Novel export and import pathways in *S. cerevisiae* identified by an engineered SUMO system

Dissertation

for the award of the degree

'Doctor rerum naturalium'

within the Molecular Biology Program

of the Georg-August-Universität Göttingen

Faculty of Biology

Submitted by

Arturo Vera Rodríguez

Born in

Mexico City, Mexico

Göttingen, April 2017

Members of the Thesis Committee

Prof. Dr. Dirk Görlich (Supervisor and referee)	Max Planck Institute for Biophysical Chemistry Department of Cellular Logistics Göttingen, Germany
Prof. Dr Marina Rodnina (Co-referee)	University Medical Center Göttingen Department of Cellular Biochemistry Göttingen, Germany
Prof. Dr. Henning Urlaub	Max Planck Institute for Biophysical Chemistry Bioanalytical Mass Spectrometry Göttingen, Germany

Additional members of the Examination Board

Prof. Halyna Shcherbata	Max Planck Institute for Biophysical Chemistry Gene Expression and Signaling Göttingen, Germany
Prof. Dr. Roland Dosch	University Medical Center Göttingen Department of Developmental Biochemistry Göttingen, Germany
Prof. Dr. Dieter Klopfenstein	University Medical Center Göttingen Department of Biophysics Göttingen, Germany

Date of oral examination: 26.06.2017

I hereby declare that I completed my thesis entitled 'Novel export and import pathways in *S. cerevisiae* identified by an engineered SUMO system' independently and with no other aids or sources than quoted.

This dissertation has not been submitted elsewhere for any academic award or qualification.

Arturo Vera Rodríguez

Göttingen, April 2017

"Just 'cause you feel it, doesn't mean it's there"
Radiohead

1 ABSTRACT

Nuclear transport receptors (NTRs) mediate the translocation of cargos between the cytoplasm and the nucleus. NTRs are classified either as importins or exportins according to the directionality of the cargo transport. To identify new interaction partners of so far poorly characterized NTRs from *S. cerevisiae*, we developed a novel protein purification approach based on an engineered SUMO-specific protease (SUMOvera protease) and its specific SUMO substrate (SUMOvera). SUMOvera can be used as a stable fusion tag in virtually all eukaryotic systems since it is not recognized by SUMO-specific proteases from yeast, plant, human, amphibian, insect and protozoa. In addition, the SUMOvera protease can be over-expressed in *S. cerevisiae* without causing toxicity unlike yeast and *B. distachyon* SUMO-specific proteases. The SUMOvera system has also an orthogonal specificity to the SUMOstar/SUMOstar protease pair to allow the purification of protein complexes with a defined stoichiometry in eukaryotic hosts.

Using the SUMOvera system, we discovered that yeast Lph2 and Pdr6 mediate the nuclear export and import of different cargos. Specifically, we showed that Lph2 exports and binds to the translation initiation factor eIF4A in a Ran-GTP-dependent manner, and that Pdr6 recognizes Ubc9 as an import substrate as well as eEF2 and eIF5A as export cargos. Lph2 and Pdr6 have been only described as importins; however, our data indicate that they act as bidirectional NTRs that shuttle distinct sets of cargos in opposite directions through the nuclear envelope.

Overall, we describe two novel bidirectional NTRs in *S. cerevisiae* in addition to Msn5. The reported findings suggest that bidirectionality in NTRs might be more common than previously assumed, and that there might be other bidirectional NTRs in different organisms that still need to be identified.

2 INTRODUCTION

2.1   Nucleus: the hallmark of eukaryotic cells

Prokaryotic cells have the simplest cellular 'architecture' as they are only defined by a cell membrane. Eukaryotic cells evolved from prokaryotes after acquiring defined cellular compartments enclosed by different membranes. The nucleus, the most defining and prominent of these compartments (Cavalier-Smith 1988), is surrounded by the nuclear envelope (NE) that separates the genetic material (genomic DNA) from the cytoplasm of a cell. The appearance of the nucleus allowed eukaryotic cells to acquire several advantages. First, the nucleus protects the genomic DNA from physical and chemical damage originated in the cytoplasm. Second, separation of nuclear transcription from translation in the cytoplasm allows eukaryotes to control gene expression in a define time and space (Gant & Wilson 1997). Third, the spatial uncoupling of transcription from translation allows cells to expand the coding potential of their genome by removing a different number of introns from pre-mRNAs by alternative splicing. Last, the nucleus prevents the translation of intron-containing mRNAs that might cause the productions of aberrant proteins by preventing the free diffusion of specific macromolecules needed for translation (*i.e.* transcription factors, tRNA and ribosomes).

Despite the fact all benefits provided by the NE in eukaryotic cells, the physical separation between the nucleus and the cytoplasm require a fine-tuned communication between the two compartments. For instance, transcription in the nucleus requires the input of proteins (*i.e.* transcription factors, RNA polymerase, mRNA biding proteins) that are synthesized in the cytoplasm. In contrast, translation relies on the supply of nuclear products such as mRNA, tRNA and ribosomal subunits for correct protein synthesis. Therefore, the constant and controlled exchange of macromolecules between the cytoplasm and the nucleus is needed to achieve complete normal cellular growth.

2.2   Overviews of the nucleocytoplasmic transport

The nuclear pore complex (NPC) is the sole gate between the nucleus and the cytoplasm, and controls the exchange of macromolecules across the NE. NPCs are large protein assemblies embedded in the NE (Watson 1959) with a molecular weight of 66 MDa (Rout & Blobel 1993) in yeast and around 125 MDa in vertebrates (Reichelt et al. 1990). NPCs

are composed of around 30 different proteins named nucleoporins (Nups), which can be divided into different classes based on their function within the NPC. On one hand, Nups of a complete defined protein structure form the eight-fold symmetric scaffold ring of the NPC (von Appen et al. 2015; Eibauer et al. 2015; Kelley et al. 2015). On the other hand, Nups containing intrinsically disordered phenylalanine-glycine domains fill the central NPC channel and form the selective permeability barrier of the NPC (reviewed in Schmidt & Görlich 2016).

Small molecules (*i.e.* nucleotides, glycerol, H_2O, glucose) and globular protein of less than 20 kDa easily cross the permeability barrier at the NPC by passive diffusion. As the size and shape of the molecules increases ($\geq$ 5 nm in diameter or 20-40 kDa), passive diffusion becomes restricted (Bonner 1975; Paine et al. 1975; Mohr et al. 2009). In this case, the active transport of large macromolecules relies on specialized proteins known as nuclear transport receptors (NTRs) (reviewed in Sloan et al. 2016; Christie et al. 2016; Matsuura 2016; Güttler & Görlich 2011).

2.3 Transport directionality and nuclear transport receptors (NTRs)

Macromolecules that do not passively diffuse through the NE due to their large size and net charge require to be transported in a facilitated manner by nuclear transport receptors (NTRs). NTRs are globular proteins of large size (90-150 kDa) that are able to shuttle continuously between the cytoplasm and the nucleus. NTRs bind to their cognate cargo in one side of the NE and transport it to the other side by interacting with the multiple FG repeats at the central channel of the NPC. The multiple interactions between the NTR and the FG meshwork at the NPC allow NTRs to transport cargos by overcoming the size limit imposed by the permeability barrier (Ribbeck & Görlich 2002; Frey & Görlich 2007; Schmidt & Görlich 2016). The active transport of large molecules is extremely efficient as a single NTR can translocate up to 1000 molecules per NPC per second (Ribbeck & Görlich 2001).

Most of the facilitated cargo transport through the NPC is mediated by NTRs of the importin β (Imp-β) superfamily (also known as β-karyopherins) which comprises 21 different members in humans and 14 in yeast (reviewed in Güttler & Görlich 2011). Although all NTRs from the importin β-like superfamily share only a modest protein sequence similarity (from 10% to 20%) (Quan et al. 2008), they all have a common

building secondary structure known as HEAT (reviewed in Güttler & Görlich 2011). As determined by different crystallographic studies, Imp-β NTRs have an α-solenoid fold that is composed by a tandem of 18-20 HEAT repeats interconnected by short loops (Vetter et al. 1999; Monecke et al. 2009; Okada et al. 2009; Aksu et al. 2016). NTRs undergo to different protein conformations between the unbound and cargo-bound (reviewed in Claustre & Maritorena 2003; Schwartz 2016). In addition, NTRs that form the Imp-β family share the ability to bind different NPC domain(s) as well as the capability to bind the small Ras-related nuclear protein GTPase (Ran) via the N-terminal region of the NTR (Rexach & Blobel 1995; Görlich et al. 1996; Görlich et al. 1997).

Ran (Gsp1 in *S. cerevisiae*) is a 25 kDa protein that switches between the GTP-bound (Ran-GTP) and GDP-bound (Ran-GDP) states (Bischoff & Ponstingl 1991). Ran is one of the key regulators of the nucleocytoplasmic transport and its function is best explained by the Ran-GTP gradient model (Görlich et al. 1996; Izaurralde et al. 1997). Ran-GTP (active form) is found predominantly in the nucleus, while Ran-GDP (inactive form) is highly enriched in the cytoplasm. The predominant subcellular localization of Ran-GTP and Ran-GDP is maintained due to the strict localization of the Ran activating protein (RanGAP1 in vertebrates, Rna1 in yeast) and the Ran guanine nucleotide exchange factor (RCC1 in vertebrates, Prp20 in yeast). RanGAP1 is found predominantly in the cytoplasm and enhances the GTPase activity of Ran that results in the increment of Ran-GDP levels in the cytoplasm (Hopper et al. 1990; Matunis et al. 1996; Mahajan et al. 1997). In contrast, RCC1 localizes only in the nucleus and facilities the exchange of GDP to GTP on Ran (Bischoff & Ponstingl 1991; Klebe et al. 1995).

Ran-GTP represents the active form of Ran as it has a higher affinity to NTRs compared to Ran-GDP (Nilsson et al. 2002; Görlich et al. 1996). Moreover, Ran-GTP also changes the conformation of NTRs upon binding. Therefore, the drastic difference in Ran-GTP concentration between the cytoplasm and the nucleus (up to 1000-fold) is known to be the main cause of the directionality during transport (Görlich et al. 1997; Görlich et al. 2003; Izaurralde et al. 1997; Görlich et al. 1996). Based on the transport directionality, NTRs are classified as importins or exportins (Kutay et al. 1997; Fornerod et al. 1997). Exportins recognize their cognate cargos in the nucleus where the concentration of Ran-GTP is high, and transport them to the cytoplasm. On the other hand, importins form a stable complex with their respective cargos in the absence of Ran-GTP in the cytoplasm, and then shuttles to the nucleus. Additionally, several NTRs can act as imports as well as exportins,

shuttling different set of cargos in both direction (Yoshida & Blobel 2001; Mingot et al. 2001; Gontan et al. 2009; Kurisaki et al. 2006; Lipowsky et al. 2000). Standard and bidirectional NTRs are described in detail in the following sections.

Table 2.1 Cargos recognized by importins in higher eukaryotes

Vertebrate Importins	Cargo(s)	Reference(s)
Importin β	Import cargos with a cNLS via importin α, histone 1 via importin 7, m₃G-capped U-snRNPs, ribosomal proteins, HIV rev, HIV tat	(Görlich et al. 1995; Jäkel et al. 1999; Huber et al. 1998; Jäkel & Görlich 1998; Truant & Cullen 1999)
Transportin 1+2 (Trn, Importin β2)	hnRNP, ribosomal proteins, TAP/NFX1, histones, c-FOS	(Pollard et al. 1996; Jäkel & Görlich 1998; Truant & Cullen 1999; Mühlhäusser et al. 2001; Arnold et al. 2006)
Transportin SR1+2 (TrnSR, Trn3)	SR proteins	(Kataoka et al. 1999)
Importin 4	Ribosomal proteins, histones	(Jäkel et al. 2002; Mosammaparast et al. 2001; Mühlhäusser et al. 2001)
Importin 5	Ribosomal proteins, histones	(Jäkel & Görlich 1998)
Importin 7	Ribosomal proteins, ERK2, SMAD3, MEK1	(Jäkel & Görlich 1998; Chuderland et al. 2008)
Importin 8	Argonaute proteins, SRP19	(Weinmann et al. 2009; Dean et al. 2001)
Importin 9	Ribosomal proteins, histones	(Jäkel et al. 2002; Mosammaparast et al. 2001; Mühlhäusser et al. 2001)
Importin 11	UbcM2, rpL12	(Plafker et al. 2000; Plafker & Macara 2002)

cNLS= classical nuclear localization signal

2.4   Nuclear import receptors (importins)

In vertebrates, there are 21 different NTRs from the Impβ superfamily where only 9 members act as standard importins (Table 2.1). In the case of *S. cerevisiae*, 10 different importins are identified out of 14 members of the Impβ superfamily (Table 2.2). Importins recognize their cognate cargos in the cytoplasm and translocate to the other side of the NE as a single unit (import complex). After translocation, Ran-GTP binds to the importin in order to trigger the dissociation of the importin complex in the nucleoplasm. The disassembly of the import complex results in the release of the cargo inside of the nucleus while the importin remains bound to Ran-GTP. Next, the remaining importin•Ran-GTP complex is transported back to the cytoplasm and further dissembled upon GTP hydrolysis promoted by the cytosolic protein RanGAP1 (Hopper et al. 1990; Matunis et al. 1996;

Mahajan et al. 1997). Finally, free importin is able to start the transport of a new cargo to the cytoplasm (Figure 2.1).

Table 2.2 Cargos recognized by *S. cerevisiae* importins

S. cerevisiae Importins	Cargo(s)	Vertebrate NTR homologue(s)	Reference(s)
Kap95	Import cargos with a cNLS via Kap60 (Srp1)	Importin β1	(Enenkel et al. 1995)
Kap104	Nab2, Hrp1 (mRNA binding proteins)	Transportin 1+2	(Kessler et al. 1997; Aitchison et al. 1996)
Yrb4 (Kap123)	SRP proteins, histones, ribosomal proteins	Importin 4	(Schlenstedt et al. 1997; Rout et al. 1997; Grosshans et al. 2001; Greiner et al. 2004)
Mtr10 (Kap111)	Npl3, tRNAs	Transportin SR1, Importin 13	(Kramer & Hopper 2013; Senger et al. 1998)
Pse1 (Kap121)	Ribosomal proteins, Yra1, Spo12, Ste12, Yap1, Pho4, histones	Importin 5, Importin 6	(Rout et al. 1997; Greiner et al. 2004; Isoyama et al. 2001; Leslie et al. 2002; Chaves & Blobel 2001; Kaffman, Rank & O'Shea 1998)
Nmd5 (Kap119)	Ribosomal proteins, histones, Hog1, Crz1, Dst1(Xu et al. 2006) import	Importin 7, Importin 8	(Greiner et al. 2004; Straube et al. 2010; Ferrigno et al. 1998; Albertini et al. 1998; Rout et al. 1997)
Sxm1 (Kap108)	Lhp1, ribosomal proteins, Pab1	/	(Brune 2005; Chaves & Rosenblum 2011; Rout et al. 1997)
Kap114	TBP, histones, Nap1, Sua7	Importin 9	(Morehouse et al. 1999; Greiner et al. 2004; Hodges 2005; Straube et al. 2010)
Pdr6 (Kap122)	Toa1 and Toa2, TFIIA	/	(Titov & Blobel 1999)
Lph2 (Kap120)	Rpf1, Ho endonuclease, Swi6	Importin 11	(Caesar et al. 2006; Kim et al. 2010; Bakhrat et al. 2006)

Kap= Karyopherin, cNLS= classical nuclear localization signal

Overall, the translocation of a cargo through the NPC does not require energy, but recycling the of importins back to the cytoplasm depends on the energy provide by the Ran-GTP (Ribbeck et al. 1999; Schwoebel et al. 1998; Englmeier et al. 1999). One molecule of Ran-GTP is converted to Ran-GDP per cargo imported into the nucleus. In order to prevent complete depletion of Ran-GTP in the nucleus, nuclear transport factor 2 (NTF2/p10) transports Ran-GDP back to nucleus to allow the conversion of GDP to GTP by the nuclear protein RCC1 (Ribbeck et al. 1998). NTF2 does not belong to the Impβ-NTR; instead, it is small homodimer and has distich protein fold (Bullock et al. 1996;

Moore & Blobel 1994). However, similar to Impβ-NTRs, NTF2 also interacts with the FG-repeats at the central channel of the NPC to freely shuttle between the nucleus and the cytoplasm.
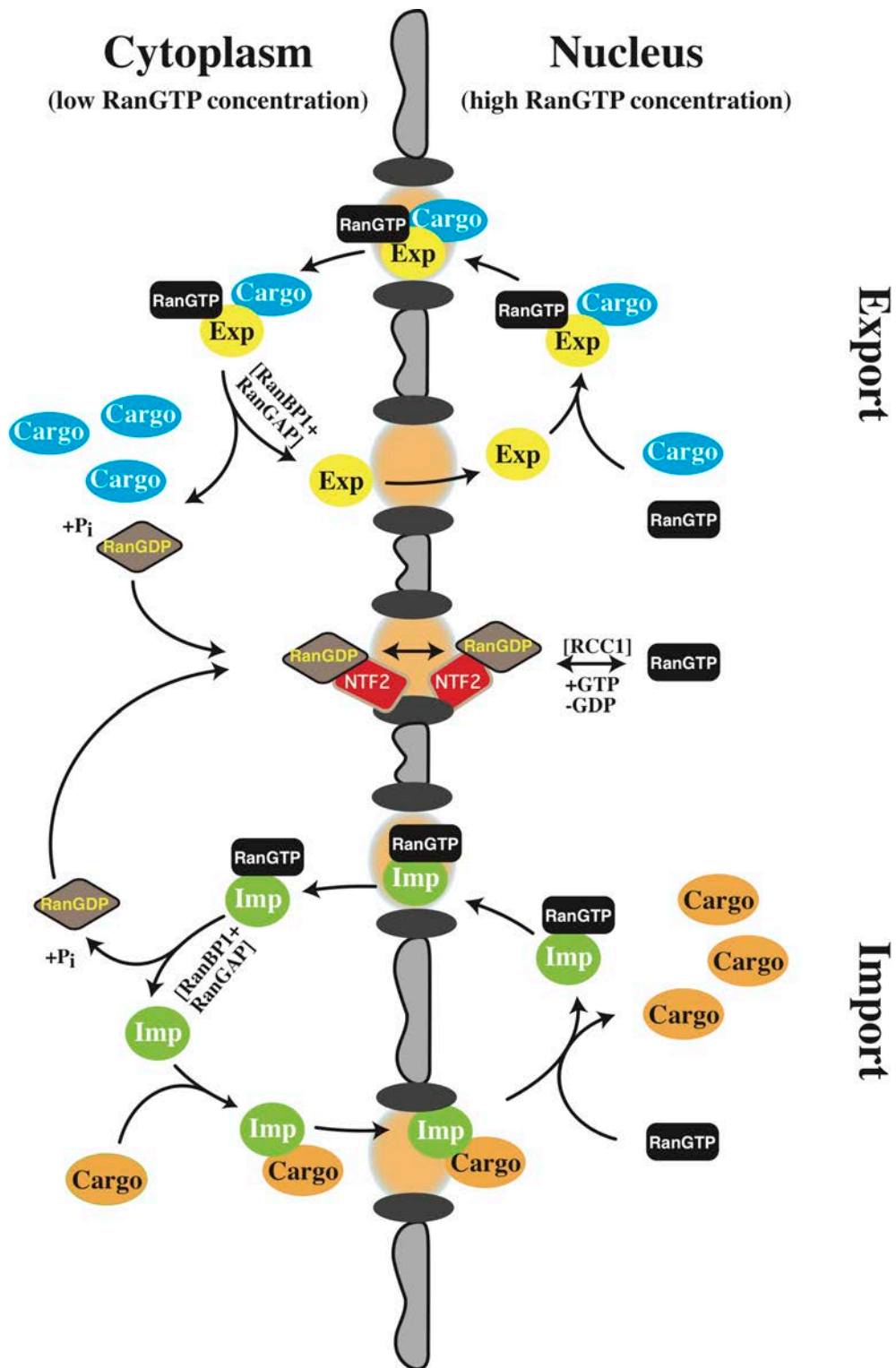
Figure 2.1 Import and export cycles. Picture taken and adapted from (Görlich & Kutay 1999). For details see sections 2.4 and 2.5 . Exp=Exportin, Imp= Importin, NTF2= Nuclear Transport Factor 2, Ran=Ras-related nuclear protein, RCC1= Regulator of Chromosome Condensation 1, RanGAP= Ran GTPase activating protein and RanBP1= Ran binding protein 1.

2.5   Nuclear export receptors (exportins)

Exportins mediate the active transport of macromolecules through the NPC from the nucleus to the cytoplasm. Eight different standard exportins that belong to the Impβ-NTR family are identified in higher eukaryotes (Table 2.3), whereas only 3 exportins are present in yeast (Table 2.4). Exportins recognize their cargos only in the nucleus where the concentration of Ran-GTP is high. After binding of Ran-GTP to a given exportin, the cargo is recognized by the exportin and consequently a trimeric export complex (exportin•cargo•Ran-GTP) is formed. Next, the export complex translocates to the cytoplasm, where the cargo is released upon hydrolysis of Ran-bound GTP promoted by the cytosolic protein RanGAP. Finally, the free exportin returns to the nucleus through the NPC in order to start the export of a new cargo (Figure 2.1).

Exportin 1 (Xpo1, also known as Crm1) is probably the NTR with the widest cargo recognition spectrum as it recognizes more than 700 different cargos in *S. cerevisiae* and more than 1000 protein in humans (Kirli et al. 2015). Most of these cargos were thought to be exclusively cytoplasmic after export by Crm1. However, the permeability barrier at the NPC is not perfect and cargos can diffuse back to the nucleus even if they exceed the exclusion size limit at the NPC (Bonner 1975; Mohr et al. 2009). As a result, Crm1 is probably the main strategy of a cell to keep a constant protein identity inside the nucleus by preventing protein accumulation inside the nucleus due to passive diffusion. For instance, Crm1 represses a premature nuclear translation of pre-mRNAs by keeping several translation factors outside the nucleus. Crm1 also exports the 60S pre-ribosomal subunit through the protein adaptor Nmd3 (West et al. 2007), and the 40S pre-ribosomal subunit via the protein factors Ltv1 and Rio2 (Fischer et al. 2015; Seiser et al. 2006). The export RanBP1 is another example of an important role of Crm1 to keep constant the protein identity in the nucleus since RanBP1 together with RanGAP is strictly required to dissemble Ran-GTP•NTR complexes in the cytoplasm (Coutavas et al. 1993; Bischoff & Görlich 1997; Kutay et al. 1997). As mentioned before, the strict cytoplasmic localization of RanBP1 and RanGAP is one of the key factors to give directionality during cargo transport. Therefore, failure to exclude RanBP1 and RanGAP from the nucleus would also impair other nuclear transport pathways.

Table 2.3 Cargos recognized by standard exportins in vertebrates

Vertebrate Exportins	Cargo(s)	Reference(s)
Exportin 1 (Crm1)	Leu-Rich NES cargos, HIV rev (adaptor for the export HIV genomic RNA), Snurportin 1, Nmd3 (adaptor for the export of 60S subunits), PHAX/CBC (adaptor for m^{7}G-capped UsnRNAs)	(Fornerod et al. 1997; Wen et al. 1995; Fischer et al. 1995; Paraskeva et al. 1999; Ho et al. 2000; Thomas & Kutay 2003; Izaurralde et al. 1995; Ohno et al. 2000)
Exportin 2 (CAS)	Importin α	(Kutay et al. 1997)
Exportin-t (Xpot)	tRNAs	(Arts et al. 1998; Kutay et al. 1998)
Exportin 5	tRNAs, eEF1A dsRNAs, pre-miRNAs	(Bohnsack et al. 2002; Calado et al. 2002; Brownawell & Macara 2002; Yi et al. 2003; Mingot et al. 2004; Bohnsack et al. 2004; Lund et al. 2004)
Exportin 6	Actin-profilin complex	(Stüven et al. 2003)
Exportin 7	p50RhoGAP, 14-3-3σ	(Mingot et al. 2004)

Crm1= Chromosomal maintenance 1

Table 2.4 Cargos recognized by standard *S. cerevisiae* exportins

S. cerevisiae Exportins	Cargo(s)	References	Vertebrate NTR homologue(s)
Exportin 1 (Crm1)	Export of cargos with a Leu-rich NES, Nmd3 (adaptor for the export of 60S subunits), Ltv1 (adaptor for the export of 40S subunits)	(Seiser et al. 2006; West et al. 2007; Kirli et al. 2015)	Exp1 (Crm1)
Cse1 (Kap109)	Kap60 (Srp1)	(Hood & Silver 1998)	CAS
Los1 (Kap127)	tRNAs	(Sarkar & Hopper 1998)	Exportin-t

Kap= Karyopherin, Crm1= Chromosomal maintenance 1

2.6 Bidirectional nuclear transport receptors

The transport of macromolecules across the NE is performed by NTRs with a defined directionality. Standard importins transport their cargos from the cytoplasm to the nucleus, whereas typical exportins carry the molecules in exactly the opposite direction. Interestingly, some NTRs act as importin as well as exportins, meaning that they can mediate the transport in a bidirectional fashion (Table 2.5). So far, two bidirectional NTRs have been described in vertebrates, importin 13 and exportin 4 (Mingot et al. 2001; Lipowsky et al. 2000; Kurisaki et al. 2006; Gontan et al. 2009), while Msn5 is the only

identified bidirectional NTR in *S. cerevisiae* (Kaffman, Rank, O'Neill, et al. 1998; Yoshida & Blobel 2001).

A distinctive feature for a standard exportin or importin is their affinity to Ran-GTP in the absence of the cargo. Interaction between exportins and Ran-GTP is known to be rather low (Kutay et al. 1997; Fornerod et al. 1997; Hellmuth et al. 1998; Hood & Silver 1998; Maurer et al. 2001), whereas importins bind to Ran-GTP with a high affinity (Görlich et al. 1996; Bischoff & Görlich 1997; Schlenstedt et al. 1997; Hahn & Schlenstedt 2011). Consequently, it is well possible that a distinct property of bidirectional NTRs is an intermediate affinity to Ran-GTP in the absence of a cargo. For instance, Msn5 has around 200-fold less affinity to Ran-GTP (52 nM; Hahn & Schlenstedt 2011) as compared to Importin β (0.23nM; Hahn & Schlenstedt 2011) and about 50-fold more affinity than Crm1 ($\cong$3 μM; Paraskeva 1999). Following this criterion, the same holds true for already described bidirectional NTRs which the affinity to Ran-GTP is ≈40nM for Exportin 4 (Lipowsky et al. 2000) and <100nM for importin 13 (Grünwald & Bono 2011).

Table 2.5 Cargos recognized by human and yeast bidirectional NTRs

Bidirectional NTR	Cargos	References	Closest NTR homologue
Human			
Importin 13 (Imp13)	eIF1A (e), Ubc9 (i), Y14 (i)	(Mingot et al. 2001)	Mtr10
Exportin 4 (Exp4)	eIF5A (e), SMAD3 (e), Sox2 (i), SRY (i)	(Lipowsky et al. 2000; Kurisaki et al. 2006; Gontan et al. 2009)	/
S. cerevisiae			
Msn5 (Kap142)	Pho4 (e), Crz1 (e), Cdh1 (e), RPA (i)	(Kaffman, Rank, O'Neill, et al. 1998; Yoshida & Blobel 2001)	Exportin 5 (Exp5)

Kap= Karyopherin, Imp= Importin, Exp= Exportin, (e)= export cargo, (i)= import cargo

2.6.1 Yeast importin Lph2 (Kap120)

Lph2 is a non-essential NTRs and has protein orthologs in *D. melanogaster*, *X. tropicalis*, *M. musculus* and *H. sapiens* (Quan et al. 2008). The closest mammalian NTR ortholog for Lph2 is importin 11 with a 27% sequence identity (Görlich et al. 1997). There are several features that characterize Lph2 (Kap120) as a NTR. First, Lph2 is a member of the Impβ-NTR super family due to the sequence homology of its N-terminal domain. Second, Lph2 is able to bind Ran-GTP *in vitro* and *in vivo* (Caesar et al. 2006). Last, Lph2 localizes

inside the nucleus and interacts with different Nups (Rout et al. 2000; Allen et al. 2001). One function of Lph2 as importin in *S. cerevisiae* was first described after the identification of the ribosomal assembly export factor Rpf1 as its import cargo (Caesar et al. 2006). Deletion of Lph2 in *S. cerevisiae* leads to the accumulation of 60S pre-ribosomal subunit in the nucleus, perhaps due to an impaired Rpf1 import. Later, it was found that Lph2 also mediates the import of Ho endonuclease and Swi6 protein (a subunit of the transcription factor SBF) (Bakhrat et al. 2006; Kim et al. 2010). Lph2 has been considered only as an importin in *S. cerevisiae* and nothing is known about Lph2 binding proteins in the presence of Ran-GTP.

2.6.2   Yeast Importin Pdr6 (Kap122)

Pdr6 (Kap122) was initially described as a member of a gene family involved in pleiotropic drug resistance in *S. cerevisiae* (Balzi et al. 1987). Overexpression of Pdr6 in yeast caused sensitivity to the antibiotics borrelidin, hygromycin b and cycloheximide (Chen et al. 1991). Pdr6 was later redefined as a NTR involved in the import of the large subunit (Toa1) and the small subunit of (Toa2) of the transcription factor IIA (TFIIA) (Titov & Blobel 1999). However, the relation between pleiotropic drug resistance and the function of Pdr6 as a importin still remains to be elucidated.

Pdr6 is not an essential NTR and was named Kap122 due to its theoretical molecular mass (123.5 kDa). Consistent with its function as an NTR, Pdr6 is localized evenly between the cytoplasm and the nucleus as it shuttles between these two compartments. Also, phylogenetic analysis show that Pdr6 does not have a corresponding ortholog in any multicellular organism (Quan et al. 2008). Little is known about the cargo recognition specificity of Pdr6, and no other transport cargos for Pdr6 have been identified over the last 20 years.

2.7   Tandem affinity purification (TAP)

Proteins regulate most of the cellular processes by forming macromolecular complexes in a specific time and subcellular localization. Every protein within the complex has a specific function that might be exhorted only upon protein-protein interaction. In many cases, the combination of all different activities within the complex provides the final cellular

function. Consequently, protein research is gradually changing the focus of study from the functional analysis of individual proteins to the evaluation of the function and composition of protein complexes that exist inside the cells (Cusick et al. 2005; Yu et al. 2008).

Yeast two-hybrid system and protein chip-based methods have been extensively used to study protein-protein interactions (Zhu et al. 2001; Uetz et al. 2000). Both methods provide information of broad protein-protein interaction networks, but they still have some major shortcomings (reviwed in, Xu et al. 2010; Van Criekinge & Beyaert 1999). For instance, the yeast two-hybrid system is restricted to be performed in *S. cerevisiae* that might create false positives or false negatives if heterologous proteins are analyzed. As for the chip-based method, the preparation of proteins needed for the analysis and their respective printing on the chip implies a time consuming and extensive task. An alternative method to analyze the formation and composition of protein complexes that circumvents these disadvantages is tandem affinity purification (TAP).

TAP, as the name implies, is a technique to purify protein complexes using two different affinity chromatography steps. In order to trap specific protein forming a stable complex, a recombinant protein tagged with two different affinity tags and a protease cleavage site (TAP-tag) is used as bait. First, the TAP-tagged bait is expressed in the host of interest or incubated in a cellular lysate in a pre-purified form. Protein complexes containing the TAP-tagged bait are purified using two affinity matrices in a sequential manner. Protein elution from the first matrix is generally mediated by a site-specific protease, whereas the second protein elution step is performed by different means (*i.e.* pH elution, addition of EGTA, competitive elution). Last, eluted proteins are subjected to mass spectrometry analysis in order to identify the exact composition of the complexes (Morris et al. 2014) (Figure 2.2).

Although TAP was developed in *S. cerevisiae* (Rigaut et al. 1999), it has been successfully used in many different organisms including mammals, plants, Drosophila, amphibians and bacteria (Van Leene et al. 2007; Veraksa et al. 2005; Kumar et al. 2004; Bürckstümmer et al. 2006). TAP is a widely used analytic method to study protein-protein interaction due to several reasons. First, the protein complexes are purified under native-like conditions. Second, TAP allows the rapid identification of protein complexes without previous structural knowledge of the bait protein to be used. Third, extensive proteomic analysis can be performed to evaluate the complete interactome of the bait of interest after TAP. Finally, the background binding caused by protein contaminants can be very low due to the

protein elution mediated by the site-specific protease and the use of two highly specific affinity matrices.
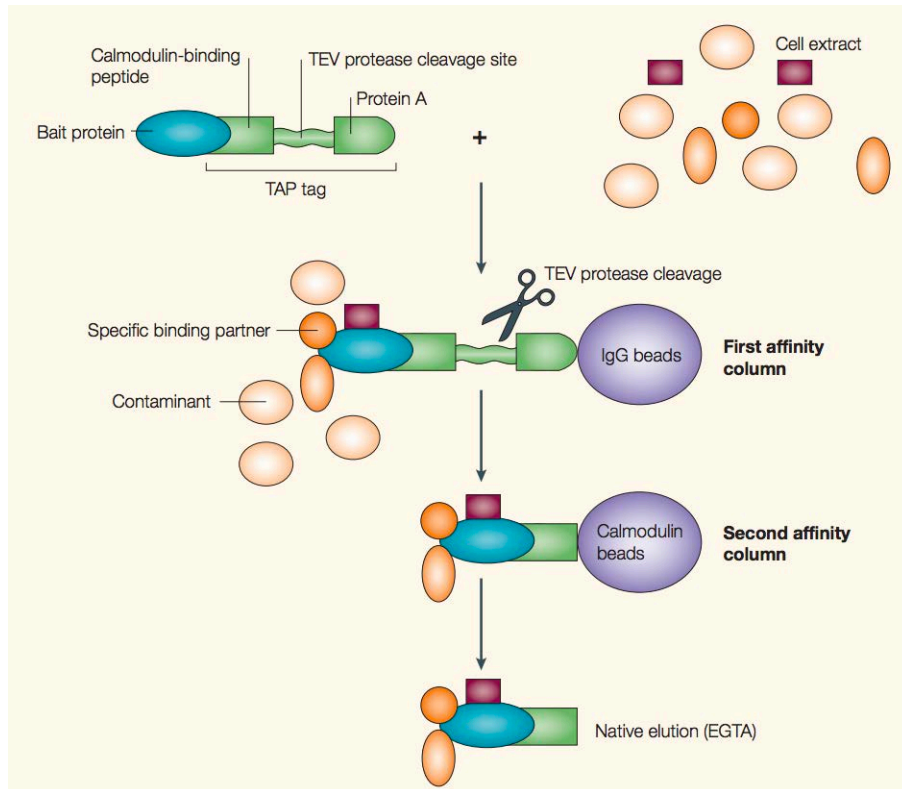
Figure 2.2 Tandem affinity purification. The tag used for the first developed tandem affinity purification protocol consisted of three components: a calmodulin-binding peptide, a TEV protease cleavage site and the IgG-binding protein A from *Staphylococcus aureus* (Rigaut et al. 1999). The TAP-tagged bait is either incubated with a pre-made cellular lysate or expressed in the appropriated host to trap specific protein binders. Protein complexes are initially purified using IgG-coupled beads. Protein contaminants are removed and immobilized protein complexes are then eluted using TEV protease. The second affinity chromatography is performed in a column that consists of calmodulin-coupled beads. Final protein elution is achieved by chelating calcium with EGTA. Picture taken and adapted from (Huber 2003).

