nargang, *Mol Biol Cell* **16** (9), 4172-82 (2005).

- 230 K. P. Kunkele, S. Heins, M. Dembowski et al., *Cell* **93** (6), 1009-19. (1998).
- 231 U. Ahting, C. Thun, R. Hegerl et al., *J Cell Biol* **147** (5), 959-68 (1999).
- 232 K. Model, T. Prinz, T. Ruiz et al., *J Mol Biol* **316** (3), 657-66 (2002).
- 233 U. Ahting, M. Thieffry, H. Engelhardt et al., *J Cell Biol* **153** (6), 1151-60. (2001).
- 234 D. Rapaport, *J Cell Biol* **171** (3), 419-23 (2005).
- 235 W. Sebald, W. Neupert, and H. Weiss, *Methods Enzymol* **55** (2), 144-8 (1979).
- 236 C. Hunte and H. Michel, *Curr Opin Struct Biol* **12** (4), 503-8 (2002).
- 237 G. Kleymann, C. Ostermeier, B. Ludwig et al., *Bio/technology (Nature Publishing Company)* **13** (2), 155-60 (1995).
- 238 S. Schmitt, *Diploma thesis* (2005).
- 239 E. Screpanti, E. Padan, A. Rimon et al., *J Mol Biol* **362** (2), 192-202 (2006).
- 240 Z. Otwinowski and W. Minor, *Methods in Enzymology* **276**, 307-26 (1997).
- 241 M. Bannwarth and G. E. Schulz, *Biochim Biophys Acta* **1610** (1), 37-45 (2003).
- 242 M. C. Wiener, *Methods* **34** (3), 364-72 (2004).
- 243 G. G. Prive, *Methods* **41** (4), 388-97 (2007).
- 244 J. P. Allen and G. Feher, *Proc Natl Acad Sci U S A* **81** (15), 4795-9 (1984).
- 245 T. Tsukihara, H. Aoyama, E. Yamashita et al., *Science* **272** (5265), 1136-44 (1996).
- 246 V. Koronakis, A. Sharff, E. Koronakis et al., *Nature* **405** (6789), 914-9 (2000).
- 247 P. W. Holloway, *Anal Biochem* **53** (1), 304-8 (1973).
- 248 H. S. Young, J. L. Rigaud, J. J. Lacapere et al., *Biophys J* **72** (6), 2545-58 (1997).
- 249 A. D. Ferguson, W. Welte, E. Hofmann et al., *Structure* **8** (6), 585-92 (2000).
- 250 M. R. Jones, *Progress in lipid research* **46** (1), 56-87 (2007).
- 251 R. M. Garavito and S. Ferguson-Miller, *J Biol Chem* **276** (35), 32403-6 (2001).
- 252 M. J. Newman, D. L. Foster, T. H. Wilson et al., *J Biol Chem* **256** (22), 11804-8 (1981).
- 253 L. Guan, I. N. Smirnova, G. Verner et al., *Proc Natl Acad Sci U S A* **103** (6), 1723-6 (2006).
- 254 M. J. Lemieux, J. Song, M. J. Kim et al., *Protein Sci* **12** (12), 2748-56 (2003).
- 255 H. Zhang, G. Kurisu, J. L. Smith et al., *Proc Natl Acad Sci U S A* **100** (9), 5160-3 (2003).
- 256 P. Gast, P. Hemelrijk, and A. J. Hoff, *FEBS Lett* **337** (1), 39-42 (1994).
- 257 H. Michel, *Trends Biochem Sci* **8**, 56-9 (1983).
- 258 B. S. Smith, B. Kobe, R. Kurumbail et al., *Acta crystallographica* **54** (Pt 4), 697-9 (1998).
- 259 H. Sui, B. G. Han, J. K. Lee et al., *Nature* **414** (6866), 872-8 (2001).
- 260 G. Chang, R. H. Spencer, A. T. Lee et al., *Science* **282** (5397), 2220-6 (1998).