2.8   Diversity of affinity tags and protease cleavage sites used in TAP

The first TAP-tag developed consisted of two IgG-binding units of protein A from *S. Staphylococcus aureus* (ProtA), the tobacco etch virus (TEV) protease cleavage site and the calmodulin-binding peptide (CBP) (Rigaut et al. 1999). ProtA and CBP acted as affinity tags, while TEV protease cleavage site was used to allow on-column protein elution mediated by the TEV protease. Over the years, numerous N- and C-terminal TAP-tags with different combinations of affinity tags sand protease cleavages sites have been used to improve the yield of purified protein complexes (Table 2.6). A common feature for more efficient TAP-tags is the replacement of CBP due its low efficiency as an affinity tag (Li 2010). In addition, several mammalian proteins containing a calmodulin-binding

domain can potentially co-purified and lead to false positive results. There are several options to replace CBP as affinity tag and increase the protein recovery efficiency in TAP (*i.e.* Strep-tag, ProtC, FLAG-tag, His-Tag, Myc-tag, HA-tag, S-tag). A more detailed description of about the different affinity tags used for TAP is summarized in table 2.7.

Table 2.6 Reported TAP-tags used in different cellular hosts

TAP-tag	Cleavage site	Organism	Reference
ProtA-CBP	TEV	*S. cerevisiae*	(Rigaut et al. 1999)
FLAG-HA	/	*A. thaliana*, human cells	(Zenser et al. 2008)
3x (FLAG-His)	/	*D. melanogaster*	(Yang et al. 2006)
ProtA-ProtC	TEV	*T. brucei*	(Schimanski et al. 2005)
ProtG-SBP	TEV	Human cells	(Bürckstümmer et al. 2006)
2x (FLAG-ProtA)	TEV	Human cells	(Tsai & Carstens 2006)
2x(His)-Strep	2x TEV	Human cells	(Giannone et al. 2007)
2x (Strep-FLAG)	/	Human cells	(Gloeckner et al. 2007)
SBP-HA	/	Human cells	(Glatter et al. 2009)
SBP-His	/	Human cells	(Li et al. 2011)
His-biotin	/	*S. cerevisiae*	(Tagwerker 2006)
GFP-S/His	TEV/HRV3C	*C. elegans*	(Cheeseman & Desai 2005)
ProtA-Myc-His	/	*N. benthamiana*	(Liu et al. 2004)
ProtA-CBP	TEV	*E. coli*	(Kumar et al. 2004)

ProtC= Protein C epitope, ProtG=*Streptococcus* protein G, SBP=Streptavidin-binding protein, HRV3C=Human rhinovirus 3C protease. Information taken and adapted from (Li 2011).

As mentioned before, a protease cleavage site is introduced in between the two affinity tags in the TAP-tag to allow on-column protein cleavage mediated by a site-specific protease. In contrast to other protein elution methods, on-column protein cleavage releases protein complexes bound to the affinity matrix under native-like conditions, and provides a purer protein sample (Frey & Görlich 2014b). The TEV protease cleavage site is the most used tool for on-column protein cleavage during TAP. Since very few mammalian proteins contain the TEV protease cleavage site (ENLYFQ|X; X≠P), protein cleavage of unspecific targets by the TEV protease is a rare event. An alternative to TEV protease is the human rhinovirus 3C protease cleavage site (HRV3C), which recognizes the amino sequence (LEVLFQ|GP) as a cleavage site (Knott et al. 1989; Cordingley et al. 1989; Cordingley et al. 1990). Unlike TEV protease, human rhinovirus 3C protease retains its enzymatic activity at 4°C, and therefore a shorter incubation time is needed to achieve complete protein cleavage (Ullah et al. 2016). So far, TEV and human rhinovirus 3C protease are the

only two documented examples of proteases used for TAP. Despite the fact that SUMO-specific proteases are the most efficient proteases used for protein cleavage *in vitro* (Malakhov et al. 2004; Frey & Görlich 2014a), there is no report of their use in TAP. Consequently, the SUMO-specific protease and its cognate cleavage site SUMO protein are described in the next sections.

2.9   An overview of SUMO as protein function modulator

The small ubiquitin-like modifier (SUMO in mammals, Smt3 in yeast) belongs to a small group of proteins that share similar protein folding and sequence homology with ubiquitin (reviewed in Kerscher et al. 2006; Herrmann et al. 2007). SUMO is covalently attached via its very C-terminal glycine residue to a lysine residue on the protein targets. SUMO conjugation, also known as SUMOylation, is considered to be a post-translational modification as it changes the properties or the localization of the modified substrates. In fact, SUMO conjugation was first described as a post-translation modification that modulates the partitioning of RanGAP1 between the cytoplasm and the NPC (Matunis et al. 1996; Mahajan et al. 1997). SUMOylation plays an essential role in many cellular functions like nuclear transport, signal transduction, protein stabilization, genome stability and regulation of transcription (reviewed in Vertegaal et al. 2004; Hendriks & Vertegaal 2016).

A single SUMO gene is expressed in unicellular eukaryotes, whereas multicellular organisms express three or more paralogs genes (Müller et al. 2001). Despite the fact that all SUMO paralogs in a cell have a high sequence identity (50-95%), they are involved in the regulation of different cellular processes due to their different cellular localization (Saitoh & Hinchey 2000). Nevertheless, all SUMO paralogs require the same enzymatic cascade in order to be conjugated to the protein targets. Similar to ubiquitin, SUMO is translated as a precursor protein that needs to be cleaved at the C-terminus in order to expose the very C-terminal double glycine motif. Next, the heterodimeric E1 enzyme (Aos1•Uba2 complex in *S. cerevisiae*) activates the mature SUMO by forming a high energy SUMO•AMP complex. After activation, SUMO is transferred to a reactive cysteine residue localized in the E2-conjugating enzyme (Ubc9, in yeast). There are two different ways how a protein is modified by SUMO conjugation in eukaryotic cells. In the first mechanism, Ubc9 recognizes the SUMO conjugation motif (ϕKXQ/E, where ϕ is an

hydrophobic residue and X can be any residue) on the substrate, and attaches SUMO by forming an isopeptide bond between the C-terminal glycine of SUMO and the ε-amino group of a lysine in the substrate (Johnson & Blobel 1997). For the second mechanism, SUMO E3 ligases conjugate SUMO to protein targets that do not have the SUMO conjugation motif. There are three different classes of SUMO E3 ligases according to their protein structure: the protein inhibitor of activated STAT family (PIAS), polycomb-group protein 2 and RanBP2/Nup358 (Kagey et al. 2003; Pichler et al. 2004; Hochstrasser 2001). These three classes do not overlap in substrate recognition and localize in different subcellular compartments.

Table 2.7 Affinity tags used for tandem affinity purification

Affinity tag	Binding partner	Comments
Calmodulin-binding protein (CBP)	Calmodulin	Calmodilin-binding proteins present in mammalian cell lysates might bind unspecifically to the matrix.
Staphylococcal protein A (ProtA)	IgG	A low pH buffer is needed for protein elution. Thus, protein elution mediated by protein cleavage site is highly recommended.
S-tag	S protein	S-tag can be used to detect protein expression levels. Protein elution is preferably performed using a site-specific protease.
FLAG-tag	Anti-FLAG antibody	The matrix couple to the anti-FLAG antibody is highly unstable and expensive.
Strep-tag II	Strep-Tactin	Strep-tag II offers a high yield of pure protein. Buffer conditions used for protein elution are quite flexible.
Hemagglutinin (HA)	Anti-HA antibody	HA tag is used to detected expression level of the tagged protein.
His-tag	Ni^{2+}	It allows a high yield of pure protein. However, the use of a His-tag does not allow the complete removal of contaminants from the cellular lysate.
Protein C epitope (ProtC)	HPC4 antibody	Alternative as an affinity tag when EGTA interferes with protein function
Streptavidin-binding peptide (SBP)	Streptavidin	Strong binding pair that enables a high yield of pure protein.
c-myc-tag	Anti-myc antibody	Normally, the tag is not for protein purification but for measuring protein expression levels.
Streptococcus protein G (ProtG)	IgG	Higher Ig-biding spectrum than ProtA. Protein cleavage is recommended for protein elution.
Biotinylation tag	Avidin	The tag has a high affinity to biotin and requires biotinylation with BirA enzyme prior to use.

Information taken and adapted from (Xu et al. 2010)

2.10 SUMO-specific proteases

Protein cleavage of SUMO is mediated by cysteine SUMO-specific proteases named Ulp (Ubiquitin-like protease) in *S. cerevisiae* and SENP (Sentrin protease) in higher eukaryotes. Cleavage at the C-terminus of SUMO serves for two different purposes. First, the SUMO precursor is cleaved by a SUMO-specific protease to convert it to the mature form prior SUMO conjugation. Second, SUMO-specific proteases also remove attached SUMO and poly-SUMO chains from the conjugated substrates in order to revert the post-translational modification.

Two different SUMO-specific proteases are identified in *S. cerevisiae* (Ulp1 and Ulp2). Ulp1 is in charge of the maturation of SUMO, localizes in the nuclear pore complex and is essential for cell cycle progression (Li & Hochstrasser 1999). In contrast, Ulp2 is not essential, resides inside the nucleus and is required for chromosome stability and cell recovery after cell cycle checkpoint arrest (Li & Hochstrasser 2000). In humans, six different SENP proteases have been described (SENP1, SEN2, SENP3, SENP5, SENP6 and SENP7). Similar to SUMO-specific proteases in yeast, the different human SENPs have preferential SUMO paralogs, differ in their isopeptidase/peptidase activities, and have a specific subcellular localization (Table 2.8). Thus, SUMO-specific proteases process a different group of SUMOylated substrates and cannot compensate for the absence of each other within the same cell.

Table 2.8 Properties of the different SUMO-specific proteases in yeast and mammals

Protease	Localization	SUMO paralog preference	Protease functions
S. cerevisiae			
Ulp1	Nuclear pore complex	Smt3	SUMO precursor maturation, SUMO deconjugation
Ulp2	Nucleoplasm	Smt3	SUMO deconjugation and poly-SUMO chain remodeling
Mammals			
SENP1	Nuclear pore complex	SUMO-1/-2/-3	SUMO precursor maturation and SUMO deconjugation
SENP2	Nuclear pore complex and cytoplasm	SUMO-2/-3 >SUMO-1	SUMO precursor maturation and SUMO deconjugation
SENP3	Nucleolus	SUMO-2/-3	SUMO deconjugation
SENP5	Nucleolus and mitochondria	SUMO-2/-3	SUMO precursor maturation and SUMO deconjugation
SENP6	Nucleoplasm	SUMO-2/-3	SUMO deconjugation and poly-SUMO chain remodeling
SENP7	Nucleoplasm	SUMO-2/-3	SUMO deconjugation and poly-SUMO chain remodeling

Information taken and adapted from (Hickey et al. 2012)

2.11 The SUMO system as a tool for protein expression and purification

2.11.1 SUMO as a fusion tag

Efficient protein expression systems are one of the rate-limiting factors for the rapid development of proteomic, genomic and structural studies. The most exploited protein expression system is the bacterium *E. coli* due to its genomic simplicity, low cost of biomass production and extensive characterization. However, *E. coli* is often not able to produce large amounts of soluble recombinant proteins, especially eukaryotic proteins. To tackle this problem, several fusion protein tags are used to increase protein expression levels, improve protein folding and prevent protein degradation (reviewed in Kimple et al. 2013; Kosobokova et al. 2016). In comparison to many other widely fusion tags (*i.e.* MBP, Trx, NusA, His-tag), SUMO protein enhances higher levels of expression and solubility of difficult-to-express eukaryotic proteins in *E. coli* (Malakhov et al. 2004; Zuo, Li, et al. 2005; Zuo, Mattern, et al. 2005; Marblestone et al. 2006). Similar to ubiquitin (Khorasanizadeh et al. 1996), SUMO protein might act as chaperone to promote proper protein folding and solubility of its recombinant fusion partners. SUMO's fast kinetic protein folding (Marblestone et al. 2006), hydrophilic surface and highly hydrophobic inner core suggest that SUMO acts as a nucleation site for the proper folding of recombinant SUMO-tagged proteins.

Recombinant SUMO-tagged proteins are generally purified by immobilized metal-affinity chromatography using a His-tag fused to the N-terminus of SUMO. In combination with a SUMO-specific protease, the His-SUMO-tagged recombinant protein can be purified and regenerated with its native N-terminal amino acid composition in a single step (Malakhov et al. 2004; Lee et al. 2008). Due to its unique features, the use of the SUMO technology is exponentially growing and many different commercial SUMO expression vectors are available containing different combinations of affinity tags.

2.11.2 SUMO-specific proteases as tool for tag removal

The removal of the fusion tag from a protein of interest after or during the purification process is often desired. The presence of the fusion tag might interfere with the protein function or structure as shown in (Amor-Mahjoub et al. 2006; Woestenenk et al. 2004;

Chant et al. 2005; Suh-Lailam & Hevel 2009). Chemical and enzymatic methods are the two options available to remove a tag from the protein of interest. Unlike chemical methods, the enzyme-based approach can have an extremely high specificity and can be used in a wide range of conditions. The enzymatic removal of the fusion tag requires the use of a site-specific protease and its cognate cleavage recognition located in between the fusion tag and protein of interest. To date, there are several commercially available proteases and that have been successfully used for tag removal (*i.e.* TEV protease, Factor Xa, enterokinase, human rhinovirus 3C protease, thrombin, SUMO-specific protease). Most of the proteases nowadays used for fusion tag removal present several problems: poor efficiency, time-consuming optimization protocols, low cleavage specificity, difficult production of the protease, or failure to restore the native N-terminus of the protein of interest (reviwed in Yan et al. 2009). SUMO-specific proteases, used to cleavage the SUMO protein from the protein of interest, overcome all the difficulties previously mentioned.

SUMO-specific proteases from different organisms are by far the most activate proteases for protein cleavage *in vitro* (Frey & Görlich 2014a). At the same conditions, TEV protease is able to cleave only a 5-fold molar excess of substrate, and SUMO-specific proteases are able to cleave up to a 10,000-fold molar excess of substrate at 4°C within 1 h, (Frey & Görlich 2014a). In addition, SUMO-specific proteases are shown to more tolerate a wide range of buffer conditions including acidic or basic pH, high ionic strength (1 M NaCl) and the presence of denaturing reagents (2 M urea) (Peroutka III et al. 2011). Another advantage of SUMO-specific protease over most of the routinely used proteases is the generation of a native N-terminal sequence after the removal of the fusion tag. Site-specific proteases such as TEV protease, thrombin and human rhinovirus 3C protease leave unwanted residual amino acids at the N-terminus of the protein of interest after protein cleavage (Table 2.9). For instance, thrombin cleaves the sequence LVPR|G after the arginine residue; therefore, the resulting N-terminus of the recombinant protein is restricted to glycine. Unlike thrombin, SUMO-specific proteases restore the native N-terminus of the protein of interest since SUMO is efficiently removed when any amino acid (except proline) occupies the +1 position of the cleavage site (Malakhov et al. 2004). Therefore, almost any wanted N-terminal sequence can be generated by SUMO-specific proteases. A good example for the importance of generating an intact N-terminal sequence is the production of functional chemokines since their activity relies on their native N-

terminus (Lu et al. 2009). Perhaps the most distinct and important feature of SUMO-specific proteases over other site-specific proteases is that they recognize the tertiary structure of a protein (SUMO) as cleavage recognition site, and consequently unspecific cleavage of protein within a small linear amino acid sequence is simple not possible.

Table 2.9 Proteases and protease cleavage sites used for tag removal

Protease	Recognition site	Residues after cleavage	Cost[1] (dollar/mg)
Immobilized subtilin BPN	C-terminus of propeptide|X	/	26.00
Enterokinase	DDDDK|X	/	45.00
Factor Xa	IEGR|X	/	4.00
Thrombin	LVPR|G	G	4.00
TEV	EXXYXQ|(G/S)	G/S	45.00
SUMO protease	XGG|X[2]	/	3.15

[1]Price based on the amount protease needed according to the manufacturers protocol.
[2]XGG represent the extreme C-terminal double glycine motif of SUMO protein.
X= any of the 20 amino acid. Information taken and adapted from (Yan et al. 2009)
\| = hydrolyzed peptide bond

2.12   Protein purification assisted by the SUMO system

As mentioned before, fusion tags are used to enhance the expression and solubility levels of recombinant proteins. Additionally, they can also be used as affinity tags to assist protein purification by affinity chromatography. For instance, MBP, Trx and GST are good examples of fusion tags with this dual function during expression and purification of recombinant proteins (McCoy & La Ville 1997; Harper & Speicher 2011; Pattenden & Thomas 2008). Unfortunately, SUMO protein serves only as a solubility enhancer but it cannot be used as an affinity tag. To circumvent this problem, different affinity tags (*i.e.* His-tag and ProtA) are coupled to the N-terminus of SUMO to ease the purification of SUMO-tagged proteins using affinity chromatography.

The elution of the SUMO-tagged protein from the affinity column is performed either by competitive elution or by on-column protein cleavage. On-column protein cleavage yields an untagged and a purer protein sample compared to proteins samples obtained by competitive elution (Frey & Görlich 2014b). Additionally, SUMO-specific proteases and different ubiquitin-like proteases are used together for the purification of binary protein complexes using two consecutive cycles of affinity chromatography and on-column protein cleavage (Frey & Görlich 2014b). In short, the purification of the protein complex requires the tagging of each subunit with an orthogonal ubiquitin like proteins (*i.e.* SUMO

and NEDD8) and an affinity tag. This purification technology is proved to be a rapid method to generate untagged and highly pure protein complexes with a defined stoichiometric composition.

2.13 Use of the SUMO system in eukaryotic cells

Hundreds of different SUMO-tagged proteins have been successfully expressed and purified in *E. coli* (reviewed in Butt et al. 2005). One of the most important aspects for the success of the SUMO technology in *E. coli* is the lack of endogenous SUMO-specific proteases. Unlike prokaryotic cells, eukaryotes possess endogenous SUMO-specific proteases, and consequently the expression of SUMO-tagged proteins results in the *in vivo* cleavage of SUMO. In order to produce stable SUMO-tagged protein in eukaryotic hosts, a SUMO mutant protein from *S. cerevisiae* (SUMOstar) that is no longer recognized by endogenous eukaryotic SUMO-specific proteases was engineered (Peroutka et al. 2008; Butt et al. 2010). The design of SUMOstar was based on the structural information of the complex between yeast SUMO and scUlp1 protease (Bohnsack et al. 2004). Specifically, two mutations (R64T and R17E) were introduced in yeast SUMO to disrupt the interaction with the wild type yeast scUlp1 protease. Consequently, SUMOstar-tagged proteins are not cleaved either *in vivo* or *in vitro* by scUlp1. In addition, SUMOstar preserved the ability to enhance protein expression and solubility as wild type yeast SUMO protein (Peroutka et al. 2008). In fact, expression levels of eGFP, mouse UBP43, Usp4, Usp15 and tryptase were significantly higher in insect cells only after fusion of SUMOstar protein at the N-terminus of the recombinant proteins (Liu et al. 2008).

In order to remove the SUMOstar tag from the protein of interest, SUMOstar protease was engineered by introducing three mutations (D451S, T452G and E455S) in scUlp1 (Butt et al. 2010). As shown by *in vitro* experiments, SUMOstar protease efficiently cleaves SUMOstar-tagged substrates, but wild type yeast SUMO fusion proteins are also cleaved. SUMOstar protease and scUlp1 are not orthogonal proteases, thus they cannot be used for the purification of dimeric complexes using two consecutive cycles of affinity chromatography as shown in (Frey & Görlich 2014b).

2.14 Aim of this work

Tandem affinity purification (TAP) seemed to us as the perfect methodology to identify potential novel export and import cargos for yeast NTRs. So far, most of the described TAP protocols use the highly inefficient TEV protease for on-column protein elution and require laborious preparation steps prior protein identification by LC/MS analysis. Here, we sought to create an alternative to conventional TAP methodologies that would generate samples ready to be analyzed directly by MS without any previous treatment. Moreover, we wanted to optimize on-column protein elution by using a highly active SUMO-specific protease and its cognate SUMO protein.

The presence of endogenous SUMO-specific proteases in eukaryotic cells is the only impediment for us to use SUMO protein as fusion tag for the identification of protein binders. In addition, we wanted to evolve a SUMO system that could be used for many other applications such as protein expression, protein purification and *in vivo* protein cleavage in eukaryotic organisms. We thus aimed to create a mutant SUMO protein and a mutant SUMO-specific protease from *B. distachyon* that could be used in yeast and any other eukaryotic host. To generate such mutants, we intended to design an *in vivo* selection system in *E. coli* that relies on bacterial survival. Moreover, the selection system should provide the possibility of selecting SUMO-specific proteases with orthogonal substrate specificity. Specifically, we wanted to isolate a mutant SUMO-specific protease with orthogonal specificity to the SUMOstar protease to purify binary protein complexes in *S. cerevisiae*.

Using the SUMOvera system and our improved TAP methodology, we aimed to identify protein binders of the poorly described yeast NTRs, Lph2 (Kap120) and Pdr6 (Kap122). We speculate that there might be high possibility for identifying more cargos since only one has been described for Lph2 (Caesar et al. 2006) and just two for Pdr6 (Titov & Blobel 1999).

3 RESULTS

3.1 Engineering of SUMO and SUMO-specific protease mutants with novel features

3.1.1 Selection system for SUMO-specific protease mutants with orthogonal substrate specificity

The initial aim of this study was to evolve a new SUMO variant (named here as SUMOvera) that is not cleaved by the SUMOstar protease or any other SUMO-specific protease, and to evolve the cognate SUMO-specific protease (SUMOvera protease) that cleaves SUMOvera, but neither SUMOstar nor any wild type SUMO. Therefore, SUMOvera protease and SUMOstar protease would have mutually exclusive specificities. In order to evolve a pair of proteases with orthogonal substrate specificity, we required to develop a system that can select "for" and "against" the cleavage of two different SUMO proteins at the same time. This means that the ideal selection system must couple a positive and a negative selection step. For that, we used the ssrA degradation signal (reviewed in Keiler 2008; Himeno et al. 2014), an N-end rule degron (degron[NER]) (Bachmair et al. 1986), and the hygromycin B phosphotransferase (Hph) that allows *E.coli* cells to survive in the presence of hygromycin B (HygB) (Rao et al. 1983).

Hph was expressed as a SUMO[1]-degron[NER]-Hph-SUMO[2]-ssrA fusion (Figure 3.1). ssrA is a small peptide (AADENYALAA) that triggers rapid degradation of Hph, unless it is removed after cleavage of SUMO[2] by a co-expressed SUMO-specific protease. "Degron[NER]" is a strong N-end-rule degradation signal (FLFVQ) (Wang et al. 2008) that remains silent in the fusion context, but gets only activated if SUMO[1] is cleaved by a co-expressed SUMO-specific protease. Thus, Hph will only be stable and the cells be HygB-resistant if the co-expressed SUMO-specific protease cleaves SUMO[2], but not SUMO[1]. In other words, bacterial are only resistant to HygB when the SUMO-specific protease has orthogonal substrate specificity for SUMO[1] and SUMO[2].
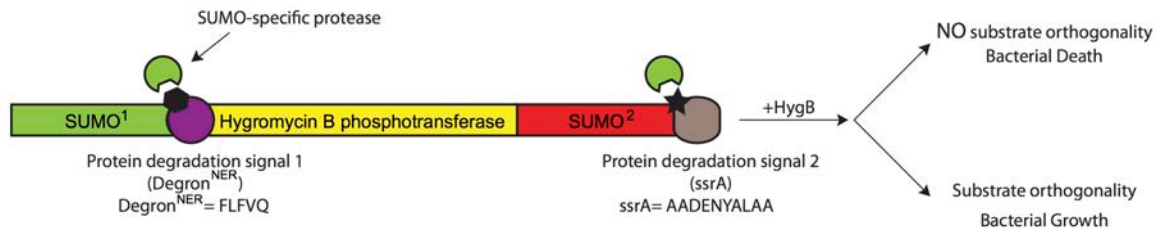
Figure 3.1 Schematic overview of the selection system for orthogonal SUMO-specific proteases. The designed *in vivo* selection system is based on the survival of *E. coli* cells on medium containing Hygromycin B (HygB). Cells co-express Hygromycin B phosphotransferase (Hph) as a selection marker and a SUMO-specific protease. The expression of the SUMO-specific protease is IPTG-inducible while Hph is constitutively expressed. Two different SUMO proteins are fused to Hph at both termini. The recognition of SUMO[1] and SUMO[2] by the SUMO-specific protease is linked to the half-life of Hph that is control by two different protein degradation signals (degron[NER] and ssrA). Degron[NER] signal that is only activated after the SUMO-specific protease cleaves SUMO[1]. On the other hand, the ssrA signal is inactivated only when SUMO[2] is cleaved by the co-expressed SUMO-specific protease. Therefore, bacterial growth is only possible when the protease recognizes only the SUMO[2], but not SUMO[1].

To validate the design of the system, we co-expressed bdSENP1 and a series of different bdSUMO-degron[NER]-Hph-bdSUMO-ssrA fusion proteins in *E. coli* cells and tested the survival of the cells in medium containing HygB and IPTG (Figure 3.2. A). To account for the absence of substrate recognition by bdSENP1 during the validation, we engineered a non-cleavable bdSUMO mutant (bdSUMO[*]) (Kuwata & Nakamura 2008). *E. coli* cells that co-expressed bdSENP1 and bdSUMO[*]-degron[NER]-Hph-bdSUMO-ssrA showed similar HygB resistance as cells expressing a Hph fusion protein lacking both protein degradation signals (Figure 3.2. E and F). In contrast, bacterial growth was not observed when wild type bdSUMO was placed in the N-terminus of the selection marker (Figure 3.2. B and D). Cleavage of the N-terminal bdSUMO by bdSENP1 led to the activation of the degron[NER] signal and thus Hph was degraded. In addition, cells expressing Hph with a C-terminal bdSUMO* tag did not grow in medium containing HygB (Figure 3.2. B and D). The inability of bdSENP1 to cleave bdSUMO* did not allow the release of the ssrA signal from the selection marker. In short, cells only survived when both protein degradation signals were inactivated by the cleavage of the C-terminal bdSUMO and the null-recognition of N-terminal bdSUMO*. These experiments clearly showed that the design of our *in vivo* system was functional and appropriated for the selection of SUMO-specific proteases with orthogonal substrate specificity.
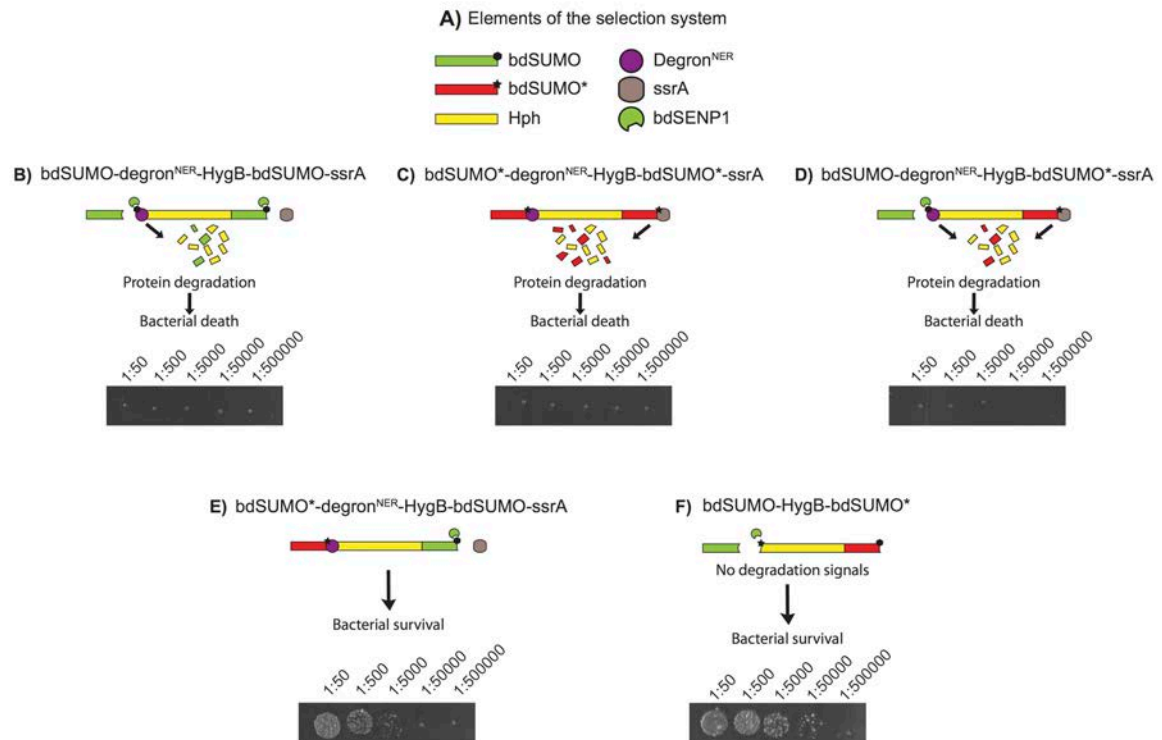
Figure 3.2 Functional test of the selection system for SUMO-specific proteases with orthogonal specificity. (A) SUMO (bdSUMO) and SUMO-specific protease 1 (bdSENP1) from *B. distachyon* were used as model proteins to test the functionality of our designed *in vivo* selection system in *E. coli*. A double mutation was introduced in bdSUMO (G96A; G97A) to generate a non-cleavable SUMO mutant (bdSUMO*). Degron[NER] and ssrA protein degradation signals were used to control the half-life of Hygromycin B phosphotransferase (Hph). **(B, C, D, E, F)** Hph was tagged at both termini with different combinations of bdSUMO and bdSUMO* (4 possible combinations) in order to test for and against protein cleavage by bdSENP1. Bacteria were co-transformed with two different plasmids encoding for bdSENP1 and a Hph fusion protein, respectively. Transformed cells were diluted in a 10-fold dilution series and further medium containing HygB and IPTG. Cells expressing Hph fused to bdSUMO at the N-terminus led to bacterial death due to the activation of the degron[NER] signal after protein cleavage (see B and D). Moreover, *E. coli* cells harboring a Hph fusion protein with the C-terminal bdSUMO* tag died because of the null inactivation of ssrA signal (see C). Bacteria survived only after the permanent inactivation of both protein degradation signals (see E). A Hph construct lacking both protein degradation signals was used as positive control for bacterial growth (see F).

3.1.2 Design of a SUMO/SUMO-specific protease mutant pair with orthogonal specificity to SUMOstar/SUMOstar protease

In practice, the approaches to obtain the bdSUMO/bdSENP1 mutant pair with new specificities would be either random or directed mutagenesis. The later method is simpler, faster and more convenient, however, it is only possible to apply when structural data of the proteins to be mutagenized are available (Packer & Liu 2015). Unfortunately, the protein structure for the bdSUMO•bdSENP1 complex is not known. However, several crystal structures of different SUMO•SUMO-specific protease complexes from several organisms are already determined (Xu et al. 2006; Reverter & Lima 2006; Shen et al. 2006; Bohnsack et al. 2004; Reverter & Lima 2004). Analysis of these structures revealed that the interaction between the different SUMOs and the respective proteases occurs

mostly through highly conserved residues in both proteins. We therefore tried to investigate whether these interacting residues were also conserved in bdSUMO and bdSENP1 in order to use directed mutagenesis for the generation of mutant pair. To do so, we performed a structure-based alignment including the catalytic domain of bdSENP1 (residues 247-481) and several other SUMO-specific proteases with known atomic structure (Figure 3.3. A). Similarly, bdSUMO was also aligned to several already crystallized SUMO proteins (Figure 3.3. B). We observed clusters of conserved residues located within all interacting domains between SUMO and the SUMO-specific protease. In contrast, long stretches of non-conversed residues were mainly observed in regions of the SUMO-specific proteases that are not involved in the recognition of SUMO. As residues involved in the interaction between SUMO and the corresponding protease were also conserved in bdSUMO and bdSENP1, we speculated that the interaction between bdSUMO and bdSENP1 is likely to happen in a similar fashion as for the rest of the SUMO•SUMO-specific protease complexes.

In order to generate a bdSUMO/bdSENP1 mutant pair with orthogonal specificity to SUMOstar/SUMOstar protease, we decided to specifically mutagenize 3 different residues in bdSUMO (T60, D67 and Q75) and their four most likely interacting residues in bdSENP1 (R269, N280, R346 and K350). These seven chosen residues are in different interacting motifs than the mutated amino acids in SUMOstar and SUMOstar protease (Butt et al. 2010) (Figure 3.3 A and B). We therefore expected higher probabilities in achieving complete orthogonal substrate specificities between a bdSENP1 mutant and SUMOstar protease.

3.1.3   SUMOvera and SUMOstar are orthogonal substrates for SUMOstar protease

First, we sought to select a bdSUMO mutant that is not recognized by SUMOstar protease using our *in vivo* selection system described in section 3.1.1 . For that, we created a library of bdSUMO mutants (bdSUMOmut) where the codons for the three previously chosen residues were randomized (T60X, D67X and Q75X). The library was then cloned at the N-terminus of the selection marker to screen against substrate cleavage by SUMOstar protease. SUMOstar protein was fused to the C-terminus of Hph to account for efficient

protein cleavage by SUMOstar protease (Figure 3.4. A). In short, the construct used during selection consisted of bdSUMOmut-degron^NER-Hph-SUMOstar-ssrA.
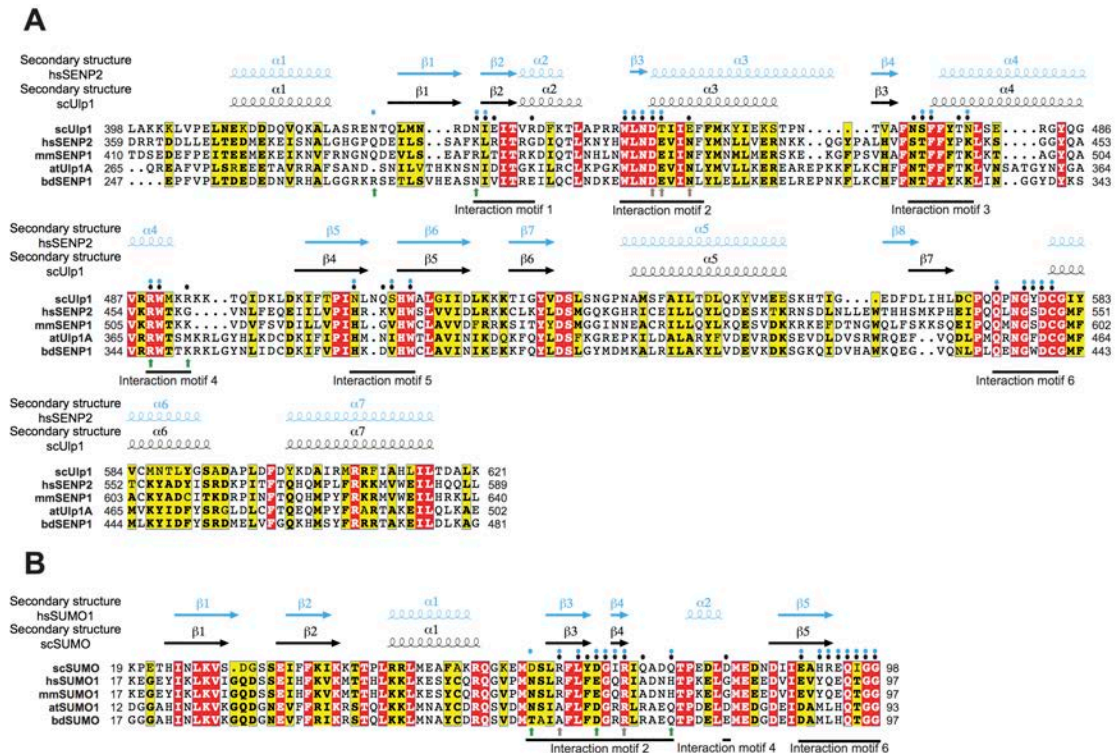
Figure 3.3 Structure-based alignment of different SUMO and SUMO-specific proteases orthologs. (A) Multiple sequence alignment between SUMO-specific proteases from *S. cerevisiae* (scUlp1), human (hsSENP2), *M. musculus* (mmSENP1) *A. thaliana* (atUlp1A) and *B. distachyon* (bdSENP1). Identical and similar side chain residues are highlighted in red and yellow, respectively. The numbering of the amino acids is according to the full-length sequence of each protease. Points in between the sequences denote gaps. α-helices and β-strands for hsSENP2 and scUlp1 are displayed and numbered above the sequences in blue and black, respectively. Side chain residues involved the in interaction between hsSENP2 and scUlp1 with their cognate substrate are marked with blue and black (•), respectively. Interactions motifs were indicated according to (Bohnsack et al. 2004). Amino acids to be mutagenized in bdSUMO and bdSENP1 are marked with green arrows. Mutagenized residues in scSUMO and scUlp1 to obtain SUMOstar and SUMOstar protease are pointed with brown arrows. ESPript 3.0 online software was used to depict the alignment (Robert & Gouet 2014). **(B)** Multiple sequence alignment of SUMO proteins from the same organisms indicated above. The numbering of the residues is assigned accordingly to the full-length SUMO sequences. Side chain residues similarities, secondary structures, interacting residues and interaction motifs are denoted as in (A).

The selection of bdSUMO mutants resulted in the isolation of 10 different colonies with a strong preference for the D67K mutation and high amino acid variability in the other two mutagenized positions (Figure 3.4. B). We tested all clones individually using the *in vivo* selection system to determine to what extent these bdSUMO mutants were recognized by SUMOstar protease. Cells were diluted and grown on selective medium containing a higher concentration of Hygromycin B than the one used during selection of the mutants (Figure 3.4. C). We did not observe cellular growth in any of the tested samples when IPTG was not included in the media. The lack of SUMOstar protease expression allowed the gradation of Hph due to the permanent presence of the C-terminal ssrA protein

degradation signal. When the stringency of selection was increased by addition of IPTG, cells expressing bdSUMO mutant 1 (bdSUMOmut1-degron^{NER}-Hph-SUMOstar-ssrA) showed the best bacterial survival compared to all other mutants. In fact, cells expressing bdSUMO mutant 1 had a similar growth behavior as positive control cells containing the non-cleavable bdSUMO* mutant (bdSUMO*-degron^{NER}-Hph-SUMOstar-ssrA). The bdSUMO mutant 1 that is not cleaved by SUMOstar protease and comprises the mutations T60K, D67K and Q75R will be referred from now on as SUMOvera.
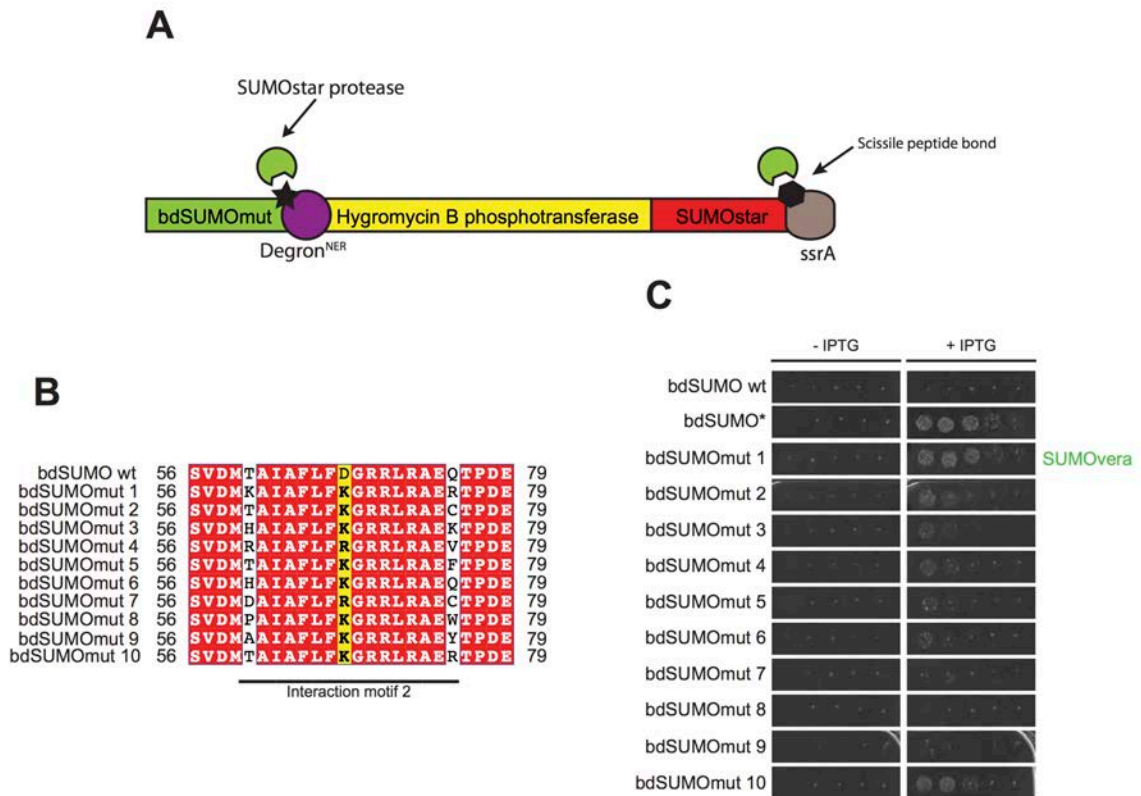
Figure 3.4 Selection of a bdSUMO mutant that is not recognized by SUMOstar protease. (A) Schematic representation of the chimeric fusion protein used to isolate bdSUMO mutants (bdSUMOmut) that are not cleaved by SUMOstar protease. **(B)** Three different residues (T60, D67 and Q75) in bdSUMO were randomized to create the bdSUMO mutant library. BdSUMO mutants were selected in medium containing Hygromycin. After selection, 10 different colonies were isolated and their sequences were analyzed by multiple protein sequence alignment. Identical and similar residues are highlighted in red and yellow, respectively. The bdSUMO wild type sequence (bdSUMO wt) was also included in the analysis. The interaction motif is indicated as in (Reverter & Lima 2004). The mutation D67K was extremely dominant in the selected bdSUMO mutants. **(C)** Cellular growth was tested individually for all 10 isolated colonies in the absence and presence of IPTG. Cells expressing the different bdSUMO mutants were diluted in 10-fold dilution series and spotted in medium containing Hygromycin B. A non-cleavable bdSUMO mutant (bdSUMO*) was used in a control sample to account for complete absence of bdSUMO recognition by SUMOstar protease. Among all samples, cells harboring bdSUMOmut 1 showed the highest growth after the overexpression of SUMOstar protease by addition of IPTG. bdSUMOmut 1 was named SUMOvera.

3.1.4   SUMOvera protease efficiently recognizes SUMOvera fusion proteins

As described in section 3.1 our main aim was to create a bdSUMO/bdSENP1 pair with orthogonal specificity to SUMOstar/SUMOstar protease. So far, we have already described

SUMOvera as a bdSUMO mutant that is not recognized by SUMOstar protease. To complete our aim, we next needed to obtain a bdSENP1 mutant (bdSENP1mut) that efficiently cleaves SUMOvera but not SUMOstar. To do so, we used our designed *in vivo* selection method (section 3.1.1). The construct used during the screen was SUMOstar-degron[NER]-Hph-SUMOvera-ssrA (Figure 3.5. A). SUMOstar was placed at the N-terminus of the Hph to select against protease cleavage, whereas SUMOvera was cloned at the C-terminus to select for efficient protein cleavage by a bdSENP1 mutant. In the presence of HygB, survival of bacteria is only possible when a given bdSENP1 mutant does not cleave SUMOstar but SUMOvera.

The bdSENP1 mutant library was created by randomization of the residues R269, N280, R346 and K350. These four amino acids are the most likely interacting partners of the mutant residues in SUMOvera. The library was then transformed into cells expressing the construct SUMOstar-degron[NER]-Hph-SUMOvera-ssrA and further plated on HygB-containing medium. After incubation of the plates, twenty different mutants were selected and tested individually using our designed selection system. The cellular survival in the presence of Hygromycin B was tested before and after induction of all 20 different bdSENP1 mutants (Figure 3.5. C). *E. coli* cells containing bdSENP1mut 5 managed to survive in selective medium even in the absence of IPTG. This result suggested that even trace amounts of bdSENP1mut 5 were enough to efficiently cleave SUMOvera and deactivate the ssrA signal. Moreover, over-expression of bdSENP1mut 5 after the addition of IPTG led to the highest survival rate of cells. All other over-expressed bdSENP1 mutants allowed bacterial growth to lower extent as compared to bdSENP1mut 5.

The sequences of all 20 different bdSENP1 mutants were analyzed by multiple sequence alignment (Figure 3.5. B). Although we did not observe clear dominant mutations among all 20 different colonies, analysis of the amino acid frequency of all bdSENP1 mutants showed that proline was the most frequent reside at first mutated position (N269P), small hydrophobic residues at second position (N280S, N280G), glutamic acid at third position (R346E), and valine, argenine or alanine at the fourth mutated position (K350V, K350R, K350A) (Figure 3.6). In fact, bdSENP1mut 5 contains two of the most frequent mutations at positions the second and third (N280S and R346E) which might explain its high efficiency for the cleavage of SUMOvera. Our experiments showed that bdSENP1mut 5 has the highest efficiency for the recognition of SUMOvera as a substrate among all 20

bdSENP1 mutants analyzed. We therefore decided to characterize bdSUMOmut 5 in more detail and named it SUMOvera protease.
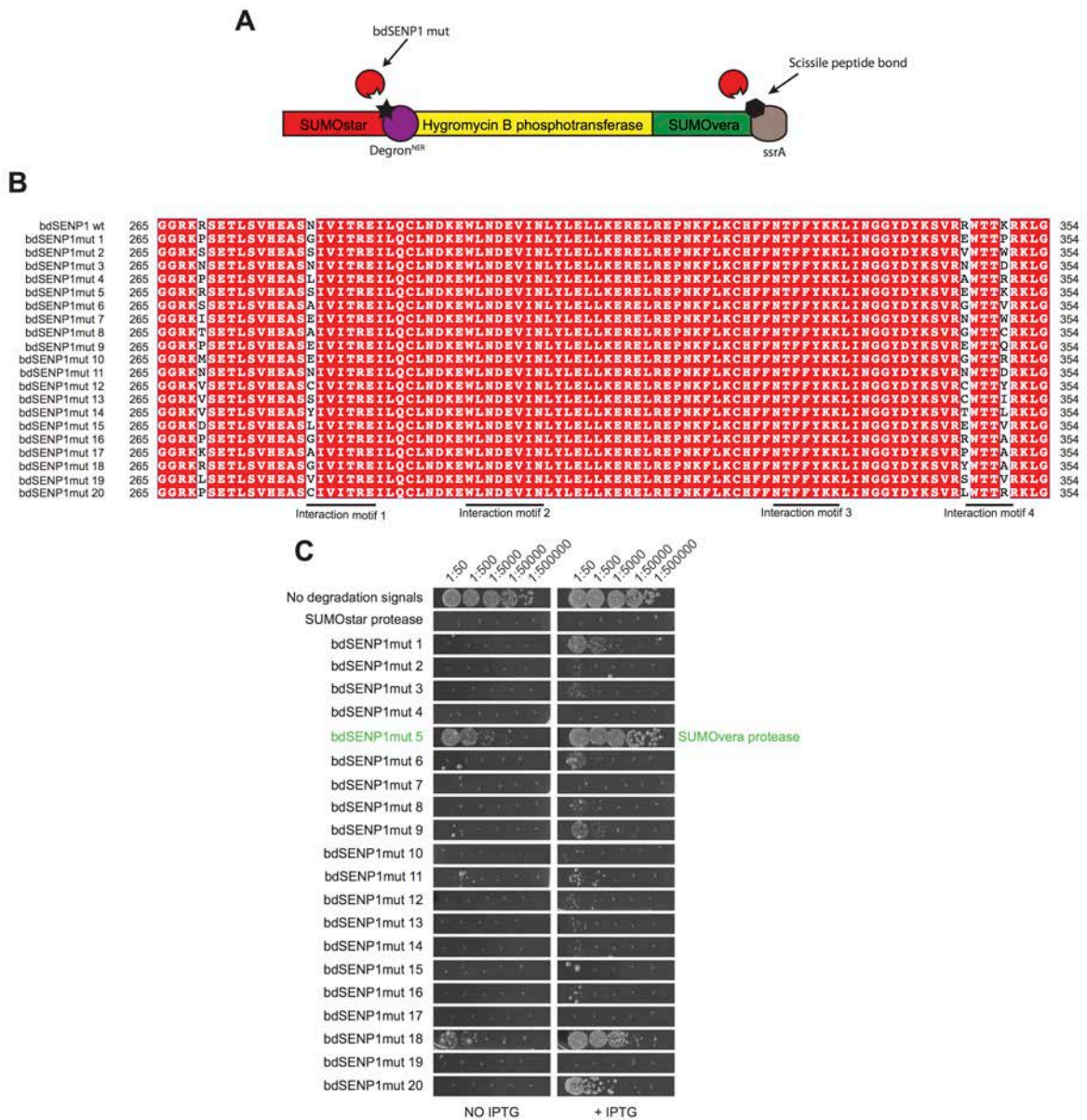
Figure 3.5 Selection and characterization of SUMOvera protease. (A) Scheme of the construct used to isolate bdSENP1 mutants (bdSENP1mut) that cleave only SUMOvera and not SUMOstar. SUMOvera was cloned at the C-terminus of Hph to select for its efficient recognition. In contrast, SUMOstar was placed at the N-terminus of Hph to select against substrate recognition by the bdSENP1 mutants. **(B)** The bdSENP1 mutants were created by randomization of four different residues (R269X, N280X, R346X and K350X) that are potentially interacting with the residues mutagenized in SUMOvera. Twenty different clones were isolated and their sequences were analyzed by multiple protein sequence alignment. Identical residues are highlighted in red. bdSENP1 wild type sequence (bdSENP1 wt) was included in the analysis. Interaction motifs between the substrate and the protease are designed as in (Reverter & Lima 2004). **(C)** The survival of all 20 isolated mutants was analyzed in plates with and without IPTG in the presence of HygB. Cells were diluted into a series of 10-fold dilutions further spotted onto the plates. Mutant bdSENP1mut 5 had the best growth behavior among all tested samples and therefore was named SUMOvera protease. A chimeric construct lacking both degradation signals was used as positive control. Cells expressing SUMOstar protease were used as a negative control sample.
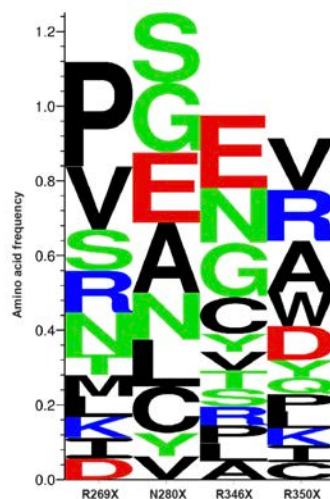
Figure 3.6 Amino acid frequencies in the four mutated positions in the bdSENP1 mutants. The x-axis in the sequence logo plot shows the four different mutated positions in bdSENP1 (R269X, N280X, R346X and K350X), whereas the y-axis represents the frequency of each residue in all 20 analyzed mutants.

3.2 In-depth characterization and applications of SUMOvera and SUMOvera protease

3.2.1 Two different sets of substrate/protease pairs with orthogonal substrate specificity

We tested the orthogonality of SUMOvera/SUMOvera protease pair and the SUMOstar/SUMOstar protease pair using an *in vitro* protein cleavage assay. We thus tagged the *E. coli* Maltose Binding Protein (MBP) at the N-terminus with either SUMOvera or SUMOstar (Figure 3.7. A). Both MBP fusion proteins were incubated for 1h at 4°C with increasing amounts of either SUMOvera protease or SUMOstar protease (Figure 3.7. B). SUMOvera-MBP was cleaved efficiently (≈95%) by SUMOvera protease at a protease concentration of 200nM. In contrast, SUMOstar protease did not recognize SUMOvera-MBP as a substrate even at a high concentration (10 µM). SUMOstar protease only cleaved SUMOstar-MBP fusion protein (≈95%) at a protease concentration of 100nM. SUMOstar-MBP was never recognized as a substrate by the SUMOvera protease even at the highest protease concentration of 10 µM. Note that 10 µM of SUMOvera protease and SUMOstar protease represent up to 1000-fold higher concentrations of protease needed for efficient cleavage of the cognate substrate and even so only protein cleavage of the cognate substrate took place. Therefore, these two proteases have truly orthogonal substrate specificity for SUMOvera and SUMOstar.
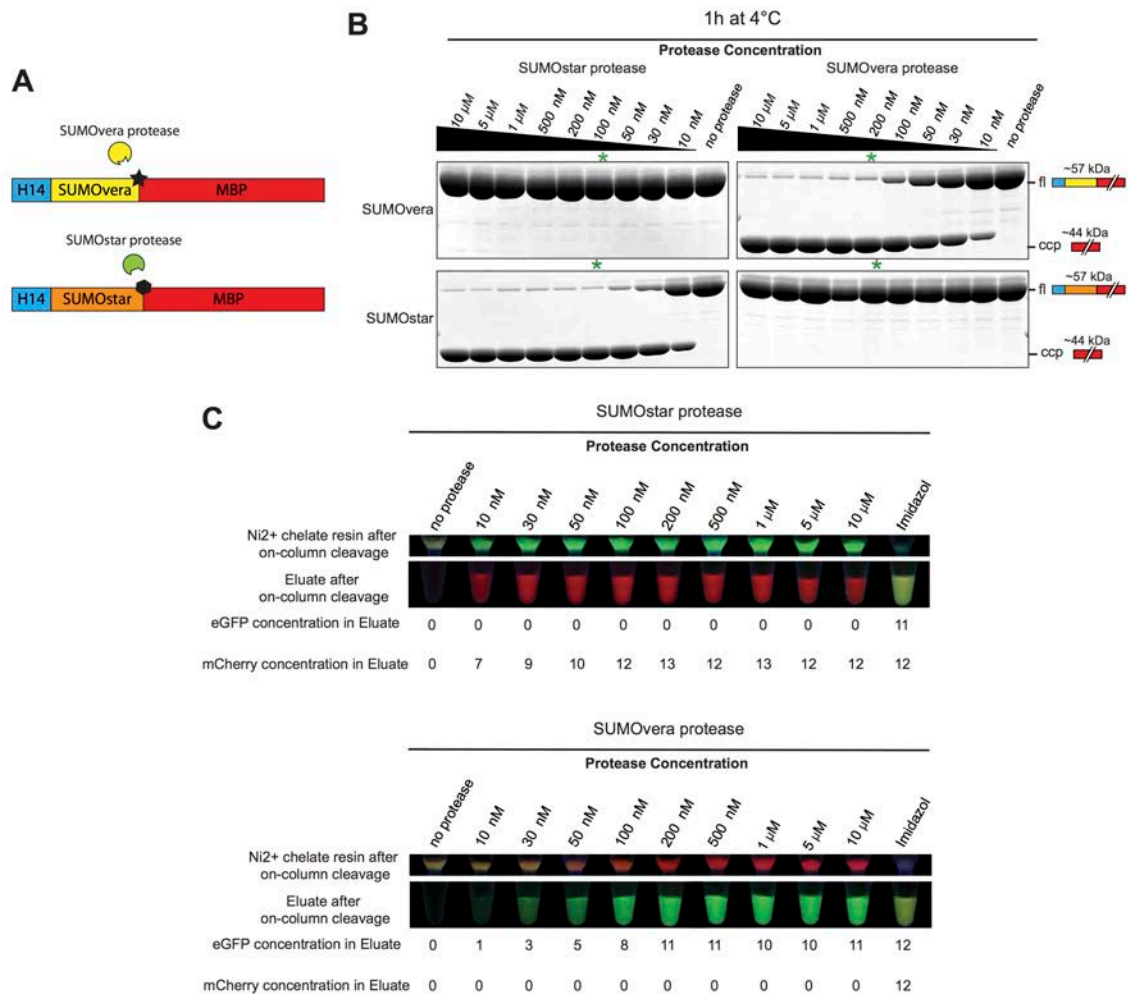
Figure 3.7 SUMOvera protease and SUMOstar protease have full orthogonal substrate specificities. (A) Graphic description of the substrates used to test protease cleavage by SUMOvera protease and SUMOstar protease in (B) (H_{14}: His-tag; MBP: maltose binding protein). **(B)** Samples with 100 μM of H_{14}-SUMOvera-MBP or H_{14}-SUMOstar-MBP were incubated with increasing amounts of SUMOvera or SUMOstar protease. Both proteases were titrated from 10 nM to 10 μM. A negative control with no protease was included in the analysis. The reactions were stopped by the addition of SDS-containing buffer. Samples were resolved by SDS-PAGE to visualize protein cleavage by the SUMO proteases. Green asterisks represent the minimal protease concentration where one substrate was ≈95% cleaved, whereas the other remained completely intact. Full-length fusion protein (fl) and C-terminal cleavage product (ccp). **(C)** On-column protein cleavage using SUMOvera protease and SUMOstar protease. Equal amounts of H_{14}-SUMOvera-eGFP and H_{14}-SUMOstar-mcherry fluorescent proteins were pre-loaded onto Ni^{2+} chelate beads. Samples with the loaded matrix were then incubated with increasing amounts of the specified SUMO-specific protease. The concentrations of eGFP and mCherry in the eluates were measured by absorption at 488 nm and 585 nm, respectively. Columns and tubes with the eluates were imaged while illuminated at 366nm. Protein elution was completely specific between the protease and its corresponding substrate. Two different control reactions, one without protease and the second with 400 mM imidazole, were included in the experiment.

We next wanted to investigate whether both proteases were able to cleave their cognate substrates even when these were immobilized on a solid support. We therefore performed a protein cleavage assay using immobilized SUMO-tagged fluorescent proteins. H_{14}-SUMOvera-eGFP and H_{14}-SUMOstar-mCherry were immobilized onto Ni^{2+}-chelate beads and incubated with different amounts of SUMOvera protease or SUMOstar protease (Figure 3.7. C). SUMOvera protease eluted only eGFP while mCherry remained

completely bound to the nickel matrix. In contrast, SUMOstar protease specifically eluted mCherry from the column whereas eGFP was never detected in the eluate even with a high concentration of SUMOstar protease. Therefore, these assays clearly showed that both proteases are perfectly suitable to be used for on-column protein cleavage. In addition, this experiment reassured that SUMOvera/SUMOvera protease and SUMOstar/SUMOstar protease pairs have true orthogonal specificity.

3.2.2   Analysis of the mutations in SUMOvera and SUMOvera protease

We obtained SUMOvera from a selected group of 10 different bdSUMO mutants that are not cleaved by SUMOstar protease (Figure 3.4). Sequence alignment of those mutants revealed that the mutation D67K was the only common feature shared between SUMOvera and all other 9 isolated mutants. The two other mutated residues in SUMOvera and the other bdSUMO mutants were variable. Since we obverse this variability, we sought to analyze in more detail the impact of every single mutation present in SUMOvera on the protein cleavage resistance by SUMOstar protease. To explore this, we purified 7 different bdSUMO-MBP fusion proteins containing one, two or three mutations present in SUMOvera. We incubated 100 µM of each MBP fusion protein for 2h at 25°C in the absence and presence of SUMOstar protease (Figure 3.8 B). In order to account for true protein cleavage resistance, we used a very high concentration of SUMOstar protease (10 µM) that represents a 2000-fold excess of SUMOstar protease needed to efficiently cleave the cognate substrate. Surprisingly, all 3 different MBP fusion proteins containing the mutation D67K remained as full-length proteins. In contrast, proteins containing only the mutations T60K and Q75R were completely cleaved by the SUMOstar protease. Mutations T60K and Q75R resulted in slight improvement on preventing protein cleavage by SUMOstar protease, but only when they were combined with the mutation D67K. Thus, these data suggest that mutation D67K is the main mutation responsible for preventing substrate cleavage by SUMOstar protease. Also, this protein cleavage assay explains why the high occurrence of mutation D67K in all 10 selected bdSUMO mutants to prevent protein cleavage by SUMOstar protease.

Subsequently, we decided to characterize the impact of the three mutant residues contained in SUMOvera (T60K, D67K and Q75K) on protein cleavage by SUMOvera protease. To do this, we incubated SUMOvera protease with the different bdSUMO-MBP fusion

proteins mentioned above (Figure 3.8. C). As we were looking for efficient protein cleavage of the different bdSUMO-MBP, we chose to use for the test the lowest amount the SUMOvera protease (200 nM) that is required for efficient cleavage of 100 μM SUMOvera-tagged proteins. The bdSUMO-MBP fusion protein containing all 3 mutations as in SUMOvera was the most efficiently cleaved by SUMOvera protease. In contrast, the MBP fusion protein containing the wild type bdSUMO sequence was the worst substrate for the SUMOvera protease. Also, the protease did not completely cleave fusion proteins containing just one or two mutations. We therefore can conclude that all 3 mutagenized residues in SUMOvera contribute to its efficient cleavage by the SUMOvera protease.
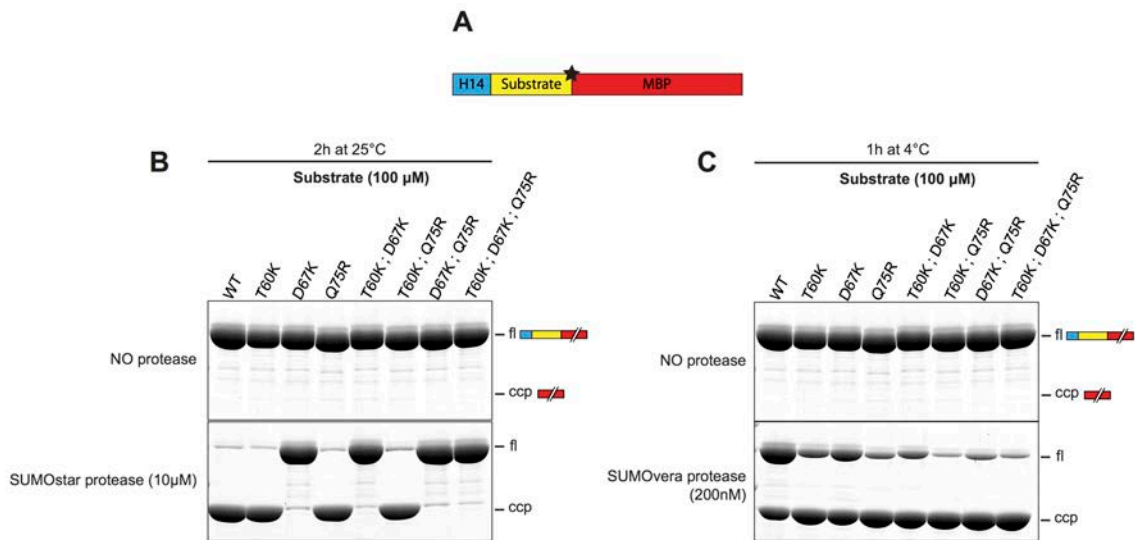
Figure 3.8 Effect of the mutations in SUMOvera during cleavage by SUMOstar protease and SUMOvera protease.
(A) Schematic representation of the fusion protein used in (B) and (C). (H$_{14}$: His-tag; substrate: bdSUMO variants having different number mutations present in SUMOvera; MBP: maltose binding protein). **(B)** Different bdSUMO-MBP fusion proteins containing one, two or three mutations as the ones located in SUMOvera were used for the protein cleavage assay. bdSUMO wild type sequence was also used in the analysis (WT). Samples containing 100 μM of the MBP chimeric protein were incubated with a large amount of SUMOstar protease (10 μM) for 2h at 25°C. Reactions were stopped by the addition of SDS-containing buffer and analyzed by SDS-PAGE. Mutation D67K had the biggest impact for complete protein cleavage resistance by SUMOstar protease. A sample without protease was used as control reaction. **(C)** The same chimeric proteins were incubated with 200nM SUMOvera protease for 1h at 4°C. After incubation, samples were treated and analyzed as in (B). SUMOvera protease cleaved the fusion protein with all three different mutations with the highest efficiency (≈95%) compared with all other samples. Full-length fusion protein (fl) and C-terminal cleavage product (ccp).
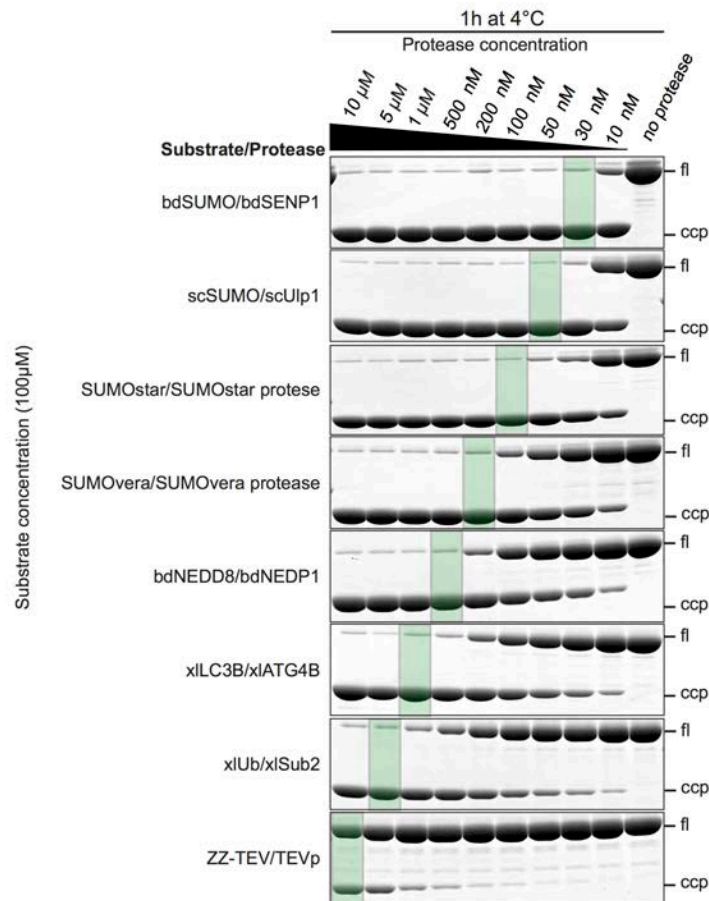
Figure 3.9 SUMOvera is a highly efficient protease compared to other commonly used proteases. The efficiency of different specific proteases (bdSENP1, TEV protease, scUlp1, bdNEDP1, SUMOstar protease, xlATG4B and xlSub2) was tested in solution for 1h at 4°C. Different amounts of a given protease (from 10nM to 10 μM) were incubated with 100 μM of its corresponding tagged substrate. After the reaction, the samples were resolved by SDS-PAGE and stained by Coomassie blue. Samples highlighted in green indicate the lowest protease concentration needed to cleave completely 100 μM of the analyzed substrate. The different experiments are ordered in the figure from the most to the least efficient protease under the same experimental conditions. Full-length fusion protein (fl) and C-terminal cleavage product (ccp).