- 261 C. Ostermeier and H. Michel, *Curr Opin Struct Biol* **7** (5), 697-701 (1997).
- 262 M. Roth, A. Lewit-Bentley, H. Michel et al., *Nature* **340**, 659-62 (1989).
- 263 C. Ostermeier, S. Iwata, B. Ludwig et al., *Nat Struct Biol* **2** (10), 842-6 (1995).
- 264 C. Hunte, J. Koepke, C. Lange et al., *Structure* **8** (6), 669-84 (2000).
- 265 Y. Zhou, J. H. Morais-Cabral, A. Kaufman et al., *Nature* **414** (6859), 43-8 (2001).
- 266 C. Hunte and A. Kannt, *Membrane Protein Purification and Crystallization: A Practical Guide*, 14 (2003).
- 267 A. M. Lesk and C. Chothia, *Nature* **335** (6186), 188-90 (1988).
- 268 S. Iwata, C. Ostermeier, B. Ludwig et al., *Nature* **376** (6542), 660-9 (1995).
- 269 R. Dutzler, E. B. Campbell, and R. MacKinnon, *Science* **300** (5616), 108-12 (2003).
- 270 Y. Jiang, A. Lee, J. Chen et al., *Nature* **423** (6935), 33-41 (2003).
- 271 M. Venturi and C. Hunte, *Biochim Biophys Acta* **1610** (1), 46-50 (2003).
- 272 C. Hunte, E. Screpanti, M. Venturi et al., *Nature* **435** (7046), 1197-202 (2005).
- 273 P. Forrer, M. T. Stumpp, H. K. Binz et al., *FEBS Lett* **539** (1-3), 2-6 (2003).
- 274 H. K. Binz, P. Amstutz, A. Kohl et al., *Nature biotechnology* **22** (5), 575-82 (2004).
- 275 J. Deisenhofer, O. Epp, K. Miki et al., *Nature* **318** (6047), 618-24 (1985).
- 276 D. A. Doyle, J. Morais Cabral, R. A. Pfuetzner et al., *Science* **280** (5360), 69-77 (1998).
- 277 V. Koronakis, J. Li, E. Koronakis et al., *Mol Microbiol* **23** (3), 617-26 (1997).
- 278 N. Dekker, K. Merck, J. Tommassen et al., *Eur J Biochem* **232** (1), 214-9 (1995).
- 279 M. Blaauw, N. Dekker, H. M. Verheij et al., *FEBS Lett* **373** (1), 10-2 (1995).
- 280 G. E. Schulz, *Curr Opin Struct Biol* **10** (4), 443-7 (2000).
- 281 C. A. Mannella, *J Bioenerg Biomembr* **29** (6), 525-31 (1997).
- 282 C. A. Mannella, *Journal of structural biology* **121** (2), 207-18 (1998).
- 283 C. Binda, P. Newton-Vinson, F. Hubalek et al., *Nat Struct Biol* **9** (1), 22-6 (2002).
- 284 G. Kohler and C. Milstein, *Nature* **256** (5517), 495-7 (1975).
- 285 E. Harlow and D. Lane, *Antibodies. A Laboratory Manual*. (Cold Spring Harbour Press, New York, 1988).
- 286 U. K. Laemmli, *Nature* **227**, 680-5 (1970).

# Lebenslauf

<b>Name</b>	Patrick Schreiner
<b>Geburtsdatum</b>	16.06.1978
<b>Geburtsort</b>	Höchstädt an der Donau
<b>Staatsangehörigkeit</b>	deutsch
<b>Familienstand</b>	verheiratet
<b>Anschrift</b>	Solnhofener Straße 2, 81249 München Tel.: 089 / 89 33 73 59 E-Mail: patschrein@gmx.de
<b>Promotion</b>	
09/2004 — heute	Promotion am Adolf-Butenandt-Institut für Physiologische Chemie, München Tätigkeiten: wissenschaftliche Arbeit im Bereich Molekularbiologie, Proteinbiochemie, Strukturbiologie Verwaltung der Äkta-Chromatographie-Systeme und der Kristallisation, Betreuung von Studenten, Präsentationen auf internationalen Kongressen Titel der Dissertation: „Structural investigation of the multifunctional proteasomal adaptor protein Rpn13 and of mitochondrial protein translocase TOM core complex”
08/2006 — heute	Anleitung und Betreuung einer technischen Assistentin
01/2005 — 06/2005	wissenschaftlicher Mitarbeiter im Max-Planck-Institut für Biophysik, Frankfurt, im Zuge einer Kooperation Tätigkeiten: Arbeiten in der Zellkultur, Herstellung monoklonaler Antikörper

## **Ausbildung, Wehrdienst, Praktikum, Schule**

09/2000 – 12/2003	Ludwig-Maximilians-Universität München, Hauptstudium in Diplom-Chemie, Schwerpunkt: Biochemie
07/2003 – 12/2003	Diplomarbeit im Bereich Proteinkristallographie in der Arbeitsgruppe von Prof. Dr. Patrick Cramer, Genzentrum, München Titel der Diplomarbeit: “Domain architecture of general transcription factor IIE“
19.12.2003	Diplom-Chemiker
10/1998 – 08/2000	Universität Ulm, Grundstudium in Diplom-Chemie
09/1997 – 09/1998	Zivildienst am Regens-Wagner-Institut in Dillingen an der Donau
09/1988 – 06/1997	Johann-Michael-Sailer-Gymnasium, Dillingen an der Donau
27.06.1997	Abschluss: Abitur

## **Publikationen und Präsentationen**

**P. Schreiner**, X. Chen, K. Husnjak et al., Ubiquitin docking at the proteasome via a novel PH domain interaction; *Nature* **453** (7194), 548-52 (2008)

Ott M., Norberg E., Walter K.M., **Schreiner P.**, Kemper C., Rapaport D., Zhivotovsky B., Orrenius S. The mitochondrial TOM complex is required for tBid/Bax-induced cytochrome c release. *Journal of Biological Chemistry* 2007 Sep 21;282(38):27633-9.

**Schreiner P.**, Schmitt S., Hunte C., Neupert W., Groll M. Crystallization of the TOM core complex from *Neurospora crassa* together with monoclonal antibodies. Poster-Präsentation auf den Kongressen “XX Congress of the IUCR” 2005 in Florenz und “38th Crystallographic Course” 2006 in Erice.

München, 17. Juli 2008