3.2.3 SUMOvera protease is a highly efficient protease

We compared the cleavage efficiencies of SUMOvera protease to those of the commonly used proteases (*i.e.* bdSENP1, TEV protease, scUlp1, bdNEDP1, SUMOstar protease, xlATG4B and xlSub2). For each protease, we incubated a defined concentration of the substrate (100 μM) in a wide concentration range of the corresponding protease (Figure 3.9). As reported in a previous study, the two most efficient proteases were bdSENP1 and scUlp1 (Frey & Görlich 2014a). Around 15-50 nM of these two proteases were required to efficiently cleave 95% of the cognate substrate for 1h at 0°C. In comparison, 100nM of SUMOstar protease and 200nM of SUMOvera protease were enough to cleave 95% of the cognate substrate at the same conditions of incubation. Therefore, around 5-fold more SUMOvera and SUMOstar proteases were needed to cleave the same amount of cognate

substrate than bdSENP1 and scUlp1. In addition, SUMOvera protease was from 2.5-fold to 25-fold more efficient as compared to some other extensively used proteases (xlAtg4, xlSub2 and bdNEDP1). Interestingly, TEV-tagged substrate was only partially cleaved ($\approx$50%) even at a high concentration of TEV protease (10 μM). Therefore, TEV protease proved to be up to 1000-fold less efficient than SUMOvera protease. Moreover, experiments carried out at 25°C proved that 10-fold less protease is required to cleaved the same amount of substrate as in 4°C (data not shown). Together, these data suggest that SUMOvera protease represents a useful tool for tag removal, but also a more efficient protease compared to the most extensively used proteases such as TEV protease.

3.2.4 SUMOvera as a stable fusion protein tag in yeast and other eukaryotic systems

The presence of endogenous SUMO-specific proteases in eukaryotic cells hampers the possibility of using SUMO protein as a fusion tag. Recombinant SUMO-tagged proteins are immediately cleaved if they are expressed in eukaryotic hosts. As previously described, SUMOstar is the only known example of a SUMO protein used as a stable tag in eukaryotic cells (Liu et al. 2008; Peroutka et al. 2008). We thus decided to test whether SUMOvera would also be resistant to proteolytic cleavage by endogenous SUMO proteases in different eukaryotic hosts.

For this purpose, we first tested whether substrates tagged with scSUMO, SUMOstar, bdSUMO or SUMOvera could be expressed as full-length proteins in *S. cerevisae* (Figure 3.10. A). To this end, substrates were over-expressed in yeast for 6h at 30°C using a galactose-inducible expression system. The stability of the tagged substrates expressed in yeast was analyzed by western blot (Figure 3.10. B). As expected, scSUMO- and bdSUMO-tagged substrates were completely cleaved and no full-length product remained in the cells. In contrast, SUMOstar- and SUMOvera-tagged substrates were mostly expressed as full-length fusion protein and almost no C-terminal cleavage product was detected. Among all tested SUMO tag proteins, SUMOvera represented the most stable fusion tag during recombinant protein expression in yeast since up to $\approx$95% of the over-expressed SUMOvera-tagged protein remained as full-length. To our surprise, these results suggest that SUMOvera represents even a better choice than SUMOstar as a fusion protein in *S. cerevisiae.*
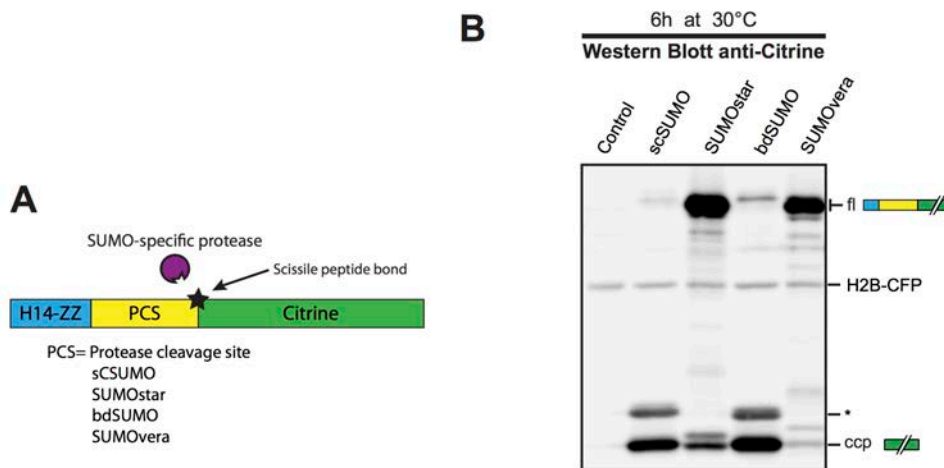
Figure 3.10 SUMOvera can be used to produce stable fusion proteins in *S. cerevisiae*. (A) Schematic representation of the substrates analyzed in B (H$_{14}$: His-tag; ZZ: Two copies of the Z-domain from Staphylococcal protein A; PCS: protease cleavage site). **(B)** scSUMO-, SUMOstar-, bdSUMO- and SUMOvera-tagged substrates were over-expressed for 6h at 30°C in *S. cerevisiae*. After expression, cell lysates were generated to analyze the stability of the fusion proteins by western blot. Samples were first separated by SDS-PAGE and further blotted. The blotted membrane was incubated with an anti-citrine primary antibody. A stably expressed cyan fluorescent protein fused to histone 2B (H2B-CFP) was used to confirm even loading of the samples. SUMOvera-tagged substrate was the most stable fusion protein expressed in yeast. Cross reactivity between the anti-citrine and the ZZ-tag is indicated by (*). Empty yeast lysate was also blotted as a negative control. Full-length fusion protein (fl) and C-terminal cleavage product (ccp).

We next decided to analyze whether scSUMO-MBP, SUMOstar-MBP, bdSUMO-MBP and SUMOvera-MBP fusion proteins were stable in different eukaryotic extracts (Figure 3.11. A). For this, we incubated separately 1 μM of each tagged MBP into different highly concentrated eukaryotic lysates for 2h at 25°C and then analyzed by western blot (Figure 3.11. B). We included samples containing a "protease mix" (scUlp1, SUMOstar, bdSENP1 and SUMOvera protease, 1 μM each) in order to control for any SUMO-specific protease inhibitory substance present in the extracts. Substrates tagged with scSUMO or bdSUMO were cleaved to a different extent in wheat germ extract, *X. laevis* egg extract, HeLa cell extract and Drosophila S2 cell extract. In contrast, substrates tagged with either SUMOstar or SUMOvera were highly stable in all cellular extracts tested. Remarkably, all SUMO-tagged substrates remained as full-length proteins in rabbit reticulocyte extract and LTE Lexsy cell extract (Mureev et al. 2009; Kovtun et al. 2010). Samples including the protease mix confirmed the absence of any protease inhibitory substance in all analyzed lysates. These results confirm that SUMOstar and SUMOvera are excellent choices as stable tag proteins to be used in many different eukaryotic extracts.
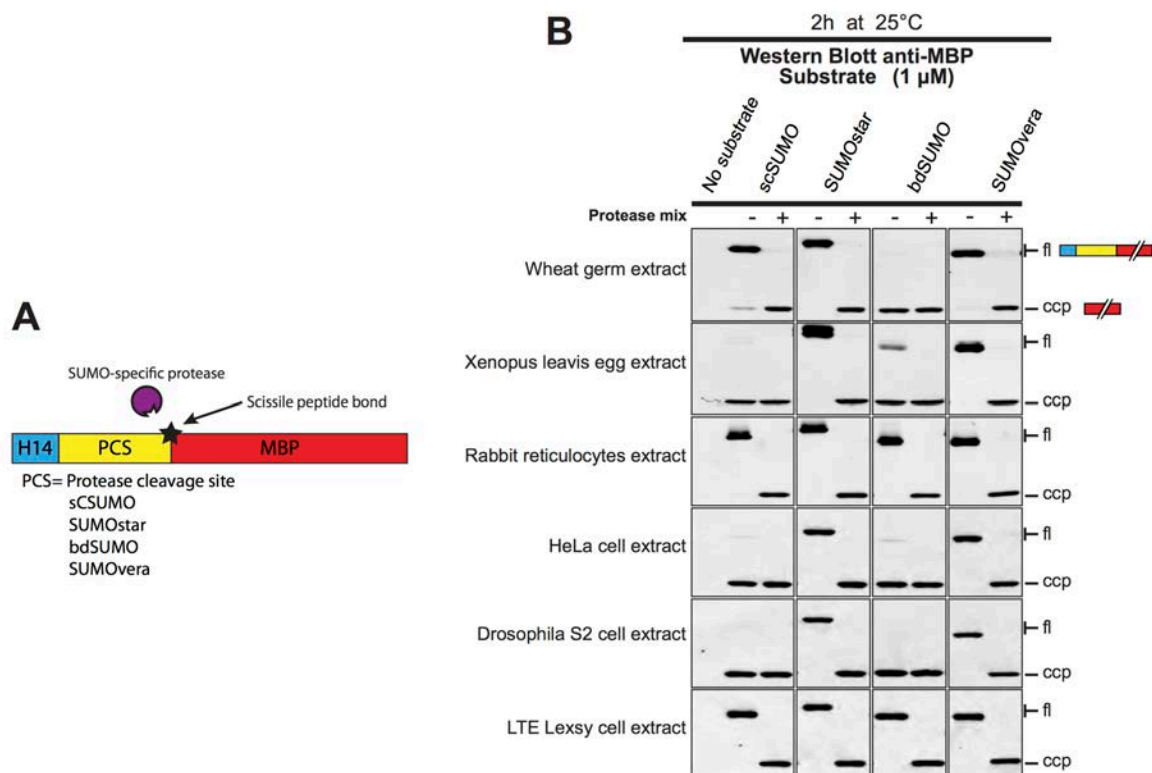
Figure 3.11 SUMOvera is a stable fusion partner in many eukaryotic lysates. (A) Graphic description of the substrates used to test fusion protein stability in (B) (H$_{14}$: His-tag; PCS: protease cleavage site; MBP: maltose binding protein). **(B)** Stability of different SUMO-tagged proteins in several eukaryotic extracts. 1μM of scSUMO-, SUMOstar-, bdSUMO- and SUMOvera-tagged MBP were incubated for 2h at 25°C in different highly concentrated eukaryotic lysates. The stability of the substrates in the extracts was analyzed by western blot using an anti-MBP primary antibody. The presence of a C-terminal cleavage product (ccp) indicated the cleavage of the tagged MBP by endogenous SUMO-specific proteases in the extracts. Samples with a protease mix (scUlp1, SUMOstar protease, bdSENP1 and SUMOvera protease; 1μM each) were used to rule out the present of any SUMO-specific protease inhibitory substance in the lysates. Negative controls for each cellular extract without the addition of the tagged MBP were included in the analysis. SUMOvera- and SUMOstar-tagged MBP remained as stable full-length (fl) fusion proteins in all analyzed extracts. Surprisingly, all fusion proteins were not cleaved in rabbit reticulocyte and LTE Lexsy cell extracts.

3.2.5 SUMOvera protease and SUMOstar protease can be used to purify protein complexes with defined stoichiometry in yeast

Sets of different proteases with orthogonal substrate specificity have been exploited to purify protein complexes with defined subunit stoichiometry. However, this technology has been applied only for protein complexes expressed in bacteria (Frey & Görlich 2014b). Here, we wanted to provide a proof of principle that SUMOstar protease and SUMOvera protease can be used to purify protein complexes over-expressed in *S. cerevisiae*. To this end, we selected the high affinity complex composed of the anti-GFP nanobody known as "enhancer" (Kirchhofer et al. 2010) and the GFP-like protein named citrine (Heikal et al. 2000) (Figure 3.12. A).
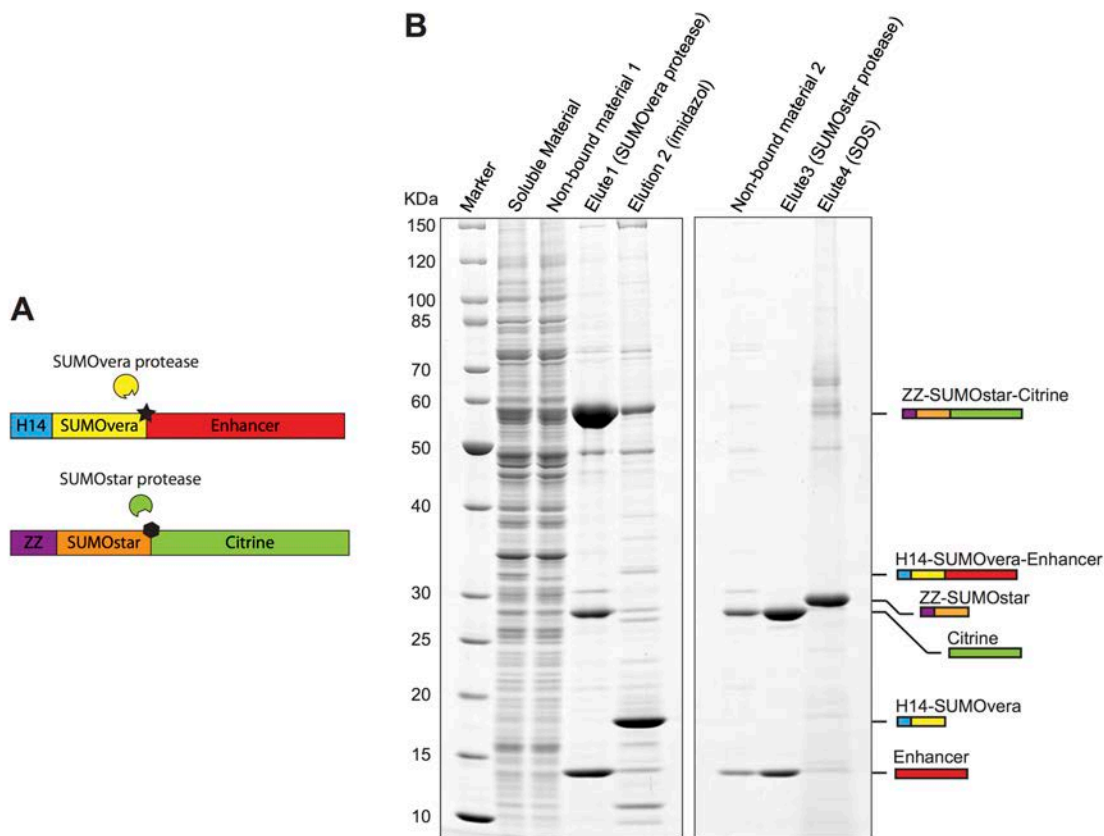
Figure 3.12 Purification of a stoichiometric protein complex in *S. cerevisiae*. (A) Graphic description of the proteins purified in (B). (H₁₄: His-tag; ZZ: Two copies of the Z-domain from Staphylococcal protein A). **(B)** A similar purification strategy as described in (Frey & Görlich 2014b) was used to purify the binary complex between citrine and the anti-GFP nanobody named "enhancer". H₁₄-SUMOvera-enhancer and ZZ-SUMOstar-citrine were co-expressed in *S. cerevisiae* for 8h at 30°C (soluble material). First, the binary complex was isolated from the yeast extract by using a Ni²⁺ chelate matrix. Elution of the binary complex was achieved by the incubation of the nickel matrix with buffer containing SUMOvera protease (elute 1). The elution of the His-tag and many other protein contaminants left in the column was done by imidazole elution (elution 2). At this point, we had a binary complex with certain excess of enhancer over citrine. For the second chromatographic step, the sub-stoichiometric complex was loaded onto a matrix coupled to anti-Z domain affibody. The 1:1 stoichiometric and highly pure complex was finally achieved after incubation of the matrix with buffer containing SUMOstar protease (eluate 3). Protein samples corresponding to 35mOD of cells or 1/1000 of the total purified protein preparation were analyzed by SDS-PAGE and stained with coomassie blue.

We co-expressed the binary complex in *S. cerevisiae* for 6h at 30°C under control of the *GAL1* promoter. Citrine was expressed with a N-terminal ZZ-SUMOstar tag, whereas enhancer contained a N-terminal H₁₄-SUMOvera tag. We used two consecutive capture-and-realize chromatographic steps, just as described in (Frey & Görlich 2014b). First, the excess of citrine, if present, was removed during the first chromatographic step using a Ni²⁺ chelate matrix (non-bound material 1). The complex was then eluted by on-column protein cleavage using SUMOvera protease (eluate 1). Notably, untagged citrine was present in the eluted complex due to partial cleavage of the ZZ-SUMOstar tag by endogenous proteases during protein expression (Figure 3.10). In a second chromatographic step, free enhancer and untagged citrine where washed off from the anti-Z domain matrix (non-bound material 2). Thus, only the critine•enhancer complex

remained bound to the matrix. The 1:1 stoichiometric complex was obtained on-column protein cleavage with the SUMOstar protease (elute 3). In order to restore the anti-Z domain matrix, the ZZ-SUMOstar tag was eluted from beads by the addition of SDS-containing buffer. Altogether, we described a simple two-step purification protocol to obtain an extremely pure and tag-free binary complex. To our knowledge, this represents the first practical example of the purification of a protein complex expressed in a eukaryotic host by using two different orthogonal SUMO-specific proteases.

3.2.6   SUMOvera protease can be expressed *in vivo* in *S. cerevisiae*

The removal of protein tags by specific proteases is typically performed *in vitro* at specific temperature and ionic conditions. However, site-specific proteolysis of fusion proteins by several proteases has been used in living cells for different biochemical assays (Chen et al. 2010; Harder et al. 2008; Sato & Toda 2007). The cleavage of a fusion protein *in vivo* implies that a specific protease is being ectopically expressed at a given time. Here, we decided to test whether over-expression of SUMOvera protease in *S. cerevisiae* could be achieved without interfering with cell viability (Figure 3.13). Yeast cells were transformed with a plasmid encoding for scUlp1, bdSENP1, SUMOstar protease or SUMOvera protease. As negative control, we used cells containing a plasmid encoding a bdSENP1 mutant that is completely inactive (C440S). Proteases were over-expressed for 72h at 30°C under the control of the galactose-inducible *GAL1* promoter. To our surprise, only *S. cerevisiae* cells expressing SUMOvera protease grew after the addition of galactose. In contrast, expression of scUlp1, bdSENP1 or SUMOstar protease caused complete cellular death. Such lethal phenotype could be easily explained by massive de-sumoylation of essential proteins caused by over-expressed proteases.

We showed here that SUMOvera protease is a SUMO-specific protease suited to perform site-directed proteolysis in living yeast cells. In addition, SUMOvera protease represents the second example besides TEV protease of a highly specific protease used for *in vivo* protein cleavage in yeast (Gruber et al. 2003; Higuchi & Uhlmann 2005).
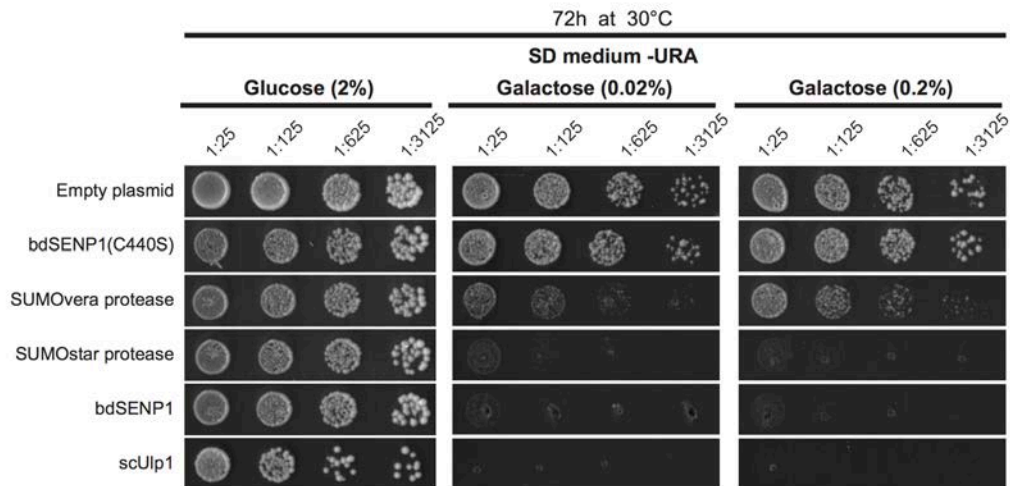
Figure 3.13 Over-expression of SUMOvera protease is possible in *S. cerevisiae*. Multicopy plasmids encoding SUMOvera protease, SUMOstar protease, scUlp1 and bdSENP1 were transformed in *S. cerevisiae*. Transformed cells were tested for cellular viability after protease over-expression driven by the addition of galactose. Yeast cells were subjected to a 10-fold dilution series, from 1:25 to 1:3125 (v/v). Then, 5 μl of the diluted samples were spotted on SD -URA medium containing either glucose (2%; v/v) to repress protease over-expression or with increasing amounts of galactose (0.02% and 0.2%; v/v). Cells transformed with either an empty plasmid or a plasmid encoding for a catalytically dead protease mutant (bdSENP1 C440S) were used as controls. Only the over-expression of SUMOvera protease was well tolerated by yeast cells. Expression of SUMOstar protease, scUlp1 and bdSENP1 caused cellular death.

3.2.7   Novel tandem affinity purification strategy

Tandem Affinity Purification (TAP) is a technique employed to isolate protein complexes from native cellular sources (Rigaut et al. 1999). TAP comprises two sequential affinity chromatographic steps to ensure low binding background. The bait protein is normally fused to two different affinity tags with a protease cleavage site in between them. The protease cleavage site allows native-like elution conditions from the first affinity column by addition of a protease. Then, protein complexes are bound to a second affinity column and eluted by competitive elution. The identity of the isolated proteins forming the complexes is normally revealed by mass spectrometry (MS) (Kaiser et al. 2008).

Tap is nowadays a well-established, rather simple and sensitive enough technique. However, TAP still has some shortcomings that need to be optimized. For instance, the technique relies on using the low efficient TEV protease during the first chromatographic step. A second disadvantage is the enormous concentration of bait protein compared to the concentration of preys in the final sample could hamper protein identification by MS. Also, the dilution of the final sample due to competitive elution might prevent the visualization of low-abundant binders during SDS-PAGE analysis. Third, tedious preparative steps are needed for the sample to be processed by MS. We thus designed a new TAP strategy to overcome all these limitations in order to identify protein-protein interactions in complex

41

biological samples. We therefore decided to tackle these disadvantages and develop a new TAP technology.
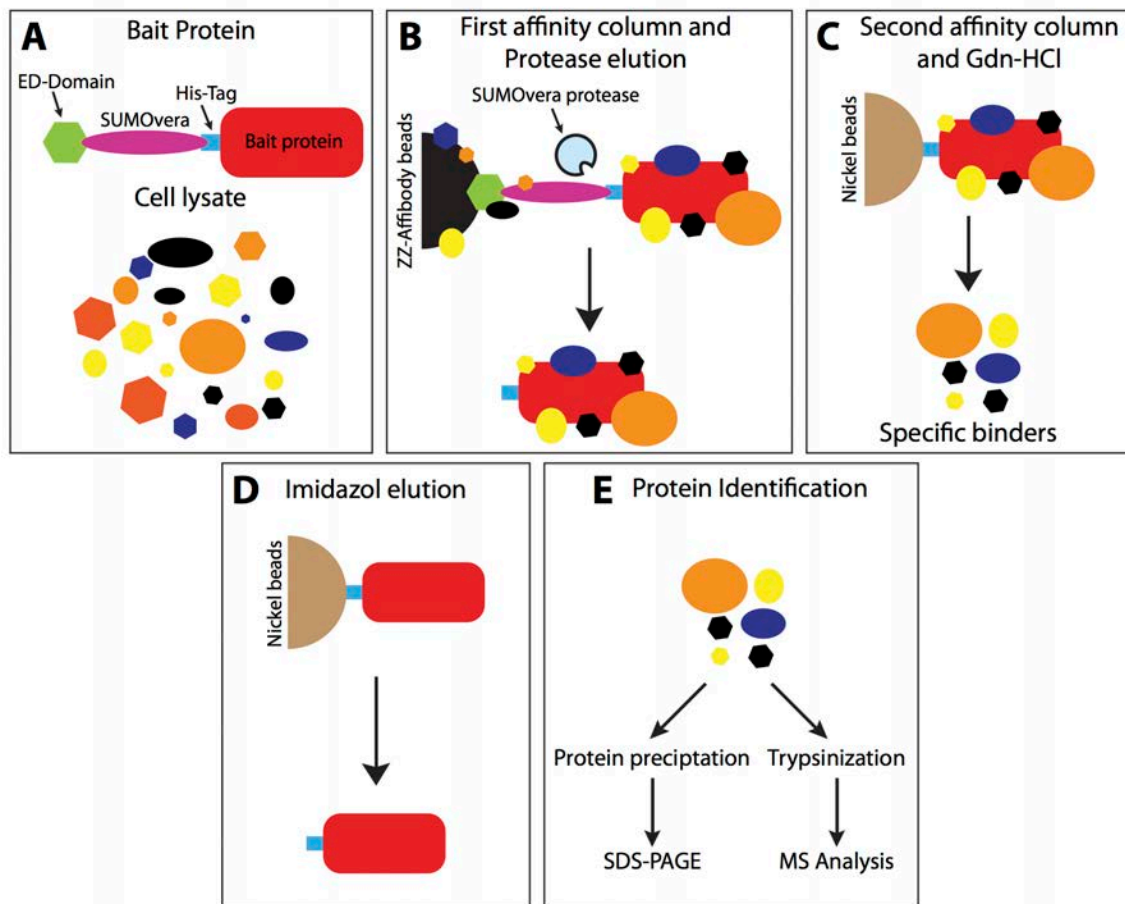
Figure 3.14 Novel tandem affinity purification technique. (A) Schematic representation of the bait and the prey used in the tandem affinity purification (TAP). The bait protein is N-terminally fused to two different affinity tags. The first tag consists of the E and D domains from the staphylococcal protein A (ED-domain). The second affinity tag is a standard His-tag (H_{12}). SUMOvera is the cleavage site that separates the two affinity tags. Therefore, ED-SUMOvera-H_{12} tag represents the TAP-tag module. The TAP-tagged bait is then incubated with cell lysate and it is during this process where the bait•prey complexes are being formed. **(B)** In the first step of the process, the complexes are trapped using an anti-Z domain affibody matrix. After washing unspecific binders away, protein complexes are specifically eluted from the column by on-column protein cleavage with the SUMOvera protease. **(C)** Further purification of the protein complexes is achieved by a second affinity purification step using a Ni^{2+} chelate matrix. Note that protein elution in this second purification is done with basic guanidinium hydrochloride (Gdn-HCl). This allows that the bait remains bound to the nickel matrix whereas the specific binders are eluted. **(D)** Finally, the bait protein can be recovered from the Ni^{2+} column by imidazole elution. **(E)** Protein identification can be done immediately by mixing directly the entire eluate with trypsin followed by mass spectrometry analysis. Alternatively, the eluted sample can be subjected to protein precipitation and further resolved by SDS-PAGE for the identification of specific protein bands.

For the design of our novel TAP protocol, SUMOvera was considered to show its proper functionality as a stable fusion tag when incubated in eukaryotic lysates. We thus decided to use the fusion protein ED-SUMOvera-H_{12} to tag the bait protein (Figure 3.14. A). The E and D domains from the staphylococcal protein A (ProtA) allow the immobilization of the bait to a matrix coupled with the anti-Z domain affibody (ZpA963). Binding of the ED tag to the anti-Z domain affibody is the first highly affinity chromatography step in the process

for the initial isolation of preys from the cell extract (Figure 3.14. B). ZpA963 was originally evolved to bind the synthetic Z domain derived from ProtA (Nilsson et al. 1987; Lindborg et al. 2013). However, we preferred to use the E and D domains from ProtA as an affinity tag because they have a higher binding affinity to ZpA963 as compared to the synthetic Z domain. Then, SUMOvera is used as the protease cleavage site to perform protein elution from the ZpA963-coupled column by addition of SUMOvera protease (Figure 3.14. B). The use of SUMOvera and SUMOvera protease brings some advantages to our new TAP strategy. First, efficient protein elution can be achieved in less than 1 h at 4°C with a nanomolar concentration of SUMOvera protease. Second, the SUMOvera-tagged bait can be incubated virtually in any cellular extract due to its resistance towards protein cleavage by endogenous eukaryotic SUMO-specific proteases. Additionally, higher purity of the eluted protein is achieved since proteins that bind unspecifically to the matrix remain bound after on-column protein elution mediated by SUMOvera protease.

For the second chromatographic step, a poly-histidine tag (H_{12}-tag) is employed to bind the eluted protein complexes to a Ni^{2+} chelate matrix (Figure 3.14. C). Protein elution is then performed suing a buffer containing 3M guanidinium hydrochloride (Gnd-HCl) and basic pH (Figure 3.14. C). Gnd-HCl causes protein unfolding and hence elution of the untagged binders from the nickel column, while the unfolded H_{12}-tagged bait remains tightly bound to the nickel column. In the last step of TAP, the His_{12}-tagged bait protein can be easily retrieved by imidazole elution from the Ni^{2+} chelate column (Figure 3.14. D).

Protein elution by Gnd-HCl brings some unique and novel aspects to our designed TAP methodology in terms of protein identification by MS. First, the complete separation of specific binders from the highly concentrated bait protein eases subsequent protein identification of low-abundant protein binders by MS. Second, samples containing Gnd-HCl are fully compatible with different proteolytic methods used in MS (*i.e.* trypsination). Therefore, buffer exchange is no longer needed prior to protein identification by MS (Poulsen et al. 2013; Saveliev & Ph 2013). Also, protein elution with Gnd-HCl eliminates the need of performing in-gel digestion procedures that might hamper protein identification of complex samples (Morris et al. 2014).

Another important aspect of our TAP system to be mentioned is the incompatibility between Gnd-HCl containing samples and SDS-PAGE. SDS precipitates immediately in the presence of Gnd-HCl. Thus, the Gnd-HCl must be removed from the protein sample before SDS-PAGE is performed. Protein precipitation with isopropanol or other similar

compounds represents an easy method to completely remove Gnd-HCl from the samples. Including a protein precipitation step in our TAP technique might represent an advantage rather than a pitfall. However, protein precipitation allows the concentration of low-abundant proteins up to several orders of magnitude and only then they can become visible after SDS-PAGE.

Altogether, we offer a rapid, very specific, MS-compatible and highly sensible TAP technique, and we assume that our novel TAP strategy represents an attractive alternative to conventional protocols.

3.3 Identification and characterization of novel Lph2 transport cargos

3.3.1 Lph2 as a novel bidirectional NTR in *S. cerevisiae*

Next, our novel TAP protocol and the SUMOvera/SUMOvera protease system were used to search for novel transport cargos of the poorly studied yeast nuclear transport receptor Lph2 (Kap120). The 60S ribosomal subunit assembly factor Rpf1 was the first identified import cargo for Lph2 (Caesar et al. 2006). This work showed for the first time that Kap120 acts as an importin in S. *cerevisiae*. So far, only two other Lph2 import cargos have been identified, the Ho endonuclease and the transcriptional regulator Swi6 (Kim et al. 2010; Bakhrat et al. 2006). To our knowledge, all published studies that identified Lph2 binding partners focused only in finding potential import cargos but not on the discovery of possible export cargos. We therefore hypothesized that there was still a high possibility of identifying novel export cargos for Lph2. Moreover, Lph2 has a medium affinity to Ran-GTP (270 nM) (Hahn & Schlenstedt 2011) that seems to be a distinct feature of bidirectional NTR.

In order to prove our hypothesis, we decided to search for novel export and import cargos for Lph2 using *S. cerevisiae* lysate (Figure 3.15, lane 1). To do so, we tagged full-length Lph2 at the N-terminus with the ED-SUMOvera-H_{12} fusion tag (Figure 3.15, lane 2). Gsp1-GTP, the yeast homologue of Ran-GTP, was included in the TAP experiment to isolate potential export cargos (Figure 3.15. lane 3). Specifically, we used a His_{12}-tagged Gsp1 mutant (H_{12}-Gsp1Q71L-GTP) that is unable to hydrolyze GTP and therefore remains trapped in the GTP-bound state (Bischoff et al. 1994; Klebe et al. 1995).

For the TAP experiment, samples were first incubated with a anti-Z domain affibody matrix in the absence and presence of Gsp1Q71L-GTP for 1h at 4°C. After incubation, protein complexes were eluted by on-column protein cleavage with SUMOvera protease (Figure 3.15, lanes 4 and 5).

The eluted samples were incubated with nickel matrix for 1h at 4°C for the second affinity chromatography step. After incubation, only protein binders were eluted from the nickel matrix by a guanidinium hydrochloride (Gdn-HCl) containing buffer (Figure 3.15, lane 6 and 7). The His$_{12}$-tagged Lph2 and Gsp1Q71L-GTP remained bound to the nickel beads and were eluted only after the addition of imidazole containing buffer (Figure 3.15, lanes 8 and 9). All eluted samples were resolved by SDS-PAGE and specific bands were isolated from the polyacrylamide gel for subsequent protein identification by MS analysis. MS analysis led to the identification of eIF4A, Kre11 and Rvs167 as potential Lph2 export cargos. Active transport to the cytoplasm of eIF4A, Kre11 and Rvs167 by Lph2 seems to be in accordance with their functions as these export cargos are solely involved in cellular processes either at the cytoplasm or the plasma membrane. The eukaryotic translation initiation eIF4A is a component of the eIF4E complex that mediates the recruitment of the 40S ribosomal subunit to the initiation codon at the mRNA (Gingras et al. 1999). Specifically, eIF4A is a helicase that uses its ATPase activity to unwind the secondary structure of the 5' leader sequence in the mRNA at promote translation initiation (Rozen et al. 1990). Kre11 is one of the three subunits of the guanine nucleotide exchanger TRAPP, a vesicle tethering complex involved in the late Golgi (Sacher et al. 2000; Liang et al. 2007). Rvs167 forms part of the cytoskeleton structure that is required for the formation of vesicles at the plasma membrane level (Friesen et al. 2005). Additionally, we identified the 60S ribosomal protein L12A (rpL12A) as a potential import cargo for Lph2 as well as the already described assembly factor Rpf1 (Caesar et al. 2006).

Identification of Gsp1-dependent and -sensitive Lph2 protein-binders is a strong indication that Lph2 might act not only as an importin but also as a exportin in *S. cerevisiae.* Therefore, Lph2 might represent the second identified bidirectional NTR after Msn5 (Kaffman, Rank, O'Neill, et al. 1998; Yoshida & Blobel 2001). We chose eIF4A and rpL12A for further characterization in order to demonstrate that Lph2 is indeed and bidirectional NTR in budding yeast.
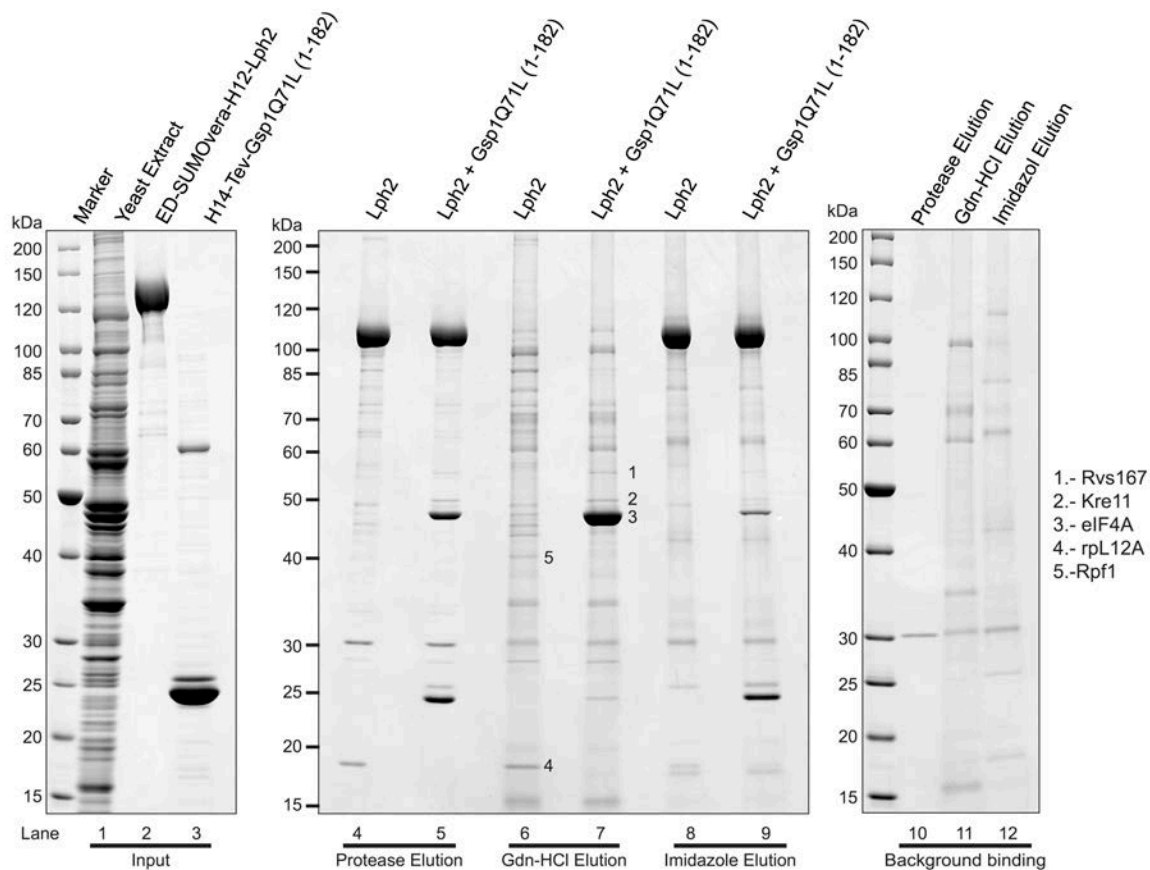
Figure 3.15 Lph2 is a bidirectional NTR in *S. cerevisiae*. ED-SUMOvera-H$_{12}$-Lph2 fusion protein (lane 2) was immobilized onto anti Z-domain matrix to isolate transport cargos from yeast extract (lane 1). TAP binding reactions were performed in the presence and absence Gsp1Q71L-GTP (lane 3). After incubation of the matrix with the yeast lysate, bound proteins were eluted with SUMOvera protease (lanes 4 and 5). Different interacting proteins were clearly visible by direct comparison between the samples with and without Gsp1Q71L-GTP. Eluted samples were further incubated with Ni^{2+} chelate matrix for the second chromatographic step. Elution of potential Lph2 cargos was done by the addition of 3M guanidinium hydrochloride buffer (lanes 6 and 7). His-tagged Lph2, Gsp1Q71L-GTP and remaining contaminants were finally eluted from the nickel beads with 400 mM imidazole (lanes 8 and 9). A control sample lacking ED-SUMOvera-H$_{12}$-Lph2 was used to analyze non-specific binding of yeast proteins from the lysate during the whole assay (lanes 10, 11 and 12). Protein samples were separated by SDS-PAGE and stained with colloidal coomassie blue. Single protein bands were cut out from the polyacrylamide gel and analyzed by mass spectrometry for protein identification. The Lph2 transport cargos identified are; Export cargos: Rvs167 (1), Kre11 (2) and eIF4A (3); Import cargos: rpL12A (4) and Rpf1 (5).

3.3.2 Lph2 mediates the nuclear export of eIF4A

As seen in Figure 3.15, the yeast translation initiation factor eIF4A was identified as a potential Lph2 export cargo. eIF4A was able to bind Lph2 in a Gsp1Q71L-GTP dependent manner. However, this experiment still left an open question whether a protein factor present in the yeast extract aided in the interaction between Lph2 and eIF4A. We thus decided to characterize the formation of the Lph2•eIF4A•GspQ71L export complex using purified recombinant components.

To this end, Lph2 was immobilized onto anti-Z domain affibody matrix via a N-terminal ED-SUMOvera tag. The binding of eIF4A to Lph2 was tested with and without the

addition of untagged Gsp1Q71L-GTP (Figure 3.16, A. Input). As a negative control, we included a sample without Lph2 to test for unspecific binding of eIF4A to the matrix. Samples were eluted with a buffer containing 200nM of SUMOvera protease for 1h at 4°C. After elution, it was clear that eIF4A formed a stable export complex with Lph2 in the presence of Gsp1Q71L-GTP (Figure 3.16, A. Bound fraction). In contrast, eIF4A failed to bind Lph2 in the absence of GspQ71L. These experiments clearly show that no extra protein factor is needed for Lph2 to form a stable export complex together with eIF4A and Gsp1Q71L-GTP, Moreover, we prove again that Lph2 might behave as the true corresponding exportin for eIF4A in yeast since cargo binding becomes efficient only upon the presence of Gsp1Q71L-GTP.

We next wanted to analyze the interaction between eIF4A and Lph2 *in vivo*. Thus, we used confocal immunofluorescence confocal microscopy to visualize the subcellular localization of GFP-labeled eIF4A in living wild type and Lph2-knockout yeast cells (ΔLph2). Yeast-codon optimized GFP (yeGFP) was inserted in-frame into wild type and ΔLph2 cells to tag eIF4A at the C-terminus (eIF4A-yeGFP). We used tetrameric cherry protein fused to a Crm1 nuclear export signal (tcherry-NES) as a nuclear exclusion marker (Xu et al. 2012; Güttler et al. 2010). We observed a perfect co-localization of eIF4A-yeGFP with the nuclear exclusion marker in wild type cells (Figure 3.16, B). Thus, wild type yeast cells expressing eIF4A-yeGFP showed perfect nuclear exclusion of eIF4A. In contrast, the single deletion of Lph2 in yeast led to a dramatic nuclear accumulation of eIF4A-yeGFP (Figure 3.16, B). These findings unquestionably prove that Lph2 is the export factor for eIF4A in *S. cerevisiae*. Remarkably, the active export of eIF4A is clearly the first evidence to confirm that Lph2 acts not only as importin but also as an exportin. Lph2 therefore needs to be regarded as a dedicated bidirectional NTR in *S. cerevisiae*.
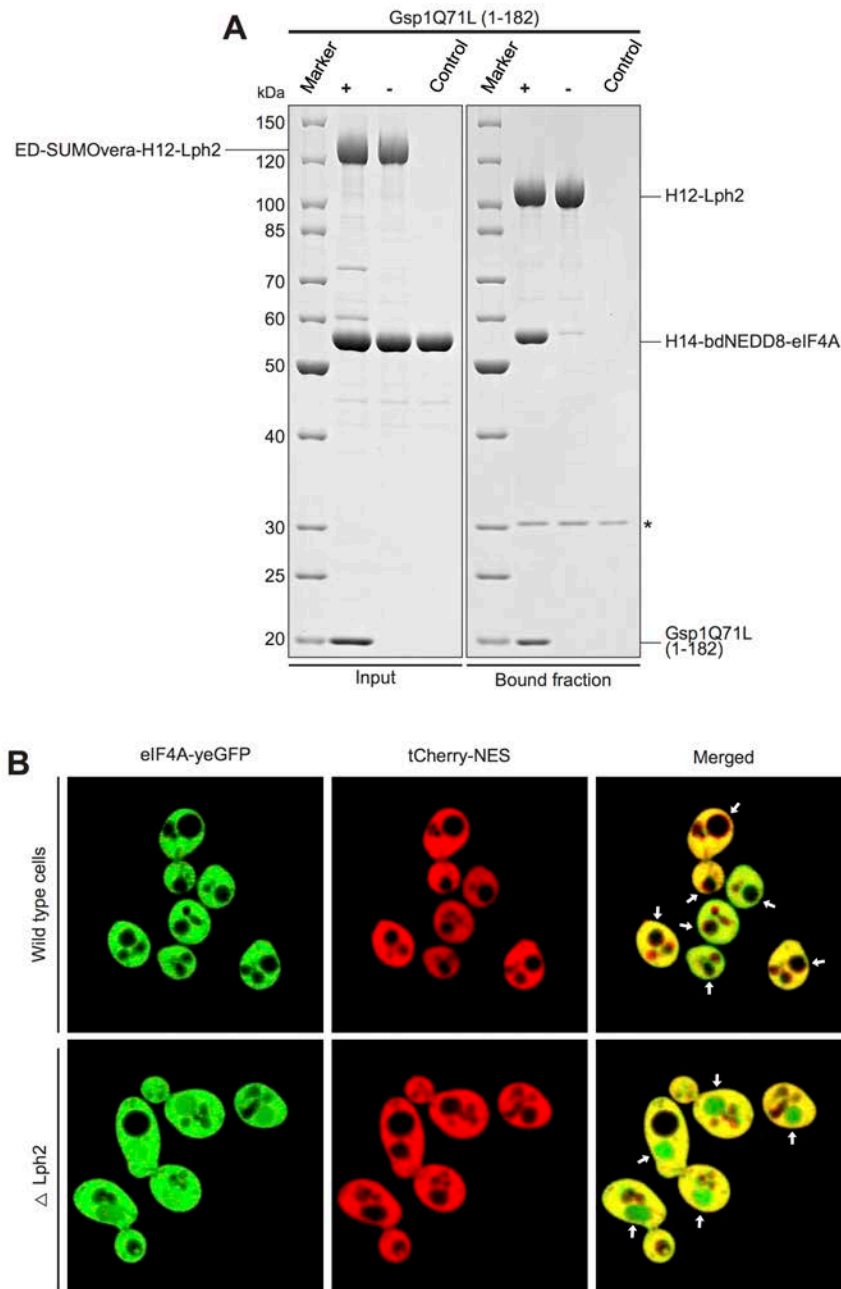
Figure 3.16 Lph2 exports eIF4A out of the yeast nucleus. (A) Characterization of the Lph2•eIF4A•Gsp1Q71L-GTPexport complex using recombinant proteins. Purified ED-SUMOvera-H_{12}-Lph2 (1.5 μM) was incubated with eIF4A (5 μM) in the presence and absence of Gsp1Q71L-GTP (5 μM) (input fraction). After incubation, ED-SUMOvera-H_{12}-Lph2 was first bound to anti-Z domain matrix and later eluted with SUMOvera protease in order to analyze protein-protein interaction (Bound fraction). A control sample was included in the analysis to test unspecific binding of eIF4A to the matrix. (*) indicates the SUMOvera protease used for on-column cleavage. Samples corresponding to 1.5μg of Lph2 were resolved by SDS-PAGE and stained with coomassie brilliant blue. Molecular weights for the bands of the protein marker used are labeled accordingly. **(B)** Localization of eIF4A was analyzed by confocal fluorescent scanning microscopy in living wild type and Lph2-knockout (ΔLph2) yeast cells. Yeast-codon optimized GFP (yeGFP) was integrated in-frame by homologous recombination at the C-terminus of eIF4A (eIF4A-yeGFP). A tetrameric mutant of cherry florescent protein fused to a well-established Crm1 nuclear export signal (tcherry-NES) was used as a nuclear exclusion marker. White arrows placed on the merged images are pointing to the yeast nuclei. The deletion of Lph2 led to the accumulation of eIF4A inside the nucleus in ΔLph2 cells. Perfect nuclear exclusion of eIF4A was observed in wild type cells.

48

3.3.3 Lph2 is the importin for the ribosomal protein L12 A (rpL12A)

We previously showed that ribosomal protein rpL12A binds to Lph2 in a Gsp1Q71L-GTP-sensitive manner, suggesting that rpL12A might be an import cargo for yeast NTR Lph2. (Figure 3.15). However, we still needed to consider whether Lph2 interacted with rpL12A directly or through a yet unidentified protein factor contained in the yeast lysate. Thus, we decided to characterize the formation of the Lph2•rpL12A import complex with purified recombinant components.

ED-SUMOvera-H_{12}-Lph2 fusion protein was incubated with rpL12A in the presence and absence of Gsp1Q71L-GTP (Figure 3.17, A. Input). ED-tagged Lph2 was immobilized onto anti-Z domain affibody matrix and subsequently eluted by on-column protein cleavage using SUMOvera protease. After protein elution, interaction between rpL12A and Lph2 was visible regardless of the presence of Gsp1Q71L-GTP (Figure 3.17, A. Bound fraction). Notably, the binding of rpL12A to Lph2 seemed to be more efficient when Gsp1Q71L-GTP was not present, suggesting that Gsp1Q71L-GTP and rpL12A might compete for a binding site in Lph2. In other words, Gsp1Q71L-GTP might limit the formation of a stable Lph2•rpL12A complex by blocking rpL12A binding site in Lph2. To clarify this issue, we used rpL12A as immobilized bait for binding assays instead of Lph2. We first incubated H_{14}-ZZ-bdNEDD8-rpL12A with untagged Lph2 in the presence and absence of untagged Gsp1Q71L-GTP (Figure 3.17, B. Input). After immobilization of rpL12A to anti Z-domain beads and further protein elution with bdNEDP1, we observed only the formation of the Lph2•rpL12A import complex in both samples (Figure 3.17, B. Bound fraction). Gsp1Q71L-GTP did not form a stable complex together with Lph2 and rpL12A. In fact, Lph2 bound more efficiently to rpL12A in the sample lacking Gsp1Q71L-GTP. It thus appeared that Gsp1Q71L-GTP and rpL12A competed for binding rpL12A.

Altogether, these binding experiments confirmed that rpL12A might be an import cargo for Lph2 in *S. cerevisiae* as it binds to Lph2 in a Gsp1-sensitive manner. However, more experiments to be performed to explain how rpL12A is completely dissociate from Lph2 when entering the nucleus where the concentration of Gsp1Q71L-GTP is high. Another experiment would be to test whether Lph2 is responsible for the nuclear import *in vivo*. We therefore analyzed the distribution of rpL12A-yeGFP in wild type and Lph2-knockout cells (ΔLph2) using the same experimental setup as in section 3.3.2 . We observed strict cytoplasmic localization of rpL12A-yeGFP in wild type and ΔLph2 cells (Figure 3.17. C).

rpL12A seemed to be perfectly excluded from the nucleus as it perfectly co-localized with the nuclear exclusion marker tCherry-NES.
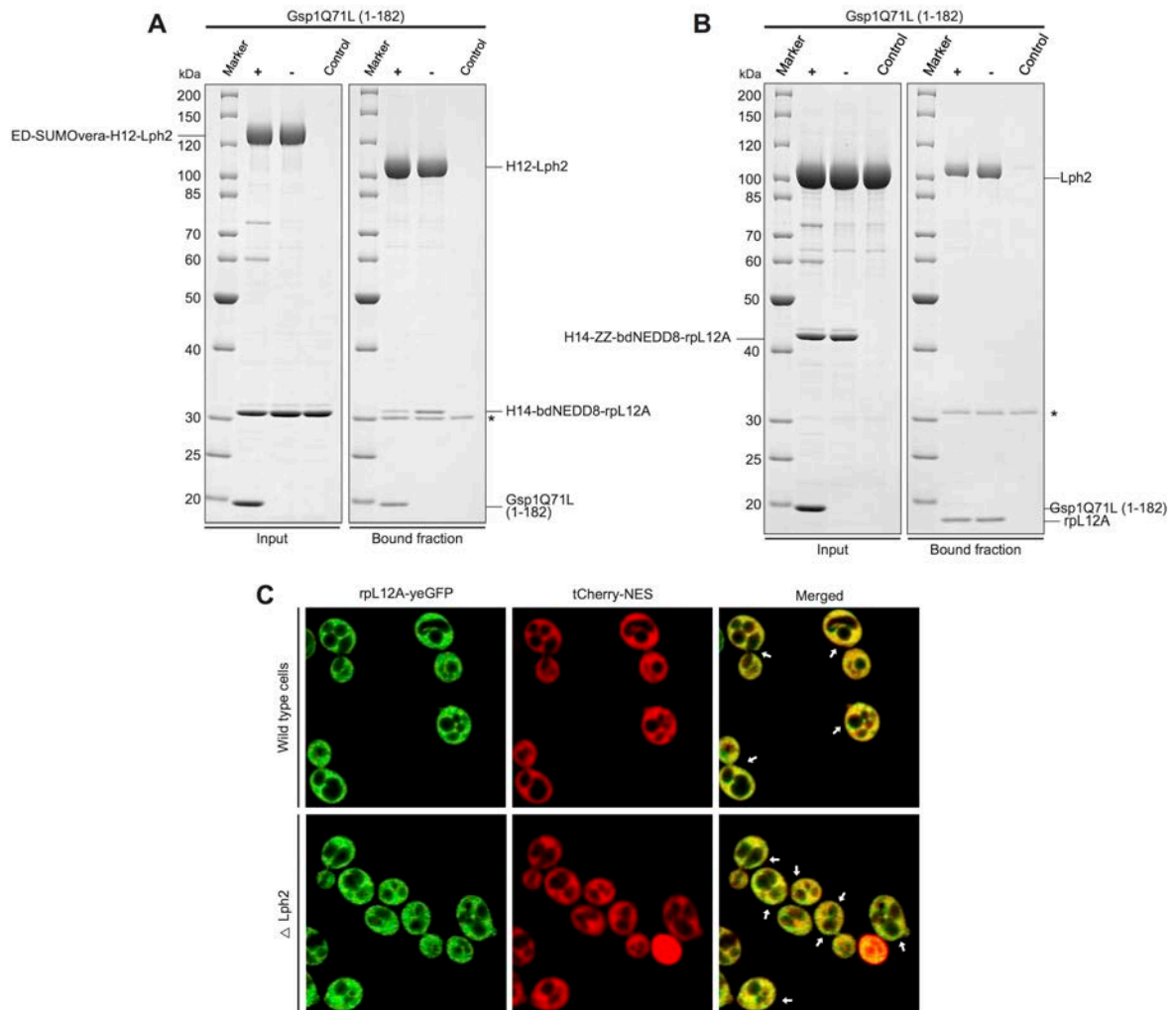
Figure 3.17 Yeast rpL12A is an import cargo for Lph2. (A) Purified components were used as starting material for the characterization of the Lph2•rpL12A import complex. The binding of ED-tagged Lph2 (1.5 μM) to rpL12A (5 μM) was performed in the presence and absence of Gsp1Q71L-GTP (5 μM) (input fraction). Samples were immobilized using an anti-Z domain matrix and further eluted with SUMOvera protease (bound fraction). Lph2 bound to rpL12A regardless the presence of Gsp1Q71L-GTP. SUMOvera protease used for protein elution is marked by (*). Input and bond fraction samples were analyzed by SDS-PAGE. **(B)** Binding experiments were performed as in (A), except that rpL12A was immobilized instead of Lph2. rpL12A recognized Lph2 as an importin without the aid of Gsp1Q71L-GTP. bdNEDP1 used for on-column protein elution is marked by (*). Control samples were used to test unspecific prey binding to the anti-Z domain matrix. **(C)** Cellular localization of rpL12A-yeGFP was analyzed by confocal fluorescent scanning microcopy. rpL12A was C-terminally tagged with yeGFP in wild type and Lph2-knockout (ΔLph2) yeast cells. Tetrameric cherry fluorescent protein fused to a nuclear export signal (tCherry-NES) was used as nuclear exclusion marker. White arrows show the localization of the yeast nucleus in the merged images. Note that there was not clear evidence of nuclear import of rpL12A neither in wild type cells nor in ΔLph2.

After performing binding assays, rpL12A seemed to be an import cargo for Lph2 as it was recognized in a Gsp1Q71L-GTP-sensitive manner. Therefore, it would be expected to observe nuclear accumulation of rpL12A in wild type cells and a cytoplasmic accumulation in the ΔLph2 cells. Now, the question as why rpL12A was strictly

cytoplasmic in wild type yeast cells remains to be answered. Strict cytoplasmic accumulation of rpL12A is well explained by the fact that rpL12A is rapidly exported by Crm1 after being assembled into 60S pre-ribosomal subunits inside the nucleus (Ho et al. 2000; Gadal et al. 2001; Lo et al. 2009). Therefore, in steady state, rpL12A would spend only a small fraction of time in the nucleus after import. More experiments should be performed to confirm that rpL12A is the fourth import cargo identified for Lph2 in *S. cerevisiae* together with Rpf1, Ho endonuclease and protein Swi6 (Kim et al. 2010; Caesar et al. 2006; Bakhrat et al. 2006).

3.3.4 eIF4A is exported exclusively by Lph2 whereas rpL12A interacts with different NTRs

Redundancy in cargo recognition by different NTRs is frequently observed in *S. cerevisiae* and mammals (Greiner et al. 2004; Mühlhäusser et al. 2001; Jäkel et al. 2002; Grosshans et al. 2001; Rout et al. 1997; Caesar et al. 2006). So far, we have only demonstrated that Lph2 recognizes specifically eIF4A and rpL12A (Figure 3.15). Until now, it was still completely unknown to us whether other NTRs in yeast could recognize eIF4A and rpL12A as transport cargos and form a stable protein complex. In order to test this, we first over-expressed in *E. coli* different NTRs from *S. cerevisiae*. Then, cleared bacterial lysates containing the different NTRs were incubated with either ED-SUMOvera-eIF4A or ED-SUMOvera-rpL12A to test the formation of export and import complexes, respectively.

For eIF4A, we decided to use the bacterial lysates containing recombinant Crm1, the major exportin in yeast (Stade et al. 1997; Adachi & Yanagida 1989; Kirli et al. 2015), and 7 less characterized NTRs that could potentially act as exportins (Figure 3.18. Input). Gsp1Q71L-GTP was included in the analysis as we were looking for the formation of trimeric export complexes. Samples were incubated with anti-Z domain affibody matrix in order to immobilize the ED-SUMOvera-tagged eIF4A. After elution of the samples with SUMOvera protease, we confirmed that only Lph2 recognized eIF4A as an export cargo (Figure 3.18. Bound fraction). We observed only the formation of a stable Lph2•eIF4A•Gsp1Q71L-GTPexport complex whereas all other tested NTRs did not bound to eIF4A using the same conditions as for Lph2. eIF4a has been previously reported to interact to Crm1 and Yrb4 (Gavin et al. 2006). However, we did not detect any

interaction between these two NTRs and eIF4A. This extremely specific interaction confirms that Lph2 is the only NTR responsible for the export of eIF4A in yeast cells.
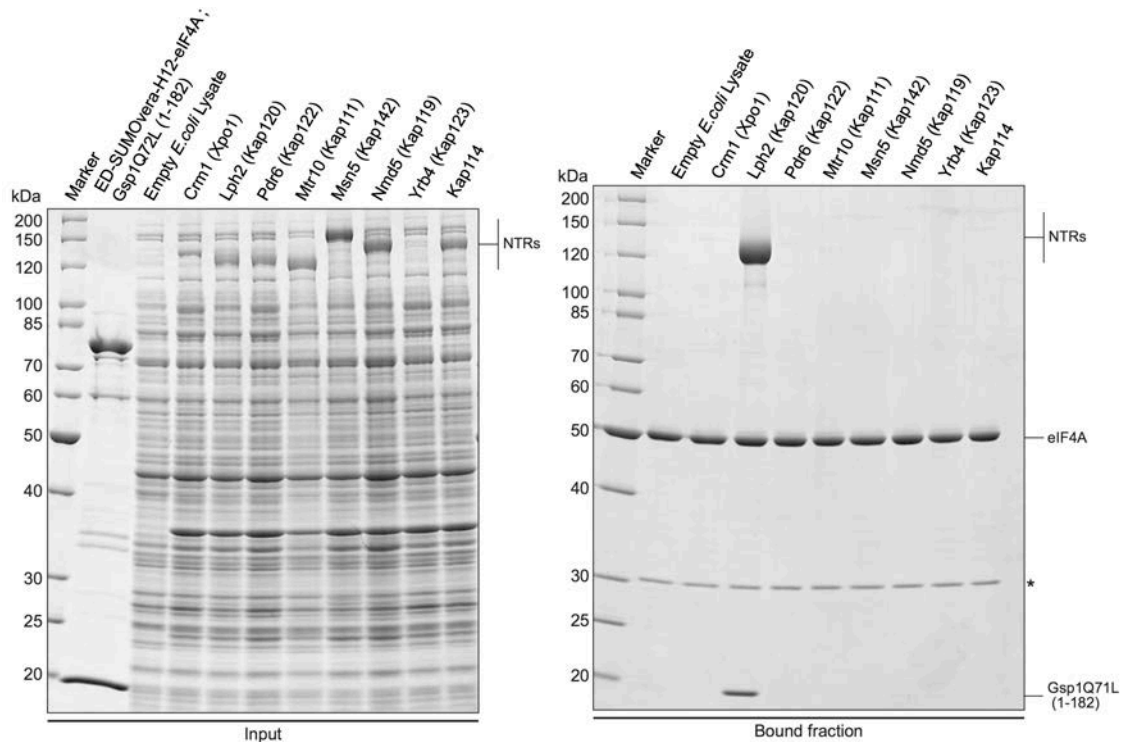
Figure 3.18 Lph2 is the only yeast NTR that recognizes eIF4A. ED-SUMOvera-H$_{12}$-eIF4A fusion protein (2 μM) was used as bait to analyze the binding specificity of different nuclear transport receptors (Crm1, Lph2, Pdr6, Mtr10, Msn5, Ndm5, Yrb4 and Kap114) to eIF4A. To form export complexes *in vitro*, the experiments were performed in the presence of Gsp1Q71L-GTP (6 μM). *E. coli* lysates containing a specific over-expressed yeast nuclear transport receptors were used as prey sources (input). ED-tagged eIF4A was immobilized anti-Z domain affibody matrix and incubated together with Gsp1Q71L-GTP and the *E. coli* lysates. A sample incubated with an "empty" *E. coli* lysate lacking a NTR was used as negative control. Elution of the immobilized eIF4A was done through SUMOvera protease elution (bound fraction). Lph2 was the only NTR able to form a stable trimeric export complex together with eIF4A and Gsp1Q71L-GTP. Samples corresponding to 2 μg of eIF4A were loaded for each sample to be resolved by SDS-PAGE and stained by coomassie blue. The (*) stands for SUMOvera protease used during protease elution. Molecular weights of the protein from the marker are labeled.

Next, we tested the binding specificity of rpL12A to different yeast NTRs. To this end, we incubated ED-SUMOvera-rpL12A with 10 different *E. coli* lysates containing one of the importins known in yeast (Figure 3.19. Input). ED-SUMOvera-rpL12A was then immobilized onto anti-Z domain affibody matrix, and subsequently subjected to on-column protein cleavage with SUMOvera protease. After resolving the samples by SDS-PAGE, it was clear that Lph2 was the most efficiently bound NTR to rpL12A (Figure 3.19. Bound fraction). In addition, other NTRs such as Mtr10, Msn5, Nmd5 and Kap14 also bound to rpL12A to different extents. These data indicate that Lph2 is the major importin for rpL12A. However, import of rpL12A might be also carried out by different parallel import mechanisms driven by different NTRs. In addition, these data also supports previous experiments showing that ribosomal proteins are one of the most common examples of

protein cargos with a broad redundancy in NTR recognition (Rout et al. 1997; Jäkel et al. 2002).
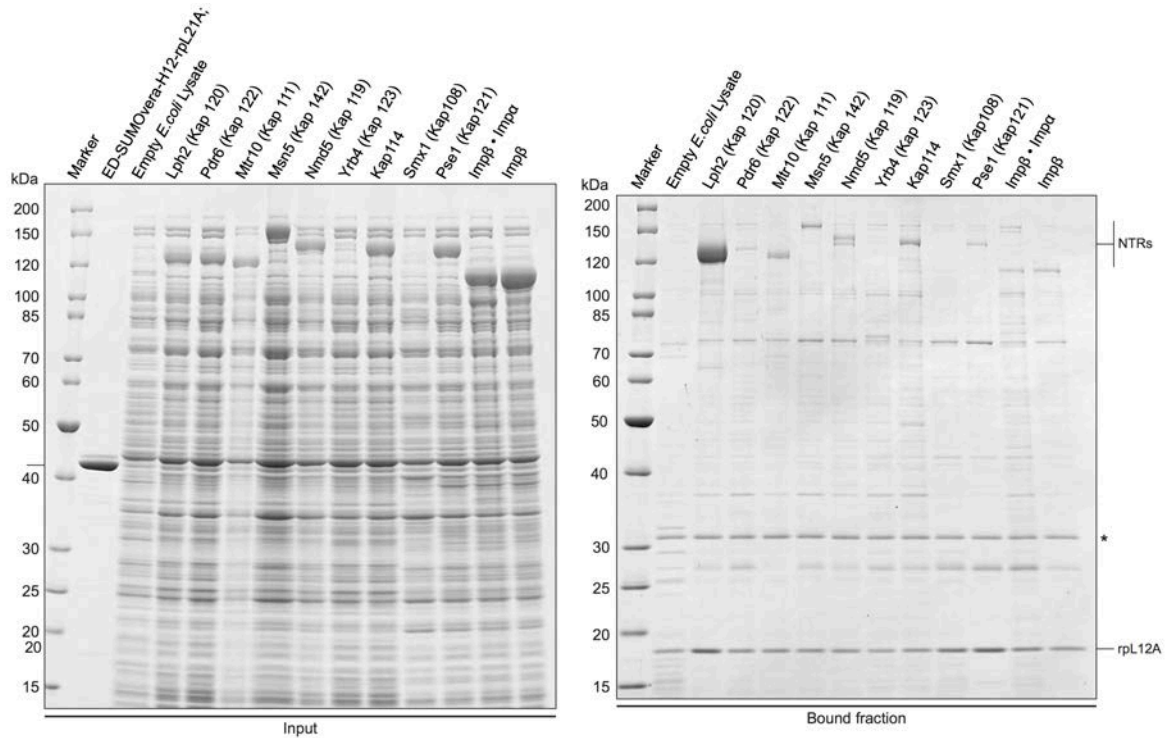
Figure 3.19 rpL12A interacts with multiple yeast nuclear transport receptors. Analysis of the interaction between rpL12A and different yeast NTRs (Lph2, Pdr6, Mtr10, Msn5, Ndm5, Yrb4, Kap114, Smx1, Pse1, Impβ•Impα dimer and Impβ) was performed (input). rpL12A was incubated with the bacterial lysates containing the recombinant importins and further immobilized onto anti Z-domain affibody matrix. Several yeast NTR bound to rpL12A (bound fraction). Lph2 was the most enriched NTR in the bound fraction as Lph2•rpL12A import complex seemed to be most prominent. On the other, several other NTRs such as Mtr10, Pdr6, Msn5 and Kap114 bound to some extent to rpL12A. A negative sample control was included in the analysis to account for binding of *E. coli* proteins to rpL12A (empty *E. coli* lysate). Samples corresponding to approximately 1.5 μg of rpL12A were analyzed by SDS-PAGE. bdNEDP1 is marked in the gel by (*).

3.4   Identification and characterization of novel Pdr6 transport cargos

3.4.1   Pdr6 is a bidirectional NTR in *S. cerevisiae*

There are 14 identified NTRs in *S. cerevisiae*. So far, only Msn5 was shown to act as a bidirectional NTR in yeast. Therefore, it was still completely unclear to us whether another yeast NTR would also behave as a bidirectional transporter (Kaffman, Rank, O'Neill, et al. 1998; Yoshida & Blobel 2001).In this study, we described Lph2 as a second example of a bidirectional NTR in *S. cerevisiae* (Figure 3.15). But whether another yeast NTR could also behave as bidirectional NTR was still an open question.

Pleiotropic drug resistance regulatory protein 6 (Pdr6/Kap122) is a yeast NTR that has been shown to import the transcription factor IIA (TFIIA) and the small subunit of the ribonucleotide reductase (R2) (Titov & Blobel 1999; Zhang et al. 2006). To the best of our

knowledge, there has not been any conclusive study showing whether Pdr6 would also act as an exportin in *S. cerevisiae*. In addition, Pdr6 has a moderate affinity to RanGTP (231 nM) (Hahn & Schlenstedt 2011) that appears to be an inherent feature of bidirectional NTRs (Lipowsky et al. 2000; Grünwald & Bono 2011). We therefore aimed to investigate in detail whether Pdr6 could recognize specific yeast proteins as export and import cargos
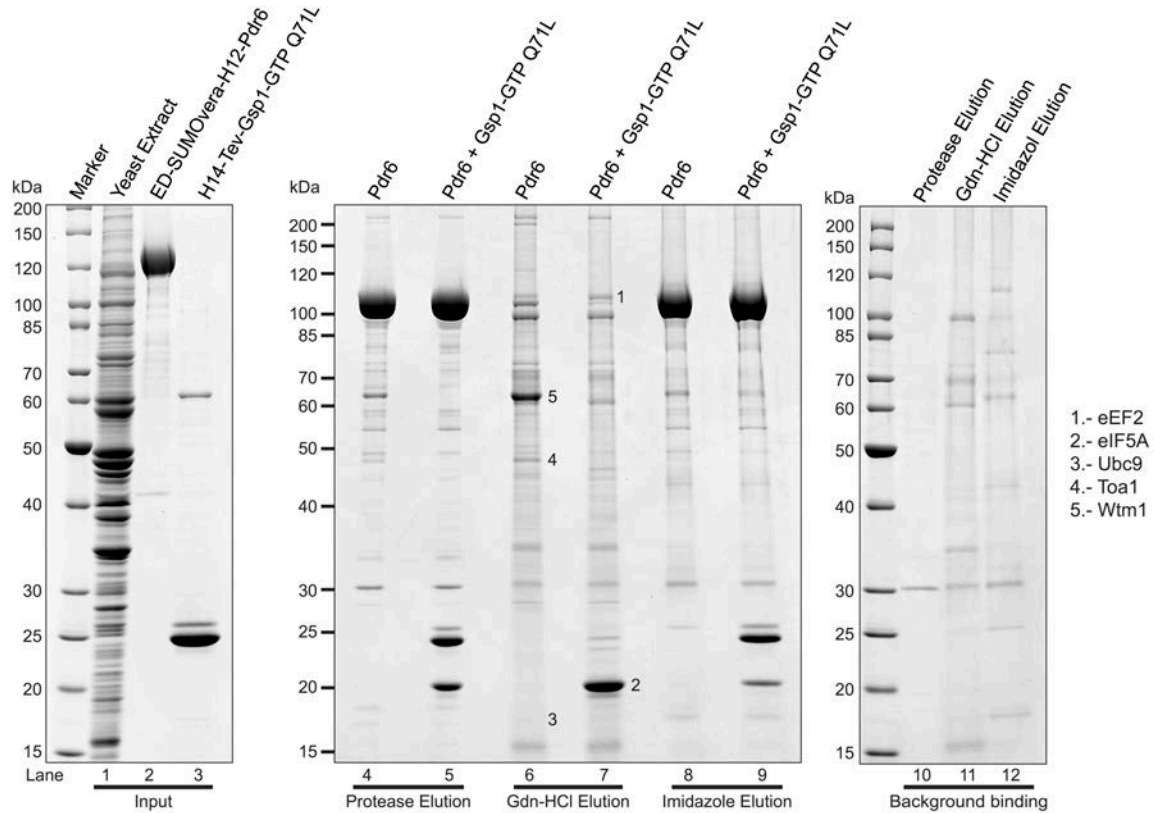
Figure 3.20 Identification of import and export cargos for Pdr6. Novel putative transport cargos for Pdr6 were identified using our designed TAP system. Pdr6 was used as bait protein (lane 2) and yeast extract (lane 1) as source of transport cargos for the NTR. ED-SUMOvera-H_{12}-Pdr6 protein (1.5 μM) was immobilized onto anti Z-domain affibody matrix and further incubated with the yeast lysate. The experiment was performed in the presence and absence of Gsp1Q71L-GTP (5 μM) (lane 3). Samples were then eluted by on-column protein cleavage with SUMOvera protease (lane 4 and 5). Next, eluted protein samples were bound to Ni^{2+} chelate matrix to perform the second affinity chromatography step. The specific elution of Pdr6 protein binders from the column was carried out by addition of guanidinium hydrochloride containing buffer (lanes 6 and 7). Distinct protein bands between the sample with and without Gsp1Q71L-GTP were excised from the polyacrylamide gel to be identified by mass spectrometry analysis. Last, His-tagged Pdr6, Gsp1Q71L-GTP and residual protein contaminants were eluted by 400 mM imidazole containing buffer (lane 8 and 9). A control sample lacking Pdr6 was included to analyze unspecific protein binding during the whole TAP procedure (lanes 10, 11 and 12). Samples corresponding to approximately 1.5μg of Pdr6 or to 35mOD of the initial cell extract were analyzed by SDS-PAGE followed by colloidal coomassie staining. These are the proteins identified by mass spectrometry analysis; Export cargos: eEF2 (1) and eIF5A (2); Import cargos: Ubc9 (3), Wtm1 (4) and Toa1 (5).

We first expressed and purified ED-SUMOvera-H_{12}-Pdr6 fusion protein in order to perform a TAP experiment as described in Figure 3.14. Pdr6 was incubated with yeast lysate to search for potential transport cargos in the presence and absence of Gsp1Q71L-GTP (Figure 3.20. Lanes 1, 2 and 3). Samples were immobilized onto anti-Z domain affibody matrix and further eluted with using SUMOvera protease (Figure 3.20. Lanes 4

and 5). We then separated specific protein binders from the His-tagged Pdr6 and Gsp1Q71L-GTP by Gdn-HCl elution from the Ni^{2+} chelate matrix (Figure 3.20. Lanes 6 and 7). Eluted samples were resolved by SDS-PAGE and the most distinct protein bands for each sample were excised from the polyacrylamide gel for protein identification by MS analysis. The elongation factors eIF5A and eEF2 were identified as the very first two export cargos for Pdr6. Nuclear export of eEF2 and eIF5A by Pdr6 is well possible in *S. cerevisiae*, as both elongation factors are known to be involved in cytoplasmic-restricted cellular functions (Perentesis et al. 1992; Gregio et al. 2009; Hyun Ah Kang & Hershey 1994). We also identified the SUMO-conjugating enzyme Ubc9 as a novel import cargo for Pdr6. Ubc9 is involved in nuclear architecture maintenance and chromosome segregation in different organisms (Nacerddine et al. 2005; Dieckhoff et al. 2004). Therefore, import of Ubc9 by Pdr6 is a well reasonable process in *S. cerevisiae*. Protein identification by MS analysis also confirmed previous reports for Toa1 and Wtm1 as import cargos for Pdr6 (Titov & Blobel 1999; Zhang et al. 2006).

The identification of potential import as well as export cargos for Pdr6 clearly shows that Pdr6 might be another example of bidirectional NTR in budding yeast. In other to assure the fact that Pdr6 is indeed a bidirectional NTR, we decided to characterize further the interaction between Pdr6 and three of the identified cargos: eIF5A, eEF2 and Ubc9.

3.4.2   Pdr6 is the export receptor for eIF5A

The hypusine-containing elongation factor eIF5A was identified as one of the major protein binders for Pdr6 (Figure 3.21). The interaction between Pdr6 and eIF5A seemed to happen strictly in the presence of Gsp1Q71L-GTP. This observation suggested that Pdr6 is the potential exportin for eIF5A in *S. cerevisiae*. However, we wanted to reconstitute the Pdr6•eIF5A•Gsp1Q71L-GTP export complex in a minimal setup by using recombinant purified components to test whether an unidentified protein contained in yeast lysate mediated the recognition of eIF5A by Pdr6.

We immobilized ED-SUMOvera tagged Pdr6 onto anti-Z domain affibody matrix in order to be used as bait. Immobilized Pdr6 was incubated with an excess of eIF5A in the presence and absence of Gsp1Q71L-GTP (Figure 3.21, A. Input). To test unspecific protein binding to the matrix, a sample without the immobilized Pdr6 was used. SUMOvera protease was used to elute bound Pdr6 and samples were further analyzed by

SDS-PAGE. Exportins form a complex of high affinity with their respective cargos only in the presence of Ran-GTP (Gsp1, in yeast) (Kutay et al. 1997; Fornerod et al. 1997). Likewise, Pdr6 only recognized eIF5A in the presence of Gsp1Q71L-GTP (Figure 3.21, A. Bound fraction). We therefore assumed that Pdr6 is the corresponding exportin for this elongation factor. In addition, we showed that only Gsp1Q71L-GTP was sufficient for Pdr6 to efficiently recognize eIF5A as its export cargo.
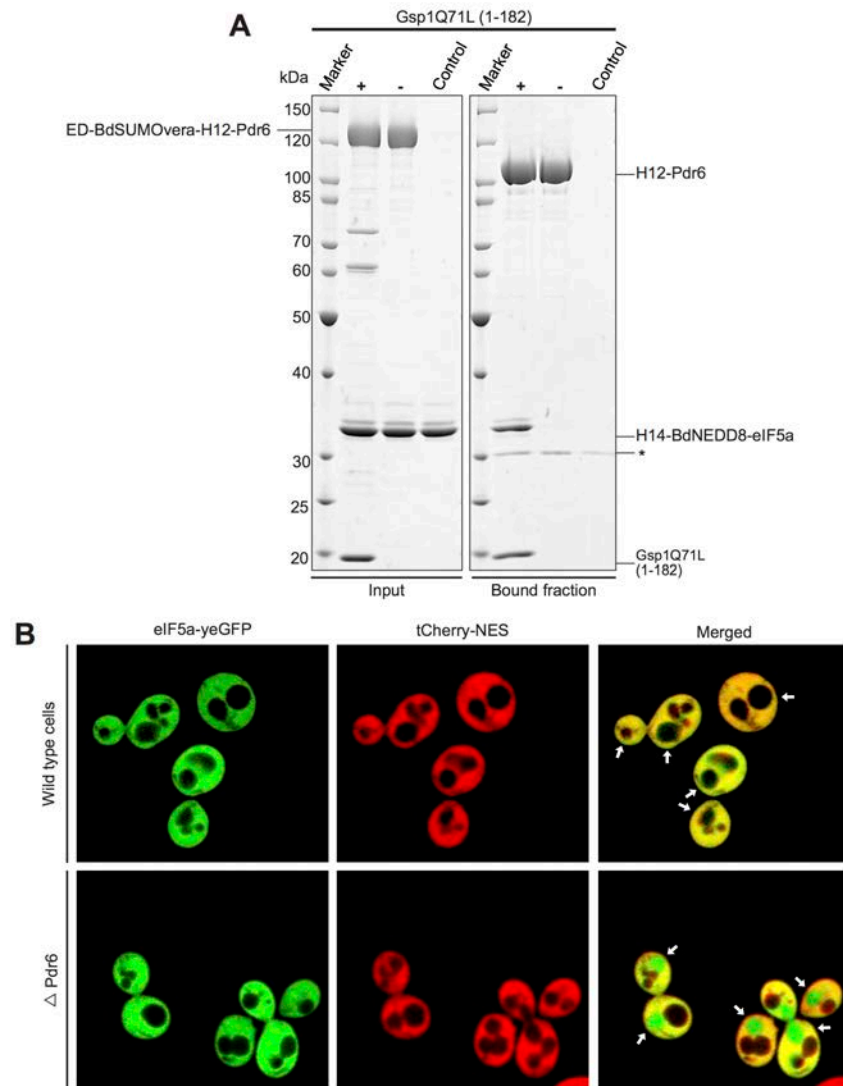
Figure 3.21 Pdr6 as a novel exportin for eIF5A. (A) Purified ED-SUMOvera-H$_{12}$-Pdr6 fusion protein (1.5 μM) was used to test the bind of Pdr6 to eIF5A. Tagged Pdr6 was incubated with eIF5A (5 μM) in the presence and absence of Gsp1Q71L-GTP (5 μM) (input). Then, ED-tagged Pdr6 was immobilized onto an anti-Z domain affibody matrix and later eluted with SUMOvera protease (bound fraction). eIF5A bound to Pdr6 in a Gsp1Q71L-GTP-dependent manner. Input and bound fractions were analyzed by SDS-PAGE followed by coomassie staining. SUMOvera protease is indicated in the polyacrylamide gel by (*). We included a negative control sample without the addition of Pdr6 to test the unspecific binding of eIF5A to the matrix. **(B)** Localization of fluorescently labeled eIF5A in wild type and Pdr6-knockout (ΔPdr6) yeast cells was analyzed by confocal fluorescent scanning microscopy. A GFP tag was introduced at the C-terminus of eIF5A by homologues recombination (eIF5A-yeGFP). In order to localize the yeast nucleus, cells were transformed with a plasmid encoding a fusion protein between tetrameric cherry protein and a Crm1 nuclear export signal (tcherry-NES). The small white arrows are placed to ease the localization of the nucleus in the merged images. Note that deletion of Pdr6 gene caused a massive nuclear accumulation of eIF5A-yeGFP. In contrast, eIF5A remained perfectly cytoplasmic in living wild type cells.

Globular proteins of up to 20-40 kDa in size can freely diffuse through the nuclear pore complex. Consequently, eIF5A can be expected to freely diffuse into the nucleus given its size of 17 kDa. However, a previous study showed that eIF5A is strictly localize in the cytoplasm of *S. cerevisiae* (Valentini et al. 2002). We therefore envisioned that Pdr6 actively exports eIF5A out of the nucleus in yeast cells. To prove this, we analyzed the subcellular localization of eIF5A-yeGFP in wild type and Pdr6-knockout (ΔPdr6) cells by confocal immunofluorescence scanning microscopy. We fused tetrameric Cherry fluorescent protein to a nuclear export signal to be used as a nuclear exclusion marker. When Pdr6 was deleted in yeast cells, eIF5A-yeGFP massively accumulated in the nucleus (Figure 3.21. B). In contrast, the accumulation of eIF5A-yeGFP did not occur in wild type cells (Figure 3.21.A). Single deletion of Pdr6 certainly proves that Pdr6 is needed the true nuclear export factor for eIF5A *in vivo*. Consequently, the export of eIF5A and the import of previously reported cargos (Titov & Blobel 1999; Zhang et al. 2006) by Pdr6 provide enough evidence to state that Pdr6 is the third so far know example of a bidirectional NTR in *S cerevisiae*.

3.4.3   Ubc9 is imported by Pdr6 in *Saccharomyces cerevisiae*

The nuclear transport receptor Pdr6 is a well-described import factor in *S. cerevisiae* (Titov & Blobel 1999; Zhang et al. 2006). Here, we identified the SUMO-conjugation enzyme Ubc9 as a new putative Pdr6 import cargo in a TAP experiment (Figure 3.20). Pdr6•Ubc9 complex might have been assisted by an unknown protein factor present in the cellular lysate during the identification. In order to test this possibility, we decided to characterize the complex using recombinant purified components.

ED-tagged Ubc9 was incubated together with Pdr6 in presence and absence of untagged Gsp1Q71L-GTP (Figure 3.22,A. Input). Pdr6 bound similar amounts of Ubc6 in the presence and absence of Gsp1Q71L-GTP, suggesting that Gsp1Q71L-GTP did not hinder drastically the recognition of Ubc9 by Pdr6 (Figure 3.22, A. bound fraction). This experiment led to the question of whether Ubc9 and Gsp1Q71L-GTPbound to Pdr6 at the same time or whether they competed for the binding of the NTR. In order to discern between these two possibilities, we immobilized Ubc9 to be used as the bait. ED-SUMOvera-Ubc9 was incubated with untagged Pdr6 in the presence and absence of untagged Gsp1Q71L-GTP (Figure 3.22, B. Input). The formation of Pdr6•Ubc9 complex

was significantly more efficient in the sample lacking Gsp1Q71L-GTP (Figure 3.22, B. Bound fraction). Therefore, it seemed that Gsp1Q71L-GTP competed with Ubc9 in the recognition of Pdr6. In addition, we did not observe the formation of a trimetric complex in the sample containing Gsp1Q71L-GTP. These experiments suggest that Ubc9 might be act as a potential functional import cargo for Pdr6. In order to confirm this, we decided to determine the subcellular localization of Ubc9 in wild type and Pdr6-knockout (ΔPdr6) cells.
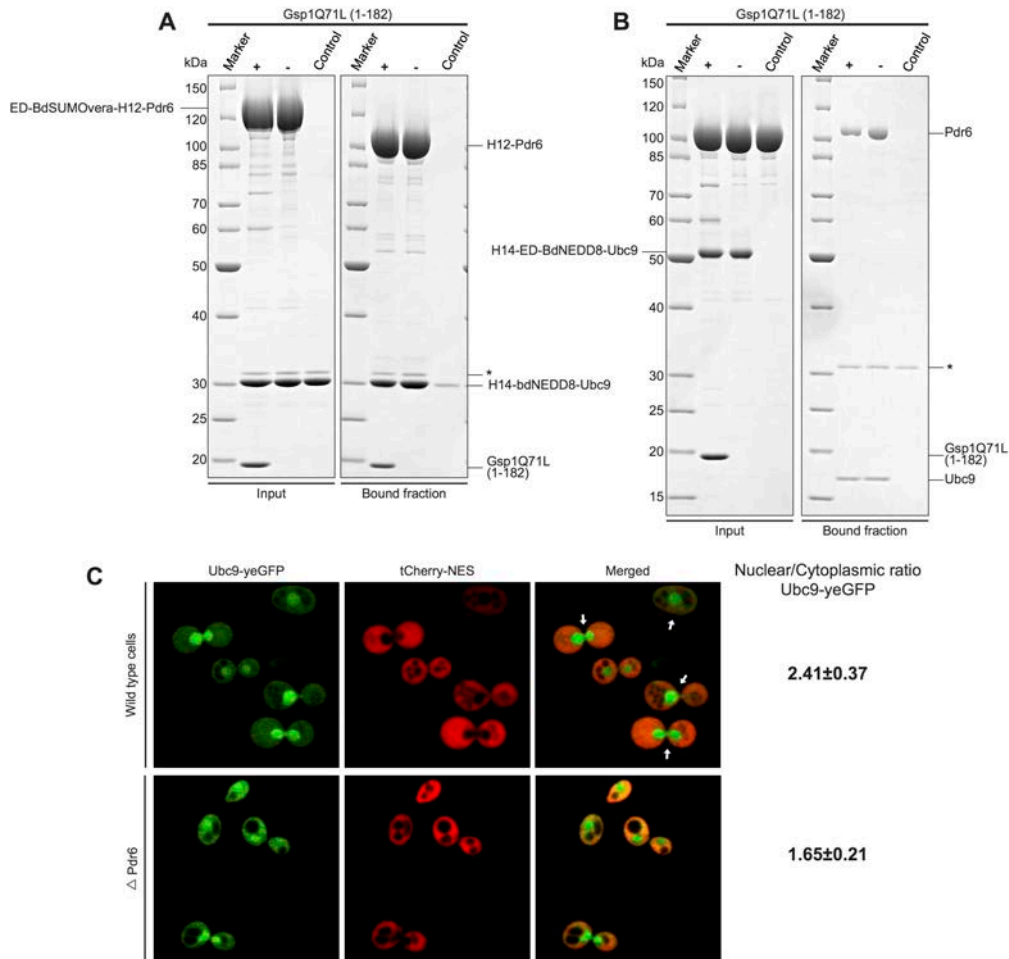
Figure 3.22 Ubc9 is recruited as an import cargo by Pdr6. (A) Interaction of ED-SUMOvera-tagged Pdr6 (1.5 μM) and yeast Ubc9 (5 μM) was analyzed in the presence and absence of Gsp1Q71L-GTP (5 μM) (input). The recombinant proteins were incubated in the corresponding combinations prior to the immobilization of Pdr6 onto anti-Z domain affibody matrix. Samples were eluted from the column by protease cleavage with SUMOvera (bound fraction). Unspecific binding of Ubc9 to the matrix was analyzed in a control sample included in the analysis. Pdr6 efficiently recognized Ubc9 and Gsp1Q71L-GTP simultaneously. SUMOvera protease is indicated in the gel by (*). **(B)** A binding experiment was performed as in (A), except that ED-tagged Ubc9 was now the immobilized bait to form the protein complex. Negative control sample lacking the bait was included in the experiment to analyze the unspecific prey binding to the used matrix. bdNEDP1 protease is indicated in the gel by (*). Protein samples for input and bound fractions were analyzed by SDS-PAGE. **(C)** The localization of Ubc9-yeGFP fusion protein was analyzed in wild type yeast and Pdr6-knockout cells (ΔPdr6). yeGFP-tagged Ubc9 was genomically expressed in both strains. In order to localize the nucleus in yeast, tetrameric cherry fused to a Crm1-type nuclear localization signal (tcherry-NES) was used as a nuclear exclusion marker. The small white arrows are placed to ease the localization of the nucleus in the merged images. The nuclear/cytoplasmic ratio of Ubc9-yeGFP of wild type and Pdr6-knockout cells was calculated using Adobe Illustrator CS5.

Data from previous studies localized Ubc9 predominantly in the nucleus of *S. cerevisiae* (Seufert et al. 1995; Dieckhoff et al. 2004). Here, we wanted to investigate whether Pdr6 is a key modulator in the nuclear export process of Ubc9 in the budding yeast. To do so, we determined the subcellular localization of yeGFP tagged Ubc9 using confocal fluorescent scanning microcopy microscopy in wild type and ΔPdr6 mutant cells. Both type of cells were also transformed with a plasmid encoding for cherry fused to a Crm1-type nuclear export signal (tCherry-NES) to be used as nuclear exclusion marker. We observed a predominant nuclear localization of Ubc9-yeGFP in wild type and ΔPdr6 mutant cells (Figure 3.22. C). However, the nuclear/cytoplasmic signal ratio of Ubc9-yeGFP in wild type cells (2.41±0.37) is significantly higher than the one in ΔPdr6 mutant cells (1.65±0.21), which suggests that the cytoplasmic levels of Ubc9-yeGFP are lower in yeast wild type cells (Figure 3.22. C). Therefore, our results demonstrate that lack of Pdr6 in *S. cerevisiae* hampers the efficient active import of Ubc9 to the nucleus.

3.4.4 Elongation factor 2 (eEF2) is transported from the nucleus by Pdr6

Exportins bind their substrates strictly in the presence of Ran-GTP (Fornerod et al. 1997; Görlich & Kutay 1999). We identified elongation factor 2 (eEF2) as a protein binder for Pdr6 only in the presence of Ran-GTP (Figure 3.20). Therefore, eEF2 seemed to be a perfect export cargo candidate for Pdr6. In order to confirm this assumption, we decided to analyze the *in vivo* localization of eEF2 by confocal fluorescent scanning microscopy. eEF2 was genomically tagged at the C-terminus with yeGFP in wild type and Pdr6-knockout (ΔPdr6) yeast cells. As used in previous experiments, tCherry-NES fusion protein was used as a nuclear exclusion marker to localize the yeast nucleus. We did not observe accumulation of eEF2-yeGFP in the nuclei of wild type cells, whereas the single deletion of Pdr6 caused the nuclear accumulation of eEF2-yeGFP (Figure 3.23). We thus show that Pdr6 is truly the nuclear export that controls the transport of the high molecular weight eEF2 in yeast cells.
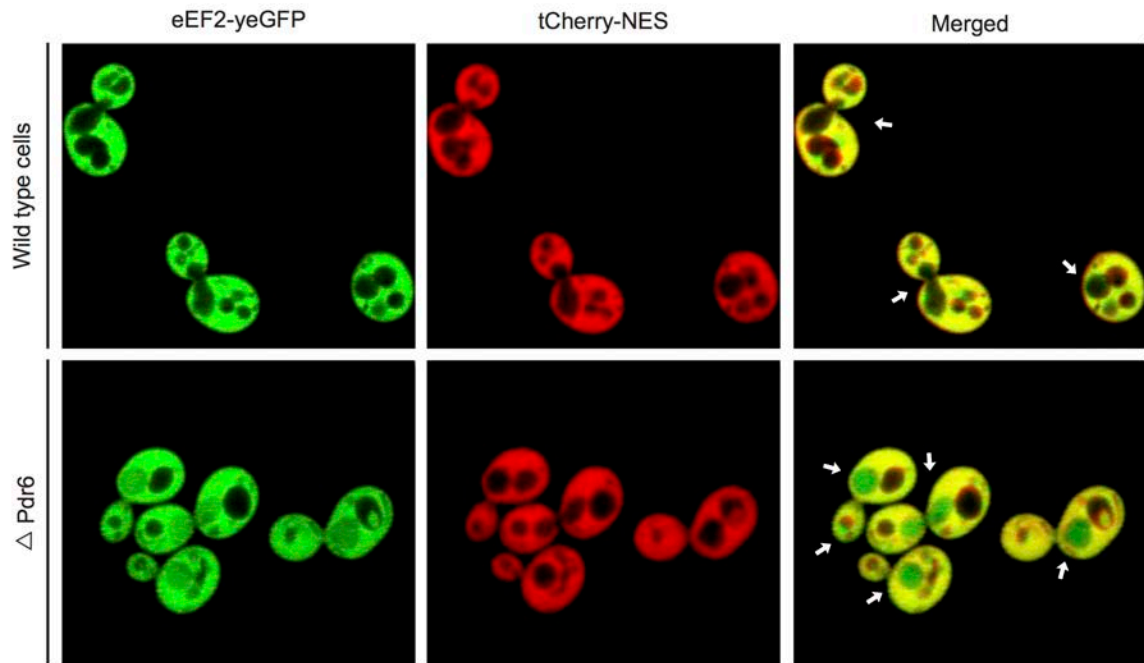
Figure 3.23 Pdr6 exports the eukaryotic elongation factor 2 in *S. cerevisiae*. *In vivo* localization of the eukaryotic elongation factor 2 (eEF2) in wild type and Pdr6-knockout (ΔPdr6) yeast cells. eEF2 was tagged with yeGFP at the C-terminus (eEF2-yeGFP). Cells stability expressing eEF2-yeGFP were transformed with a multicopy plasmid encoding for the nuclear marker composed of cherry protein and a PKI-type nuclear export signal (tcherry-NES). Transformed cells were grown at 30°C in liquid synthetic medium until the exponential growth phase was reached and further imaged by confocal fluorescent scanning microscopy. Arrows in the merged images are pointing to the position of the nucleus inside the cells. Deletion of Pdr6 caused the nuclear accumulation of eEF2-yeGFP.

Pdr6 specifically controls the export of two different translation factors, eEF2 and eIF5A. We therefore assumed that Pdr6 serves the purpose of suppressing any possible translation process in the nucleus by actively exporting essential translation factors. We also speculated that the export of eEF2 and eIF5A by Pdr6 might be an active measure to prevent the early interaction between these factors and immature ribosomal assemblies in the nucleus. In addition, we also proved that Ubc9 is a novel Pdr6 import substrate. Along with the TFIIA and Wtm1, Ubc9 is the third characterized import cargo so far for Pdr6. We therefore conclude that Pdr6 is not only an importin but also a newly described and highly specialized exportin.

3.4.5 Pdr6 is a highly specialized nuclear transport receptor

As mentioned in section 3.3.4 , recognition of a single transport cargo by different NTRs is a recurrent phenomenon in different organisms (Greiner et al. 2004; Mühlhäusser et al. 2001; Jäkel et al. 2002; Grosshans et al. 2001; Rout et al. 1997; Caesar et al. 2006). Therefore, several transport pathways might occur in parallel for the orchestrated transport

of a given substrate in a single cell. We showed that yeast nuclear transport receptor Pdr6 regulates the transport of Ubc9 and eIF5A (sections 3.4.2 and 3.4.3). Now, we investigated to what extent other yeast NTRs might interact with Ubc9 and eIF5A. To do so, we overexpressed different yeast NTRs in *E. coli* and used the bacterial cleared lysates as source of NTRs. We first incubated eIF5A together with Gsp1Q71L-GTP and 8 different bacterial lysates containing a single recombinant yeast NTR (Figure 3.24. Input). We included in the analysis the broad-spectrum exportin Crm1 (Adachi & Yanagida 1989; Stade et al. 1997; Kirli et al. 2015) and 7 less-characterized NTRs that might act also as exportins (Lph2, Pdr6, Mtr10, Msn5, Nmd5, Yrb4 and Kap114). The elongation factor eIF5A only formed a complex with Pdr6 and Gsp1Q71L-GTP, whereas the other 7 tested NTRs did not recognized eIF5A as an export cargo (Figure 3.24. Bound fraction). Pdr6 seems to be the only NTR responsible for the nuclear export of eIF5A in budding yeast. In fact, we didn't find in the literature any annotated interaction between eIF5A and a yeast NTR.
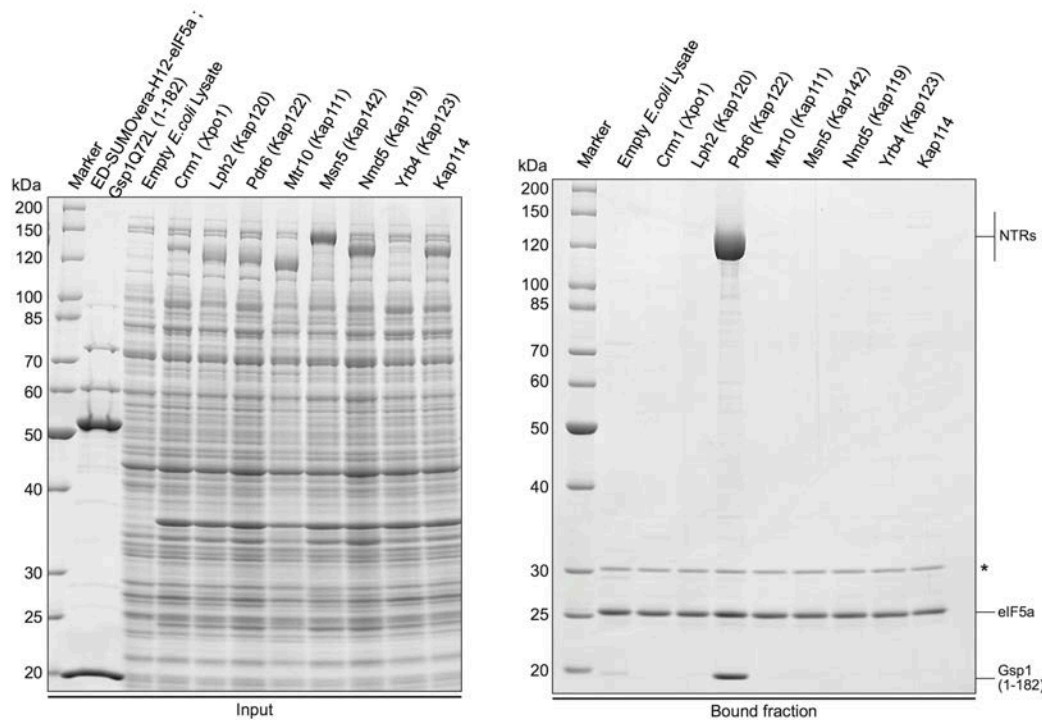
Figure 3.24 Pdr6 is the only nuclear transport receptor that recognizes eIF5A as export cargo. Different *E. coli* lysates containing different recombinant yeast NTRs (Crm1, Lph2, Pdr6, Mtr10, Msn5, Nmd5, Yrb4 and Kap114) were used to test the interaction between the receptors and eIF5A. Gsp1Q71L-GTP was added to the samples in order to form potential export complexes between the NTRs and eIF5A (input). All samples were incubated for 1h at 4°C prior to immobilization of purified ED-SUMOvera-H$_{12}$-eIF5A. SUMOvera protease was used to elute the export complexes from the anti-Z domain affibody matrix (bound fraction). Pdr6 was the only transport receptor able to recognize eIF5A as an export cargo. An empty *E. coli* lysate lacking the yeast NTR was included as negative control. Samples from the input and bound fraction were analyzed by SDS-PAGE and followed by coomassie staining. SUMOvera protease is represented by (*) in the figure.
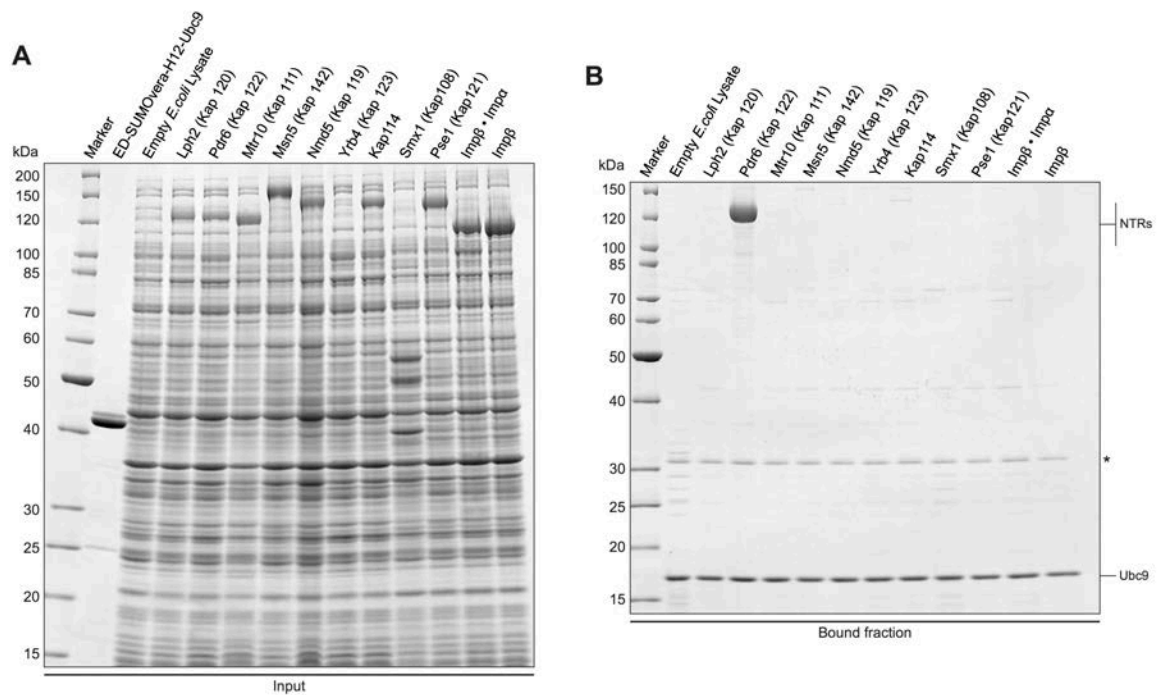
Figure 3.25 Ubc9 is exclusively imported to the yeast nucleus by Pdr6. ED-SUMOvera-H$_{12}$-Pdr6 (2 μM) was incubated with several *E. coli* lysates containing an over-expressed yeast NTR (input). Samples were incubated with anti-Z domain matrix to immobilize Pdr6. Note that Pdr6 was the only NTR able to form an import complex with yeast Ubc9. No interaction was observed between the other tested NTRs and Ubc9. SUMOvera protease used for on-column cleavage during the binding assays is marked with (*).

Next, we analyzed the specificity of Ubc9 recognition as an import cargo by different yeast NTRs. For this analysis, we used 9 different yeast NTRs described as importins (Lph2, Pdr6, Mtr10, Msn5, Nmd5, Yrb4, Kap114, Smx1 and Pse1). We also included as part of the experiment the major importin complex in yeast, Importin β•Importin α complex (Görlich et al. 1997; Görlich & Kutay 1999; Lange et al. 2015). ED-SUMOvera-H$_{12}$-Ubc9 fusion protein was immobilized onto anti-Z domain affibody matrix and further incubated with bacterial lysates containing one of the over-expressed yeast NTR (Figure 3.25. Input). After protein elution, we observed that Pdr6 was the only NTR able to recognize Ubc9 as a cargo, whereas the remaining 9 tested NTRs completely failed to bind Ubc9 (Figure 3.25. Bound fraction). Although Kap114 was reported to interact with Ubc9 to some extent (Rothenbusch et al. 2012), we did not detect any significant binding between them at the tested conditions. We confirmed with this experiment that Ubc9 is actively transported to the nucleus by a single import pathway driven by Pdr6.

3.4.6 Export of eIF5A and import of Ubc9 are two interesting cases about the evolution of NTR functions in yeast and human cells

Multiple bioinformatics studies have shown the loss, divergence and multiplication of NTR genes during the course of evolution (Malik et al. 1997; Görlich et al. 1997; Quan et al. 2008). One interesting example of the diversification of NTR genes during evolution is the absence of ortholog genes of yeast Pdr6 in multicellular organisms. Pdr6 does not share sequence similarity with any of the 20 different vertebrate NTRs that would go beyond that between paralogs. Yet, the cellular functions of Pdr6 are still conserved in humans in two different NTRs. First, human exportin 4 (Exp4) (Lipowsky et al. 2000; Aksu et al. 2016) as well as yeast Pdr6 (Figure 3.21) efficiently export eIF5A from the nucleus to the cytoplasm. Second, it was already proved before that human Importin 13 (Imp13) recognizes Ubc9 (Mingot et al. 2004; Grünwald & Bono 2011) as Pdr6 does in *S. cerevisiae* (Figure 3.22). We therefore hypothesized that the export of eIF5A and the import of Ubc9 were "invented" at least two times during the course of gene evolution in two different organisms. In order to provide some experimental evidence for this observation, we performed binding assays to analyze whether Exp4 and Pdr6 were able to recognize human and yeast eIF5A as an export cargo. Likewise, we analyzed whether Pdr6 and Imp13 bind to human and yeast Ubc9.

First, Pdr6 and Exp4 were incubated with the yeast (sceIF5A) and human (hseIF5A) export cargo in the presence of Gsp1Q71L-GTP. Both non-homologous NTRs recognized sceIF5A and hseIF5A (Figure 3.26. A and B). Since yeast and human eIF5A have a significant sequence similarity (64% sequence identity), we speculated that highly conserved residues are the recognition sites for Pdr6 and Exp4 in both cargos. In addition, we also consider that the mechanism of cargo recognition of Pdr6 and Exp4 must differ drastically as both NTRs are non-homologous proteins. The interaction between Exp4 and hseIF5A has been described in detail (Aksu et al. 2016), however, further experiments have to be performed in order to map the exact recognition motifs of the sceIF5A by Pdr6.

Subsequently, we investigated whether Imp13 and Pdr6 would recognize both yeast and human Ubc9 as import substrates. To this end, we incubated Pdr6 and Imp13 with both Ubc9 homologs in order to let them form a stable import complex. Even though yeast and human Ubc9 have a rather high sequence identity (57%), we observed that Imp13 and Pdr6 were able to recognize only their putative cargo (Figure 3.26. C and D). Residues in human

Ubc9 that make contact with Imp13 are perfectly conserved in *S. cerevisiae* (Grünwald &
Bono 2011), however, yeast Ubc9 was not able to bind Imp13. We therefore speculate that
Pdr6 and Imp13 are not ortholog proteins with conserved cargo interactions, but paralogs
with different evolved Ubc9-binding modes. Several experiments still need to be
performed in order to identify the interaction site between Pdr6 and yeast Ubc9.
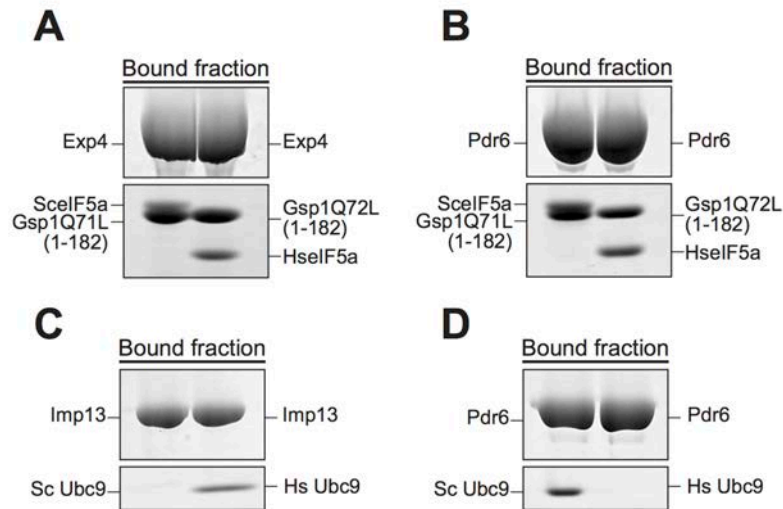
Figure 3.26 Yeast Pdr6 might have evolved into two different non-homologous human NTRs. (A) Analysis of the
interaction between human exportin 4 (Exp4) and homologues proteins of the translation initiation factor 5A from human
(hseIF5A) and yeast (sceIF5A). To form export complexes, ZZ-tagged Exp4 (1 μM) was incubated with hseIF5A (3 μM)
and scIF5A (3 μM) in the presence of Gsp1Q71L-GTP (3 μM). After immobilization of Exp4, protein elution was carried
out by protease on-column cleavage bdSENP1. Exp4 bound equally to hseIF5A and sceIF5A in the presence Gsp1Q71L-
GTP. **(B)** Characterization of the interaction between yeast Pdr6 and both eIF5A export cargos homologues was
performed as described in (A). Note that Pdr6 also formed stable export complexes with hseIF5A and sceIF5A in the
presence of Gsp1Q71L-GTP. **(C)** Recognition of human (hsUbc9) and yeast (scUbc9) Ubc9 as import cargos by human
importin 13 (Imp13). H_{14}-ZZ-SUMOvera-Imp13 fusion protein (1 μM) was incubated with hsUbc9 (3 μM) and scUbc9
(3 μM) for 1h at 4°C. Import cargo recognition by Imp13 was species-specific as it only bound to hsUbc9 in the
condition tested. **(D)** Interaction of Pdr6 with hsUbc9 and scUbc9 was tested as in (C). scUbc9 was only recognized by
Pdr6 as an import cargo. Interaction of hsUbc9 and Pdr6 was not detected. Samples corresponding to 1-2μg of the tested
NTRs were resolved by SDS-PAGE followed by coomassie staining. Only bound fraction samples are shown in the
figure for a better clarity of the experiment.

4 DISCUSSION

4.1 Efficient *in vivo* selection system for proteases with orthogonal specificity.

Here, we have described an *in vivo* selection system based on bacterial survival to obtain SUMO-specific proteases with orthogonal substrate specificity. In our system, one plasmid encodes for a SUMO-specific protease under the control of IPTG, whereas a second plasmid constitutively expresses an antibiotic selection marker (Hph) fused to a different SUMO protein at the N- and the C-termini. The designed system links the half-life of Hph with the cleavage of the different SUMO proteins by a given SUMO-specific protease. To do so, Hph was expressed as a SUMO[1]-degron[NER]-Hph-SUMO[2]-ssrA fusion protein under the control of the constitutive GADPH promoter from *E. coli*. After expression of the SUMO-specific protease, Hph remains intact only if SUMO[2] is cleaved and SUMO[1] is not. In other words, cells survive in selective medium containing HygB only if the SUMO-specific protease has orthogonal specificity for both SUMO proteins. Our results prove that our designed selection system could successfully be used in *E. coli* to isolate SUMO-specific proteases with orthogonal substrate specificity. The main aspects and the advantages of our selection system as compared to other methods used for the engineering of proteases are discussed in the next paragraphs.

The bacterium *E. coli* was selected as a model organism to develop this selection assay for different reasons. First, *E. coli* is cheap to produce in large scale and easy to manipulate by standard laboratory techniques. Second, the high transformation efficiency of *E. coli* cells facilitates the screening of libraries with a complexity of up to 10^{11} members (Dower et al. 1988) . Last, the absence of the SUMO modification pathway in *E. coli* allows engineering of SUMO proteins and SUMO-specific proteases with new specificities without disturbing the bacterial growth.

Several methods have been developed for the evolution of proteases (reviewed in Turanli-Yildiz et al. 2012). Among these, methods based on bacterial survival in antibiotic selective medium allow higher throughput selection as compared to spatial separation based methods where the maximum library size that can be analyzed is limited to 1×10^4 (Packer & Liu 2015). Unlike *in vitro* evolution methods (*i.e.* ribosome display, phage display), *in vivo* selection methods based on cellular growth offer the possibility to select for mutants with an improved or novel enzymatic activity. In addition, our *in vivo* system

ensures that proteases are evolved against non-specific cleavage of *E. coli* proteins. Moreover, our system select for eukaryotic proteases that are efficiently folded by bacterial chaperones at 37°C which would provide a high yield of soluble protease for further protein purification.

In order to base our selection system on bacterial survival, we generated an efficient protease-sensitive antibiotic selection marker by adding two different protein degradation signals that regulate its half-life. Although we employed Hph as a reporter for protein cleavage by the protease of interest, other protein reporters that can be tagged at the N- and C-terminus without losing significant enzymatic activity could be used. At the beginning of this project, we tried to use bleomycin resistance protein (Semon et al. 1987) and tetracycline efflux protein as antibiotic selection markers, however, they did not tolerate SUMO-tagging at the N-terminus (data not shown). A feasible option to be used as a selection marker in our system would be chloramphenicol acetyltransferase (CAT), which confers resistance to chloramphenicol. CAT has been proven to be tolerant against tagging at both ends without losing activity and confers one of the most important features for any *in vivo* selection system which is a wide dynamic range of selection (Rackham & Chin 2005). Therefore, CAT remains an excellent choice for further optimizations of our selection system.

Similar to our work, previously reported selection systems showed that protein cleavage by a specific protease could be linked to a specific cellular phenotype in *E. coli* (*i.e.* fluorescence or bacterial growth) (Sandersjöö et al. 2014; Kostallas et al. 2011; Kostallas & Samuelson 2010). In contrast to our selection system, these studies did not provide any fruitful application of the system and did not offer the possibility to select against protein cleavage. Along these lines, the opportunity to select for and against the cleavage of two different proteins at the same time by a single protease is by far the most important feature of our novel selection system. In fact, the possibility to test for the independent cleavage of two different proteins is what offers the opportunity to isolate proteases with orthogonal specificities.

Although we used our system so far only to evolve novel SUMO proteins and SUMO-specific proteases variants, the design of our novel selection method tolerates the use of other specific proteases and their cognate recognition sites. The major restriction of our system is that the protease of interest needs to leave a specific amino acid (Arg, Lys, Phe, Leu, Trp or Tyr) after the cleavage of the N-terminal recognition site in order to allow

selection against protein cleavage. Proteases such as human rhinovirus 3C protease and thrombin could not be used in our system since protein cleavage results in a residual amino acid that is incompatible with the activation of the N-terminal degradation signal (Petrassi et al. 2005; Cordingleys et al. 1990).

Last, we applied our novel selection system to evolve a SUMO/SUMO-specific protease pair with orthogonal specificity to the SUMOstar/SUMOstar protease pair. However, we believe that our *in vivo* selection system would also represent a valuable tool to improve enzymatic activity of known site-specific proteases, evolve more efficient protein recognition sites, and reveal novel protein sequences that can be recognized by a protease of interest.

4.2 SUMOvera system as an efficient protein fusion tool for eukaryotic hosts

In this study, we aimed for the evolution, isolation and characterization of a SUMO-fusion system that is fully compatible with different assays using eukaryotic extracts. From the beginning of this work, we sought to engineer the eukaryotic-compatible SUMO system based on the SUMO protein and SUMO-specific protease from *B. distachyon* (Frey & Görlich 2014a). Also, we pursued to evolve a SUMO system that has complete orthogonality to the already described SUMOstar system (Butt et al. 2010; Yan et al. 2009). With these ideas in mind, we evolved the SUMOvera system that is composed of the SUMOvera protein and its cognate protease named SUMOvera protease. Our results show that the SUMOvera system is fully functional for protein expression and purification using different eukaryotic hosts. Moreover, we proved that SUMOvera protease is the first SUMO-specific protease that can be over-expressed in yeast without causing cell death. Last, we showed that SUMOvera and SUMOstar systems have fully orthogonal specificities that can be exploited for the purification of equimolar protein complexes in *S. cerevisiae*. The main applications of the SUMOvera system as well as the comparison between SUMOvera protease and other commonly used site-specific proteases are discussed in the next sections.

4.2.1 SUMOvera as a tool for the enhancement of protein solubility

The low solubility of recombinant proteins is one of the most encountered problems during protein expression in *E. coli*. Normally, fusion protein tags (*i.e.* MBP, the N-terminal domain of IF2, TX-tag and GST-tag) are fused to the recombinant protein of interest to increase its expression levels and solubility (reviewed in Walls & Loughran 2011). SUMO proteins from *S. cerevisiae* (scSUMO) and *B. distachyon* (bdSUMO) have also been used to increase the solubility of recombinant proteins expressed in *E. coli* (Marblestone et al. 2006; Frey & Görlich 2014a). We show here that although bdSUMO was mutated in order to create SUMOvera, this did not impair its properties as a solubility enhancer fusion tag. Similar to bdSUMO and scSUMO proteins, SUMOvera also enhanced protein expression levels and solubility of eGFP in *E. coli* (data not shown). Despite the fact that SUMOvera was designed for being used in eukaryotic hosts, SUMOvera is also a good option as a fusion tag for the expression of recombinant proteins in *E. coli*. In fact, we have purified in high yields more than 30 different SUMOvera-tagged eukaryotic proteins expressed in *E. coli*.

4.2.2 SUMOvera protease activity

The removal of fusion tags from recombinant proteins is often mediated by site-specific proteases. In agreement with a previous study (Frey & Görlich 2014a), we showed that ubiquitin-like proteases are the most efficient proteases for tag removal. We observed that bdSENP1 is the most active ubiquitin-like protease *in vitro*, being 4-fold more efficient than SUMOvera protease. However, we showed that SUMOvera protease is from 3- to 25-fold more active than bdNEDP1, xlATG4B and xlSub2. Therefore, we believe that SUMOvera protease is a highly attractive alternative for efficient tag removal.

SUMOvera protease also represents a better tool for tag removal compared to commercially available proteases, i.e. TEV protease, thrombin, human rhinovirus 3C protease and scUlp1. For instance, SUMOvera protease is around 100-fold more active than TEV protease under the same *in vitro* conditions. Although we did not test the cleavage efficiency of thrombin and human rhinovirus 3C protease, both proteases would require longer incubation times and higher substrate/protease ratios for complete protein cleavage (Rubio et al. 2005; McCluskey et al. 2007) as compared to SUMOvera protease. Furthermore, SUMOvera protease is able to tolerate 5-fold higher salt concentration in

comparison to scUlp1 protease derived from *S. cerevisiae* (data not shown). During the course of this study, we have routinely used SUMOvera protease to cleave SUMOvera-tagged proteins expressed in *E. coli* and *S. cerevisiae*. Therefore, we consider that SUMOvera protease is a real alternative for tag removal compared to conventional site-specific proteases such as TEV protease, human rhinovirus 3C protease and thrombin.

We believe that bdSENP1 mutants with a higher proteolytic activity than SUMOvera protease can be obtained by using our *in vivo* system with more stringent conditions of selection than the one used during the evolution of SUMOvera protease. In addition, SUMOvera protease could be subjected to random mutagenesis in order to drastically improve its proteolytic activity. Last, a more active bdSENP1 mutant could be achieved by creating a protease that contains the 4 most frequent mutations obtained during the selection of SUMOvera protease (section 3.1.4).

4.2.3   SUMOvera as a fusion tag in *S. cerevisiae*

As described in section 2.13 , the use of the SUMO-fusion technology is completely hampered in yeast by the presence of endogenous SUMO-specific proteases. So far, the only available option to circumvent this problem was the protease-resistant SUMOstar tag and its cognate SUMOstar protease (Yan et al. 2009). Here, we offer SUMOvera as the second SUMO fusion protein that is completely stable in *S. cerevisiae*. Our results demonstrate that SUMOvera-tagged proteins are not cleaved *in vivo* by endogenous yeast SUMO-specific proteases. In fact, SUMOvera-tagged substrates expressed in yeast turned out to be more stable than SUMOstar-tagged proteins. In addition to SUMOvera and SUMOstar, other ubiquitin-like proteins (xlLC3B and bdNEDD8) have also been shown to work as stable fusion tags in yeast (Frey & Görlich 2015). It is worth mentioning that experiments performed in our study showed that proteins tagged with SUMOvera and xlLC3B are more stable in yeast compared to SUMOstar- and bdNEDD8-tagged proteins (data not shown).

SUMOvera tag can be exploited as a tool to produce recombinant proteins that require to be expressed in a eukaryotic host, such as *S. cerevisiae*, for proper protein folding or posttranslational modification. In combination with an N-terminal His-tag, SUMOvera-tagged proteins can be over-expressed in yeast and further purified using a nickel chelate matrix. During protein purification, SUMOvera fusion proteins can be eluted by on-

column cleavage using SUMOvera protease in order to achieve a higher degree of purity compared to standard elution methods. In fact, we have exploited SUMOvera and SUMOvera protease for the successful expression in yeast and further purification of yeast eEF2, *Xenopus laevis* Nup93, citrine, Nmd3, MBP, and human NAP1L1 (data not shown).

Although SUMOstar protease is also a suitable choice for tag removal of recombinant proteins produced in eukaryotic hosts, SUMOvera protease can be produced in higher yields with more purity using a single purification step. Unlike SUMOvera protease, SUMOstar protease co-purifies with different *E. coli* proteins and nucleic acids after imidazole elution, which results in the need of extra purification steps to obtain a protease preparation with decent quality (data not shown).

In summary, we are confident to say that the SUMOvera system is the most efficient technology for protein expression in yeast and purification due to its extremely protease-resistant feature. Furthermore, the high efficiency of SUMOvera protease (5-fold more active than xlLC3B cognate protease) enables the fastest method for on-column or in-solution tag removal of proteins that needed to be over-expressed in in *S. cerevisiae*.

4.2.4   SUMOvera tag can be used in multiple eukaryotic systems

We tested the applicability of the SUMOvera protein in different cellular extracts derived from higher eukaryotes. In combination with previous studies for SUMOstar fusion proteins (Zuo, Mattern, et al. 2005; Liu et al. 2008), our work shows that SUMOvera- and SUMOstar-tagged proteins are highly stable in the following cellular extracts: in *Xenopus laevis* egg extract, rabbit reticulocyte extract, wheat germ extract, HeLa cell extract, *Leishmania tarentola* cell extract and Drosophila S2 cell extract. Therefore, these results suggest that SUMOvera and SUMOstar proteins are not recognized by endogenous SUMO-specific proteases from amphibians, plants, insect, protozoa and human cells. Moreover, SUMOvera, in combination with another fusion tag or any affinity tag, represents a powerful technology to improve protein expression, protein purification and protein localization in eukaryotic organisms. To sum up, we believe that SUMOvera-fusion technology can be applied to a countless number of different applications using virtually any eukaryotic model system described so far. In fact, we are currently testing different applications involving the expression of SUMOvera-tagged proteins in eukaryotes.

4.2.5 Purification of protein complexes in *S. cerevisiae*

Recently, a multi-step affinity chromatography technique was developed to purify dimeric protein complexes with a well-defined stoichiometric composition (Frey & Görlich 2014b). After the co-expression of the tagged subunits, the purification technique relies on the use of two different affinity matrices and two proteases with orthogonal substrate specificity for on-column protein elution. Although this method represents an elegant and a streamlined process, it is restricted to the purification of proteins over-expressed in *E. coli*. Here, we showed that the fully orthogonal and eukaryotic protease-resistant SUMOvera and SUMOstar systems are compatible tools to purify dimeric protein complexes in *S. cerevisiae*. Specifically, we co-expressed in yeast and further purified the citrine•enhancer (Heikal et al. 2000; Kirchhofer et al. 2010) and the anti-MBP Darpin•MBP complexes (Binz et al. 2004) with a high degree of purity and a well-defined subunit stoichiometry. Along these lines, we foresee that SUMOvera and SUMOstar tags can be used together with other eukaryotic protease-resistant fusion tags (*i.e.* xlLC3B and bdNEDD8) (Frey & Görlich 2015) for the purification of protein complexes with at least 4 different subunits. In combination, these four fusion tags would improve existing purification protocols for difficult-to-express heterotetrameric or hererotrimeric protein complexes such as the endosomal sorting complex (ESCRT-I) and the Protein G complex, respectively (Kostelansky et al. 2007; Katzmann et al. 2001; Chan et al. 2011).

As mentioned before, the purification of protein complexes with defined stoichiometry requires two proteases with strict orthogonal substrate specificity. Despite the fact that SUMO-specific proteases from different species are the most active site-specific proteases described so far (Frey & Görlich 2014a), they could not be considered in the original design of the method due the cross-reactivity between them (Frey & Görlich 2014b). Instead, the method was developed in *E. coli* using two orthogonal ubiquitin-like proteases, bdSENP1 and bdNEDP1 from *B. distachyon*. SUMOvera protease and SUMOstar protease are completely orthogonal and 3- to 6-fold more active than bdNEDP1, we therefore believe that the use of both SUMO-specific proteases is a more efficient alternative for the purification of protein complexes not only in yeast but also in *E. coli*. In fact, we successfully used SUMOvera protease and SUMOstar protease for the purification of import and export complexes expressed in *E. coli*.

4.2.6   SUMOvera protease can be overexpressed in *S. cerevisiae*

Site-specific proteases are typically used for *in vitro* assays. However, an obvious and interesting approach would be to exploit the high specificity of these proteases *in vivo*. So far, this idea has already been implemented, but only for TEV protease (Gruber et al. 2003; Higuchi & Uhlmann 2005; Chen et al. 2010; Harder et al. 2008; Sato & Toda 2007). Despite its poor catalytic activity (Frey & Görlich 2014a), TEV protease is frequently used for protein cleavage *in vivo* for the simple reason that its short recognition site (ENLYFQS) is practically absent in all eukaryotic organisms. In contrast to TEV protease, yeast SUMO-specific proteases cannot be used for *in vivo* protein cleavage as over-expression of such proteases results in cell death (Bohnsack et al. 2004; Li & Hochstrasser 2003). Here, we showed that unlike wild type SUMO-specific proteases and SUMOstar protease, SUMOvera could be over-expressed in *S. cerevisiae* without causing cellular death. We speculate that the over-expression of SUMOvera protease is tolerated in *S. cerevisiae* because it does not cause massive deSUMOylation of endogenous proteins as compared to all other tested SUMO-specific proteases. However, the cause for cell death in yeast upon expression of a SUMO-specific protease still needs to be determined.

We considered that SUMOvera protease is a better choice for *in vivo* site-specific protein cleavage as compared to the TEV protease due to several reasons. First, SUMOvera protease is a more active enzyme. Second, SUMO proteases can work efficiently in very different conditions (pH, temperature, ionic strength), meaning that SUMOvera could be highly active in organelles such as mitochondria or peroxisomes. Last, SUMOvera generates cleaved proteins without an exogenous N-terminal amino acid. This last feature becomes essential for cases where a specific sequence at the N-terminus of the target protein plays a crucial role for the intended *in vivo* assay (Taxis & Knop 2012).

To sum up, we showed here the first successful example of the over-expression of a highly active SUMO-specific protease in *S. cerevisiae*. We believe that SUMOvera protease in combination with SUMOvera protein represent so far the best option for site-directed protein cleavage in eukaryotic living cells. Currently, we are designing experiments to target specific SUMOvera-tagged proteins for degradation in the proteasome upon protein cleavage by SUMOvera protease.

4.3   A new tandem affinity purification approach

In this study, we present a novel tandem affinity purification (TAP) protocol that uses a novel affinity protein tag. This tag is composed of the E and D domains form the staphylococcal ProtA as the first affinity tag, SUMOvera as a protease cleavage site and a His-tag as the second affinity tag. In the next paragraphs, we discuss the advantages of our novel TAP technology over conventional protocols.

Most of the reported TAP protocols use a single or double copy of the synthetic Z domain derived from ProtA (Nilsson et al. 1987) for the initial binding of the tagged bait to a IgG-coupled matrix. Here, we improved this initial binding step by replacing the low binding capacity and expensive IgG-coupled matrix for a solid support coupled to the anti-Z domain affibody named ZpA963 (Lindborg et al. 2013). Production and coupling of anti-Z domain affibody to sepharose beads is cheaper compared to the commercially available IgG-coupled matrices. Also, anti-Z domain affibody sepharose matrix has a higher protein binding capacity as compared to IgG-coupled matrices (unpublished data). We also improved the initial immobilization of the bait protein by replacing the double copy of the synthetic Z domain for the E and D domains from ProtA. The E and D domains represent a better affinity tag as they have a higher binding affinity to the ZpA963 affibody than the doubled Z domain (Lindborg et al. 2013). Together, the ZpA963 affibody and the ED domains from ProtA represent the first advantage of our novel TAP protocol.

Most reported TAP protocols employ a protease-mediated elution step to recover the protein binders from the first affinity matrix used. Although protein elution using a site-specific protease considerably increases the selectivity by reducing background binding, only highly inefficient proteases such as TEV protease and human rhinovirus 3C protease have been employed so far (Rigaut et al. 1999; Cheeseman & Desai 2005; Giannone et al. 2007; Rubio et al. 2005; Bürckstümmer et al. 2006). Therefore, using TEV protease or human rhinovirus 3C protease for TAP results in complete elution of the target protein and extremely long incubations times ($\cong$16 h). To solve this problem, we envision that a SUMO-specific protease would drastically reduce the time needed for efficient and complete protein elution. SUMOvera protein in combination with SUMOvera protease allowed us to achieve complete protein elution of the SUMOvera-tagged bait from the affinity matrix within less than 1 h at 4°C. Also, only nanomolar concentrations of SUMOvera protease are needed to completely elute all protein bound to the matrix,

whereas protein elution using TEV protease would require high micromolar concentrations (Schimanski et al. 2005). Therefore, we believe that the use of SUMO protease is the second advantage of our novel TAP system as it represents a cheaper and a faster way to achieve complete and highly selective protein elution.

The use of a His-tag gives the opportunity to elute only the isolated proteins binders and keep the His-tagged bait bound to the Ni^{2+} chelate matrix when using a denaturing elution buffer. In our assays, we used as denaturing elution buffer a solution containing Gdn-HCl and a basic pH. The separation of binder proteins from the His-tagged bait using a Gdn-HCl-containing solution offers several advantages for further analysis of the eluted samples. First, Gdn-HCl-containing samples are compatible with trypsinization protocols prior to MS analysis (Proc et al. 2010). Therefore, there is no need of extra preparation steps for protein identification by MS. Second, the absence of the high amount of bait in the final sample improves the identification of isolated low abundant protein binders by LC/MS. Third, protein elution with Gdn-HCl helps to visualize binders that would co-migrate on SDS-PAGE with the His-tagged bait. For instance, the separation of the His-tagged Pdr6 from its protein binders was an extremely useful step for the identification of eEF2 as export cargo. The large amount of Pdr6 used as bait masked the visualization of eEF2 (Figure 3.20. Lane 5). Thus, eEF2 factor was only visible in the polyacrylamide gel only after Gnd-HCl elution (Figure 3.20. Lane 7).

In summary, we consider that our novel method to identify protein binders of a specific bait is a more efficient, highly selective, MS-compatible and a less time consuming alternative to most of the available TAP methodologies described so far.

4.4   How many bidirectional NTRs are present in *S. cerevisiae*?

Exportins have an rather low affinity to Ran-GTP and their interaction is only detectable in the presence of the export cargo (Hellmuth et al. 1998; Maurer et al. 2001). In contrast, importins have a high affinity to Ran-GTP and that might explain why the import cargos are rapidly released upon translocation into the nucleus (Görlich et al. 2003). In this study, we asked the question whether NTRs with a medium affinity to Ran-GTP might work as exportins as well as importins.

As mentioned earlier, a medium affinity to Ran-GTP seems to be a particular characteristic for NTRs that can act as an import as well as an exportin. A previous work found that 7 yeast NTRs (Kap114, Yrb4, Nmd5, Msn5, Mtr10, Pdr6 and Lph2) have a medium affinity Ran-GTP (Hahn & Schlenstedt 2011). However, the authors of this work completely ignored the possibility to acknowledge that these 7 NTRs might act as bidirectional NTRs. We therefore performed a series of *in vitro* and *in vivo* experiments and proved that Lph2 and Pdr6 act as bidirectional NTRs in budding yeast. Specifically, we found that eIF4A is an export cargo for Lph2 whereas Pdr6 exports eIF5A and eEF2.

Conventional importins or exportins consume one molecule of GTP per every cargo transported across the nuclear envelope. In contrast, bidirectional NTRs transport two cargos per every GTP molecule consumed. Bidirectionality therefore seems to be an improved feature for NTRs as it represents a reduction in the energy consumed for nucleocytoplasmic cargo transport. However, it has been proposed that bidirectional transport might have the shortcoming of limiting the amount cargo accumulated in the final cellular comportment (Mingot et al. 2001). Consequently, it is reasoned that bidirectionality in transport might be present in exceptional cases rather than a common feature for NTRs. Here, our results support the idea that bidirectional NTRs are more recurrent phenomenon that previously believed, and that bidirectionality in cargo transport might represent a more common process for cells to consume less metabolic energy to transport two different cargos into two distinctive cellular compartments. Our findings might open a very important question in the field of nucleocytoplasmic transport: are there still other bidirectional NTRs to be characterized in other higher eukaryotes?

At this point, Msn5 (Kaffman, Rank, O'Neill, et al. 1998; Yoshida & Blobel 2001), Lph2 (3.3.2) and Pdr6 (3.4.2 and 3.4.4) are the only described bidirectional NTRs in *S. cerevisiae*. Currently, we are testing whether other poorly characterized yeast importins (*i.e.* Mtr10, Nmd5, Yrb4 or Kap114) might also act as novel exportins.

4.5   Lph2 (Kap120) is a novel bidirectional NTR in yeast

Here, we described that Lph2 is not only a importin but it also acts an export in *S. cerevisiae*. On one hand, we identified that Lph2 exports and binds in a Ran-GTP dependent manner to the yeast initiation factor eIF4A. On the other hand, we found that rpL12A is a novel import cargo for Lph2. Although only eIF4A was fully characterized *in*

vivo and *in vitro* as a new export cargo, our data suggest that several other yeast proteins (Kre11-1, Rvs167 and Rvs161) could also be exported by Lph2. Our findings strongly suggest that Lph2 is not a NTR highly specialized in the transport of few cargos but an NTR with a wider cargo recognition spectrum.

The yeast translation initiation factor eIF4A binds to the 40S subunit at the beginning of translation and unwinds secondary structures of the mRNA (Rogers et al. 1999). eIF4A also recognizes, in complex with other translation factors, the 5' cap region of the mRNA in order to promote an efficient translation initiation (Jaramillo et al. 1991; Neff & Sachs 1999). Therefore, we speculated that Lph2 prevents the interaction between immature ribosomal subunits and different initiation factors in complex with eIF4A to prevent interference with ribosomal biogenesis. We also envision that Lph2 is actively suppressing premature nuclear translation of non-spliced mRNAs by actively excluding eIF4A from the nucleus. In fact, the argument that Lph2 as chaperone protein to prevent nuclear translation are supported by the fact that nuclear export of several translation factors by different NTRs is a highly recurrent process in several organisms (Lipowsky et al. 2000; Bohnsack et al. 2002; Calado et al. 2002; Kirli et al. 2015). Last, eIF4A is the first reported export cargo for the NTR Lph2. Thus, Lph2 is the third example of an NTR in *S. cerevisiae* that can import and export different sets of proteins across the nuclear envelope.

Besides new export cargos, we also identified rpL12A as a potential import cargo for Lph2 (3.3.1). While characterizing the Lph2•rpL12A import complex, we observed that Gsp1Q71L-GTP did not cause the disassociation of rpL12A from Lph2. One could therefore assume that rpL12A might not be a real import cargo but just a Lph2 binding protein. However, it is possible that the intrinsic low affinity of Lph2 for Gsp1-GTP (270 nM, Hahn & Schlenstedt 2011) is insufficient to cause the proper release of rpL12A. If so, the binding of an export cargo (*i.e.* eIF4A) and Gsp1-GTP to Lph2 might be required to assist the complete disassociation of rpL12A from Lph2. In fact, import cargo release aided by Gsp1-GTP and an extra protein factor have been observed for different importins (Greiner et al. 2004; Mingot et al. 2001; Senger et al. 1998; Caesar et al. 2006). This situation could be beneficial for the cell as Lph2 might serve as a chaperone and deliver rpL12A to the final destination where it is transferred to other binder partners (*i.e.* other nuclear chaperones or eventually the ribosomal 60S subunit). Further experiments have to be performed in order to prove this plausible scenario.

4.6   Pdr6 (Kap122) acts as a bidirectional NTR in *S. cerevisiae*

Pdr6 was first identified as a member of a protein family that causes pleotropic drug resistance in *S. cerevisiae* (Chen et al. 1991; Balzi et al. 1987). Later, completely different properties were described for Pdr6 (Titov & Blobel 1999). First, Pdr6 was defined as a protein that belongs to the Impβ-like NTR family due to the sequence homology of its N-terminal domain (Görlich et al. 1997). Second, it was shown that Pdr6 interacts with different nucleoporins (Nup1 and Nup2) and is evenly distributed in the cytoplasm and nucleus of *S. cerevisiae*. Last, Pdr6 was described as an importin for the transcription factor TFIIA. Despite the fact that Pdr6 might act as a bidirectional NTR due to its medium affinity to Ran-GTP (231 nM, Hahn & Schlenstedt 2011), there has not been any reported study trying to identify potential export cargos for Pdr6. In this study, we identified that translation factors eIF5A and eEF2 are the first two export cargos for Pdr6. In agreement with their function, translation elongation factors eEF2 and eIF5A are localized in the cytoplasm of *S. cerevisiae* (Valentini et al. 2002; Perentesis et al. 1992; Gregio et al. 2009). Therefore, *S. cerevisiae* counts on Pdr6 to actively export eEF2 and eIF5A out of the nucleus.

Two criteria are met by eIF5A and eEF2 to be considered Pdr6 export cargos. First, both translation factors bind to Pdr6 in a Gsp1-GTP dependent manner. Second, export of eIF5A and eEF2 in inhibited *in vivo* by the absence of Pdr6. Additionally, our results also indicate that export of eIF5A is mediated specifically by Pdr6 unlike several other cargos that are transported by many independent pathways (Caesar et al. 2006; Mühlhäusser et al. 2001; Greiner et al. 2004; Rout et al. 1997).

Splicing of pre-mRNAs in the nucleus must be achieved before the mRNA can be translated in the cytoplasm. Premature nuclear translation of pre-mRNAs containing introns could cause the production of aberrant proteins with a dominant negative mutant function. Efficient nuclear translation would require a complete set of translation factors in optimal concentrations. Translation factors "leak" into the nucleus even if their size is above the passive diffusion limit at the NPC (≈40kDa). In addition, translation factor can be located in the nucleus after the mixing of cytoplasm and nuclear contents during mitosis. As a consequence, active export of translation factors is one of the strategies used by eukaryotic cells to prevent translation inside the nucleus (Lipowsky et al. 2000; Bohnsack et al. 2002; Calado et al. 2002; Kirli et al. 2015). Our results support the idea by

showing that Pdr6 suppresses nuclear translation in *S. cerevisiae* by actively exporting essential translation factor (eEF2 and eIF5A). Export of eiF5A is achieved by Exportin 4 in humans (Lipowsky et al. 2000), however, the NTR responsible for the export of eEF2 in humans has not been identified yet.

Besides the identification of eEF2 and eIF5A as novel export cargos, we also described the discovery of new import cargos for Pdr6. We showed that Pdr6 is a NTR responsible for the import of the E2 SUMO-conjugating enzyme Ubc9 in *S. cerevisiae*. As most of the protein SUMOylation takes place in the nucleus (Kamitani et al. 1997; Seeler & Dejean 2003), the active import of Ubc9 represents an important process for the proper function of the cell. In addition, Ubc9 is the only E2-conjugating enzyme present in *S. cerevisiae*, therefore, this essential protein needs to be transported to nucleus as it is the only choice in the cell for the SUMOylation of target proteins. Therefore, our results suggest that Pdr6 plays an important role in the SUMO pathway by actively importing Ubc9 to achieve proper SUMOylation in time and space of target proteins. The import of Ubc9 (3.4.3) and TFIIA (Titov & Blobel 1999) as well as the export eEF2 (3.4.4) and eIF5A (3.4.2), proves that Pdr6 is another bidirectional NTR together with Lph2 and Msn5 (Kaffman, Rank, O'Neill, et al. 1998; Yoshida & Blobel 2001) in *S. cerevisiae*. Moreover, our results demonstrate that Pdr6 has broader cargo recognition spectrum than as previously thought (Titov & Blobel 1999).

Even though Pdr6 does not have any ortholog protein in human (Malik et al. 1997; Görlich et al. 1997; Quan et al. 2008).It is likely that the common ancestor of yeast and human had no Pdr6 and consequently *S. cerevisiae* had a gene duplication that did not occur in higher eukaryotes. Nevertheless, we noticed that the cellular functions of Pdr6 (export of eIF5A and import of Ubc9) are conserved and taken over by two different humans NTRs. In human cells, exportin 4 (Lipowsky et al. 2000) and importin 13 (Mingot et al. 2001) mediate the export of eIF5A and the import of Ubc9, respectively. In addition, human exportin 4 and importin 13 evolved further as they mediate the translocation of several cargos using transport pathways that have not been yet described in *S. cerevisiae* (Kurisaki et al. 2006; Gontan et al. 2009; Mingot et al. 2001). In the light of these observations, we bring two questions to the discussion. Why were export of eIF5A and import of Ubc9 also "developed" in human cells after millions of years of evolution? and why did exportin 4 and importin 13 evolve and acquire extra transport functions that don't seem to be performed by Pdr6? We speculate that the export of eIF5A and import of Ubc9 were kept

during evolution since both cargos are involved in fundamental cellular processes, translation and SUMOylation, respectively. Therefore, eIF5A and Ubc9 are required to be in the right subcellular location at the right time for optimal cellular performance. We also believe that acquirement of new functions by exportin 4 and importin 13 as compare to Pdr6 offers the advantage of creating new and tighter ways of regulating the subcellular localization of specific cargos. Moreover, the further evolution of exportin 4 and importin 13 allows the opportunity to transport a higher number of different proteins in more complex organisms such as humans.

4.7 Perspectives

For the last 10 years, it was thought that Lph2 and Pdr6 were highly specialized importins since only very few specific cargos were known for each NTR (Titov & Blobel 1999; Caesar et al. 2006; Kim et al. 2010). In this study, we found that Lph2 and Pdr6 are in fact bidirectional NTRs with wide cargo recognition spectra. Lph2 recognizes Rpf1, rpL12A and eIF4A as cargos, whereas Pdr6 transport TFIIA, Ubc9, eIF5A and eEF2. We now aim to investigate in detail how these two NTRs are able to recognize their specific cargos. Do Lph2 and Pdr6 recognize a similar linear sequence shared by their cargos? Is a conformational motif in the cargo that is recognized by Lph2 and Pdr6? Is the recognition of each cargo by Lph2 or Pdr6 completely different from each other? In order to answer these questions, we are currently trying to crystalize Lph2 and Pdr6 in complex with their different cargos and Gsp1Q71L-GTP. In fact, Dr. Metin Aksu has gotten promising crystallization conditions with Pdr6•eIF5A•Gsp1Q71L-GTP and Pdr6•Ubc9•Gsp1Q71L-GTP complexes. We are currently trying to optimize the expression conditions in *E. coli* of recombinant Lph2 in order to crystalize Lph2•eIF4A•Gsp1Q71L-GTP and Lph2•rpL12A complexes. Furthermore, we will also try to obtain the protein structure of the "free" Lph2 and Pdr6 to understand how conformational changes in the NTRs are induced upon cargo binding.

The identification of low abundant cargos for Lph2 and Pdr6 is another import aspect that we want to investigate further. We plan to analyze by mass spectrometry the interactome of Lph2 and Pdr6. Specifically, we want to identify what are the proteins most enriched in the presence and absence of Gsp1Q71L-GTP as compared to protein amount found the input material. We believe that the identification of several cargos Lph2 and Pdr6 will give

us a great opportunity to understand how these two different NTRs impact and regulate specific cellular processes in *S. cerevisiae*.

As mentioned before, we strongly consider that there might be other bidirectional NTRs in yeast to be identified. Therefore, we are currently performing binding experiments using poorly studied yeast NTRs with a medium affinity to Gsp1-GTP as baits to identify novel export and import cargos. Overall, we seek to demonstrate that bidirectionality in NTRs is more common phenomenon than currently thought.

5 MATERIALS

5.1 Reagents

All chemical reagents used for this study were purchased from the following companies: Calbiochem (San Diego, CA, USA), Qiagene (Hilden, Germany), Roche (Mannheim, Germany), Roth (Karlsruhe, Germany), Sigma-Aldrich (St. Louis, MO, USA), AppliChem (Germany), New England Biolabs (USA), Serva (Germany), GibcoBRL-Life Technologies (Paisley, UK), Merck (Darmstadt, Germany), MoBiTech (Göttingen, Germany), Pharmacia (Uppsala, Sweden) and Promega (Madison, WI, USA).

5.2 Buffers and solutions

Table 5.1 Buffers and solutions used during manipulation of nucleic acids (DNA and RNA)

Buffer or solution	Composition
Orange G sample buffer	10 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 50 % (w/v) glycerol and 25 % (w/v) Orange G
TAE buffer (50x)	242 g Tris base, 57.1 ml acetic acid, and 100 ml 0.5 M EDTA (pH 8.0) in 1 l H_2O
Ethidium bromide	20 mg/ml ethidium bromide in water
T4 DNA Ligase buffer (10x)	500 mM Tris (pH 7.5), 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP and 250 μg/ml BSA
Phusion HF buffer (5x)	Thermo Scientific, USA
PfuS Triple Mix (100x)	Self-made mixture containing 100 ng/μl PfuS polymerase, 15 ng/μl pyrophosphatase, and 6 ng/μl dUTPase
dNTPs (10x)	2.5 mM of each desoxynucleotide in ddH_2O
Gibson Assembly mix (2x)	Self-made mixture containing 5' Exonuclease, DNA polymerase and DNA ligase
DNA ladder	GeneRuler 1 kb Plus (Thermo Scientific, USA)
Tth buffer	100 mM Tris-HCl (pH 8,9), 15 mM $MgCl_2$, 1M KCl, 500 μg/mL BSA, 0,5% Tween20
10× TE (Tris-EDTA) buffer	100 mM Tris/HCl (pH 8.0), 5 mM EDTA
LiAc/TE Solution	0,1 M LiAc in 1×TE buffer
PEG/LiAc/TE solution	40% (w/v) PEG6000, 1 M LiAc in 1×TE buffer
Salmon sperm DNA	10 mg/ml salmon sperm DNA in 1x TE buffer

Table 5.2 Buffers or materials used during protein biochemistry work.

Buffer or material	Composition
SDS sample buffer	3% (w/v) SDS, 125 mM Tris/HCl (pH 6.8), 50 mM DTT, 1 M sucrose, coomassie brilliant blue G250
Protein ladder (SDS-PAGE)	PageRuler Unstained (Thermo Scientific, USA)
Protein ladder (Western Blot)	PageRuler Plus Prestained (Thermo Scientific, USA)
SDS-running buffer (10x)	150 g glycine, 30 g Tris base, and ddH$_2$O to 1 l volume
Coomassie staining solution	2 % (w/v) Coomassie brilliant blue G250 in 3% (v/v) acetic acid
Amido Black staining solution	0.2 % (w/v) Amido Black in 2 % acetic acid
1x PBS	137 mM NaCl, 2.7 mM KCl, 8.1 mM Na$_2$HPO$_4$, 1.76 mM KH$_2$PO$_4$
Cell resuspension buffer	45 mM Tris/HCl pH (7.5), 150-500 mM NaCl, 2-20 mM imidazole, 5% (v/v) glycerol, 4.5 mM MgCl$_2$, 0-5 mM DTT
Washing buffer	45 mM Tris/HCl pH (7.5), 500-2000 mM NaCl, 4.5 mM MgCl$_2$, 2-20 mM imidazole, 0-5 mM DTT
Ni^{2+}-elution buffer	45 mM Tris/HCl pH (7.5), 150-500 mM NaCl, 400 mM imidazole, 4.5 mM MgCl$_2$, 250 mM sucrose, 0-5 mM DTT
Binding buffer	45 mM Tris/HCl pH (7.5), 100 mM NaCl, 2 mM MgCl$_2$, 5 mM DTT
Gdn-HCl elution buffer	3 M guanidinium chloride, 50 mM Tris/HCl pH (8.0)
GK75 buffer	20 mM HEPES-KOH pH (7.9), 1.5 mM MgCl$_2$, 75 mM KCl, 5% (v/v) glycerol, 0.01% (v/v) NP40, 0.5 mM DTT)
Protein cleavage buffer	45 mM Tris/HCl pH (7.5), 250 mM NaCl, 2 mM MgCl$_2$, 250 mM sucrose, 10 mM DTT
Colloidal Coomassie stock solution	0.08% (w/v) coomassie brilliant blue G250, 1.6% (v/v) ortho-Phosphoric acid, 8% (w/v) ammonium sulfate, 20% (v/v) methanol
Blotting buffer	100 ml 10x SDS-PAGE running buffer, 200 ml MetOH, 0.03% (w/v) SDS, add to 1L with ddH$_2$O
Washing and blocking buffer	4% (w/v) powdered milk in 1x PBS
Protease-inhibitor mix (500x)	5 mg/ml aprotinin, 5 mg/ml leupeptin, 2.5 mg/ml elastatinal, 2.5 mg/ml chymostatin, 0.5 mg/ml pepstatin A

5.3 Laboratory instruments

Table 5.3 Equipment used during this study

Instrument	Manufacturer
Analytical scale	Sartorius, Germany
Arium® Pro UV ultrapure water system	Sartorius, Germany
Agilent 2100 Bioanalyzer	Agilent Technologies, USA
Eppendorf 5415R and 5424 tabletop centrifuge	Eppendorf, Germany
MicroPulser™ electroporation apparatus	Bio-Rad Laboratories Inc., USA
NanoDrop 2000c	PeqLab, Germany
Thermomixer comfort	Eppendorf, Germany
Epson Perfection V700 Photo scanner	Epson, Japan
Sonifier 450	Branson Ultrasonics, UK
Odyssey infrared imaging system	LI-COR Biosciences, USA
Eppendorf Biophotometer	Eppendorf, Germany
Incubator/ Climo-Shaker ISF1-X	Adolf Kühner AG, Switzerland
TCS SP5 confocal laser scanning microscope	Leica, Germany
M165C stereo microscope	Leica, Germany
Thermocycler	SensoQuest, Germany
DNA gel documentation system	Vilber Lourmat, Switzerland
Äkta Purifier + Äkta Explorer	Pharmacia, Sweden
MDF 793 -80 °C freezer	Sanyo (Osaka, Japan)
Perfection V700 photo scanner	Epson (Long Beach, California, USA)
UV table	Benda Laborgärate (Wiesloch, Germany)
Multifuge 3L-R	Heraeus, Germany
Sorval RC6+ centrifuge	Sorvall/Thermo Scientific, USA
Sorval Lynx 6000	Sorvall/Thermo Scientific, USA
Sorval Wx Ultra ultracentrifuge	Sorvall/Thermo Scientific, USA
Discovery M120 SE ultracentrifuge	Sorvall/Thermo Scientific, USA
Research pro multi channel pipette	Eppendorf AG, Hamburg, Germany

5.4 Software and online tools

Table 5.4 Software used during this study

Software	Developer
Adobe Illustrator + Photoshop CS5.1	Adobe Systems Inc., USA
Microsoft Excel for Mac 2011	Microsoft Corp., USA
UCSF Chimera 1.10.1	Resource for Biocomputing, Visualization, and Informatics, UCSF, USA
CFX Manager^{TM} 3.1.1517.0823	Bio-Rad Laboratories Inc., USA
Odyssey 3.0.30	LI-COR Biosciences, USA
Oligo 7.58	Molecular Biology Insights Inc., USA
SeqBuilder; Protean; MegAlign; SeqMan; Gene Quest – all version 11.2.1	DNASTAR, USA
Gene Designer 2.0	DNA2.0, USA
Scaffold 4	Proteome Software, USA
LASAF	Leica, Germany
UNICORN 5.0	Amersham Biosciences, Sweden
Mendeley	Mendely Ldt., USA
Python 2.7.10	Python Software Foundation
Mac OS X Yosemite Vers.10.10.5	Apple, USA
Image J	National Institute of health (USA)

Table 5.5 Online tools used during this work

Tool	Website
Biogrid	https://thebiogrid.org/
BLAST	http://blast.ncbi.nlm.nih.gov/Blast.cgi
ClustalW	http://ebi.ac.uk/Tools/msa/clustalw2
Uniprot	http://uniprot.org
PsiPred	http://bioinf.cs.ucl.ac.uk/psipred
Protein Data Bank	http://rcsb.org/pdb
Saccharomyces genome database	http://www.yeastgenome.org/
Saccharomyces genome deletion project	http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3.html
TermiNator	http://www.isv.cnrs-gif.fr/terminator3/index.html
ESPript 3.0	http://espript.ibcp.fr/ESPript/ESPript/
NEB tools	https://www.neb.com/tools-and-resources

5.5   *E. coli* and *S. cerevisiae* strains

Table 5.6 *S. cerevisiae* strains used in this study.

Strain	Purpose	Genotype	Supplier
SFY123	Protein expression	MATα ADE2 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 can1-100 H2B::CFP	Provided by S. Frey. (Frey & Görlich 2015)
BY4741	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3	Provided by B. Schwappach. (Brachmann et al. 1998)
AVY01	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔLph2	Provided by B. Schwappach (Winzeler et al. 1999)
AVY02	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔPdr6	Provided by B. Schwappach (Winzeler et al. 1999)
AVY04	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔLph2 TIF1::GFP	This study
AVY05	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔLph2 TIF1prL12A::GFP	This study
AVY06	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔPdr6 Ubc9::GFP	This study
AVY07	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔPdr6 EFT2::GFP	This study
AVY08	Confocal microscopy analysis	MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6 his3-11, 15 leu2-3, 112 LYS2 trp1-1 ura3 ΔPdr6 HYP2::GFP	This study

Table 5.7 List of *E. coli* strains used in this study.

Strain	Purpose	Genotype	Supplier
5-alpha	Cloning	fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	New England Biolabs, USA
10-beta	Cloning	Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14-φ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (Str^R) rph spoT1 Δ(mrr-hsdRMS-mcrBC)	New England Biolabs, USA
Bl21 (DE3)	Protein expression	B F$^-$ *ompT gal dcm lon hsdS$_B$(r_B^- m_B^-)* λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB$^+$*]$_{K-12}$($λ^S$)	Stratagene, USA
Rosetta	Protein expression	F$^-$ *ompT gal dcm lon? hsdS$_B$(r_B^- m_B^-)* λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB$^+$*]$_{K-12}$($λ^S$) pLysSRARE[*T7p20 ileX argU thrU tyrU glyT thrT argW metT leuW proL ori*$_{p15A}$](Cm^R)	Novagen, USA
NEB express F'	Protein expression and selection method	*fhuA2 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr)114::IS10*	New England Biolabs, USA
Top10 F'	Protein expression	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ$^-$	Invitrogen, USA

5.6   Culture media for bacterial and yeast growth

Table 5.8 Culture media used for bacterial growth

Media	Composition
SOB liquid medium	20 g tryptone, 5 g yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, and H$_2$O to 1 l
LB liquid medium	10 g tryptone, 5 g yeast extract, 10 g NaCl, and H$_2$O to 1 l
2YT liquid medium	16 g tryptone, 10 g yeast extract, 5 g NaCl, and H$_2$O to 1 l
TB liquid medium	12 g tryptone, 24 g yeast extract, 0.4% (v/v) glycerol, 72 mM K$_2$HPO$_4$, 17 mM KH$_2$PO$_4$

Table 5.9 Culture media used for yeast growth

Media	Composition
YPD medium	20 g tryptone, 10 g yeast extract, 2.0 % (v/v) glucose and H$_2$O to 1 l
CSM-URA	1.7 g yeast nitrogen base without (NH$_4$)$_2$SO$_4$, 50 g (NH$_4$)$_2$SO$_4$, 2.0 % (v/v) glucose, 770 mg CSM-mix without URA and H$_2$O to 1 l

All media were autoclaved before use with the following conditions (15 min, 121°C, 15 psi). Table 5.10 shows antibiotics working concentration used to select and cultivate transformed *E. coli* and *S. cerevisiae* cells. In order to prepare culture plates, 2YT, CSM-URA or YDP media were supplemented with 15 g/l of agar. Media for the cultivation of *S. cerevisiae* was supplemented before use with 2% (w/v) glucose from an autoclaved stock solution of 40% (w/v).

The antibiotics were added right before using the media. For the preparation of selective agar plates, the antibiotics were added when the media were approximately at 60°C. All antibiotic stock solutions were prepared according to the instructions provided by the manufacturer and immediately sterilized by filtration.

Table 5.10 Antibiotics used for this study

Antibiotic	Working concentration	Purpose
Ampicillin (Amp)	100 µg/ml	Bacterial growth
Kanamycin (Kan)	50 µg/ml	Bacterial growth
Spectinomycin (Spec)	100 µg/ml	Bacterial growth
Hygromycin B (HygB)	600 µg/ml	Bacterial growth
Hygromycin B (HygB)	300 µg/ml	Yeast growth
Geneticin (Gen)	100 µg/ml	Yeast growth

6 METHODS

6.1   Standard methods in molecular biology (plasmid preparation)

6.1.1   Electro-competent *E. coli* cells

All *E. coli* cells used in this study (Table 5.6) were produced by Gabriele Hawlitschek and Uwe Hoffmann according to (Sambrook et al. 1989). Bacterial cells were grown overnight in SOB medium at 37°C with the respective antibiotic(s). Next day, overgrown bacterial cultures were diluted with fresh SOB medium and further incubated at 37°C. Bacterial cells were incubated until the exponential growth phase was reached ($OD_{600} \approx 0.8$). Next, bacteria were pelleted by centrifugation at 5,000 rpm for 10 min at 4°C (RC6 plus centrifuge, F10 rotor; Sorvall, USA). After centrifugation, the supernatant was removed by

decantation and cells were resuspended in 1.4 % (w/v) glycerol. Bacteria were centrifuged and washed with glycerol one more time. Finally, cells were resuspended in 14 % (w/v) glycerol and aliquoted in 100 μl fractions. Aliquots were frozen in liquid nitrogen and incubated at -80°C until further use.

6.1.2   Transformation of electro-competent *E. coli* cells by electroporation

Typically, 40 of μl electro-competent cells were transformed with 50 ng of plasmid by electroporation. The transformation was performed in an electroporation cuvette using a GenePulser device (BioRad, Burlington, USA). Electroporated cells were then resuspended in 1ml of 2YT medium supplemented with 0.5 % (w/v) glucose. After suspension, cells were incubated for 1h at 37°C with continuous shaking. Next, cells were plated on agar containing the appropriated antibiotic(s) and incubated overnight at 37°C in order to select transformed cells. Single bacterial colonies were picked and further incubated in selective medium for DNA extraction or protein expression as described in 6.1.5 and 6.3.1 , respectively.

6.1.3   Purification and concentration determination of DNA

DNA was extracted and purified according to manufacture's instructions. For small-scale purification (mini-prep), DNA was purified using NucleoSpin Plasmid Kit (Marcherey Nagel, Germany). NucleoBond PC100 kit was used (Marcherey Nagel, Germany) when higher amounts of DNA were required (midi-prep). After purification, the concentration of DNA was determined using a NanoDrop 2000C spectrophotometer (PeqLab, Germany) according to the manufacturer's instructions.

6.1.4   Agarose gel electrophoresis

Agarose gel electrophoresis was performed as described in (Sambrook et al. 1989). Double stranded DNA fragments resulting after enzymatic restriction reactions or PCR were separated according to their size by agarose gel electrophoresis. The agarose concentration in the gel varied from 0.8% (w/v) to 2.0% (w/v) depending on the size of the analyzed DNA fragments. For the preparation of the gel, the corresponding amount of agarose was

dissolved in 1X TAE buffer. Ethidium bromide (3 μl of a 20mg/ml solution stock) was added to 100 ml of melted agarose solution. DNA samples were mixed with Orange-G sample buffer prior loading into the gel. For reference, a standardized DNA maker (1kb DNA ladder, Fermentas, St.Leon-Roth, Germany) was included during the analysis. Electrophoresis was carried out at 170V for 30 min in 1X TAE buffer using a Consort EV233 power supplier (Sigma-Aldrich, Germany). After the run, the DNA inside the agarose gel was visualized with UV light.

6.1.5   DNA extraction from agarose gels

DNA fragments loaded in agarose gels were extracted and further purified using the Zymoclean Gel DNA recovery kit (Zymo Research CA, USA) according to the manufacturer's instructions. DNA concentration was determined using a NanoDrop 2000C spectrophotometer (PeqLab, Germany).

6.1.6   Enzymatic restriction of DNA

Enzymatic restriction of DNA was performed for further cloning or analytical purposes. All restriction enzymes and the corresponding restrictions buffer used were purchased from NEB (New England Biolabs, USA). Typically, the reactions were performed as indicated by the manufacturer. For the complete cleavage of double stranded DNA, the restriction enzymes were used at least in a 2-fold excess per microgram of DNA during 2h at 37°C. Restriction reactions were stopped after the addition of Orange-G sample buffer. Finally, samples were analyzed by agarose gel electrophoresis as described in 6.1.4 .

6.1.7   Dephosphorylation of DNA

Digested DNA samples were treated with fast alkaline phosphatase (Thermo Scientific, USA) in order to dephosphorylate the 5' end of vector backbone. DNA was incubated with 2 μl of alkaline phosphatase (1 Unit/μl) for 1h at 37°C. The reactions were stopped with the addition of Orange-G sample buffer. Samples were analyzed by agarose gel electrophoresis and further purify as described in 6.1.5 .

6.1.8 Ligation of DNA fragments

Ligation of DNA fragments with blunt or sticky ends was performed using T4 ligase (prepared in the lab, 100ng/µl). For a standard DNA ligation reaction, 30 fmol of DNA used as backbone were mixed with 60 fmol of DNA insert. The mixture was then supplemented with 1 µl of T4 ligase stock solution and 1 µl of ligation buffer (500 mM Tris (pH 7.5), 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP and 250 µg/ml BSA). The ligation reaction was then filled up with ddH_2O to a total volume of 10 µl. Reaction was incubated for 1h at 37°C for the ligation of fragments with sticky end or overnight at 16°C for fragments with blunt ends. A sample lacking the insert was always used as a negative control. The ligation reaction was stopped by heat-inactivation of T4 ligase for 10 min at 70°C. Next, 1 µl of a inactivated reaction was used to transform electro-competent *E. coli* cells by electroporation as in 6.1.2 .

6.1.9 Polymerase chain reaction (PCR)

A homemade PfuS protein mix (PfuS polymerase, pyrophosphatase and dUTPase) was used for PCR. Typically, a 100 µl PCR reaction contained the following components: 50 ng template DNA, 1 µl PfuS Tripple Mix, 1 µl (100 µM) each of forward and reverse primer, 5 µl of 10x dNTPs mix (2.5 mM each), 20 µl of 5x Phusion HF buffer (Thermo Scientific, USA) and the corresponding amount of ddH_2O to make up to the final volume. PCR reactions were performed with a SensoQuest Lab cycler (SensoQuest GmbH, Göttingen, Germany). If not specified otherwise, PCR reactions were run using the following protocol: (1) initial denaturation at 98.5°C for 2 min, (2) denaturation at 98.5°C for 30 sec, (3) annealing at 58-62°C for 30 sec, (4) elongation at 72°C for 30 sec per 1 kb of PCR product length and (5) final extension at 72°C for 2 min. Steps (2)-(4) were repeated 25-30 times. The annealing temperatures were calculated using the Oligo 7.58 (Molecular Biology Insights Inc., USA). Restriction sites or Gibson assembly overhangs were introduced to both ends of the amplified DNA fragments with the primers for further cloning.

6.1.10 Gibson assembly reaction

Gibson assembly is a cloning technique that allows the joining of DNA molecules in a single isothermal step (Gibson et al. 2008; Gibson et al. 2009). DNA fragments are annealed and ligated in just one single reaction. The specific annealing of the fragments was guided through complementary overhangs (10-18 nucleotides long) contained at both ends of each fragment. Overhangs were introduced with the primers used during PCR as described in 6.1.9 . For optimal cloning efficiency, the overhangs were designed using Oligo 7.58 software (Molecular Biology Insights Inc., USA) in order to avoid the formation of secondary structures in the primers and achieve an annealing temperature of 48°C. After PCR, the samples were treated with 1µl of DpnI for 1h at 37°C to remove the DNA template. Amplified fragments were purified via MSB Spin PCRapace Kit (Stratec, Germany). For a standard Gibson assembly reaction, equimolar (30 fmol) amounts of each DNA fragment were mixed with ddH$_2$O to a final volume of 2.5 µl. Next, 2.5 µl of a 2X Gibson assembly protein mix (prepared in the lab) were added to the previously mixed DNA fragments. A reaction without the addition of one fragment was included as a negative control. Samples were incubated at 46°C for 30 min. After the reaction, electro-competent cells were transformed with 1 µl of the Gibson assembly reaction and resuspended with 1 ml of 2YT medium supplemented with 0.5% (w/v) glucose. Cells were then inoculated on selective agar plates overnight at 37°C to select for cells containing the DNA construct of interest.

6.1.11 Blunt end mutagenesis PCR

If not mentioned otherwise, mutagenesis of a specific DNA plasmid was perfumed by PCR and followed by blunt end ligation (Sambrook et al. 1989). PCR primers were designed so they introduced the desired mutation to the target DNA vector. Mutations were introduced by amplifying the whole vector using PCR. Oligonucleotides were synthesized with a 5' phosphoryl group to be able to perform blunt end ligation by the T4 ligase as described in 6.1.8 . PCR reaction was carried out just as described in 6.1.9 . The amplified plasmid was on-column purified using the MSB spin PCRapace kit (Stratatec, Germany). The purified DNA product was measured as described in 6.1.3 and further treated with DpnI for 2h at 37°C. 250 ng of amplified product was ligated overnight at 37°C. After incubation, the

reaction was diluted to a final DNA concentration of 10ng/µl. Last, 1µl of the diluted reaction was transformed into 50 µl electro-competent cells as described in 6.1.2 .

6.1.12 Protein test expression

Protein expression in small-scale was used as a quick test to screen for bacterial colonies containing the plasmid of interest after Gibson assembly reaction or restriction enzyme based cloning. Colonies with the desired plasmid would express a given protein of known molecular weight. To do this rapid test, individual colonies were picked from the plate and resuspended into 200 µl of 2YT medium containing appropriated antibiotic(s). Resuspended cells were grown for 3-4h at 37°C with continuous agitation. Next, 100µl of the bacterial culture were mixed with 100µl of fresh 2YT supplemented with antibiotic(s) and 0.4 mM IPTG in order to induce protein expression. Bacterial cells were incubated again for 2h at 37°C while shaking. After protein expression, cells were centrifuged at 13,000 rpm for 3 min using a tabletop centrifuge. Bacterial pellets were resuspended in 100µl of SDS sample buffer and analyzed by SDS-PAGE as described in 6.3.9 . Positive colonies were inoculated using the remaining non-induce culture for DNA extraction (6.1.3) or for protein expression in large scale (6.3.1).

6.1.13 PCR colony

The insertion of the DNA sequence of interest into a given vector after Gibson assembly reaction or restriction enzyme based cloning was confirmed by colony PCR. Colony PCR consisted in using a resuspended bacterial colony as source of DNA template for PCR. A specific DNA product is amplified by PCR only in positive samples. For colony PCR, primers were designed so they aligned up- and downstream of the inserted DNA fragment. Before starting the PCR reaction, cells were picked from the agar plate and resuspended in 30µl of ddH2O. Then, PCR reactions were performed in a total volume of 20 µl using 1 µl of the bacterial suspension as DNA template. After PCR, the reaction was mixed with 1/10 volume Orange G sample buffer. Samples were finally analyzed by agarose gel electrophoresis. Positive colonies were inoculated using the resuspended bacteria for further DNA extraction as in 6.1.3 .

6.1.14 Oligonucleotide synthesis and DNA sequencing

Oligonucleotide design was performed using the Oligo 7.58 software (DBA Oligo, Inc., USA). Designed sequences were then synthesized by Sigma-Aldrich (Germany) and dissolved in ddH$_2$O to a final concentration of 100 µM. If not stated otherwise, oligonucleotides were purchased as simple desalted material without any extra purification step.

For the design of larger DNA fragments such as entire codon-optimized genes, Gene Designer 2.0 software (DNA2.0, USA) was used. Further DNA synthesis was carried out by GenScript (USA).

Cloned plasmid and specific PCR products were sequenced using the appropriated primers at Seqlab (Germany). Sequencing results were analyzed using the SeqMan software (DNAStar, USA).

6.2   Methods for Saccharomyces cerevisiae genetics
6.2.1   Long-term storage of *S. cerevisiae* cells

Glycerol stocks were prepared for the long-term storage of all yeast strains used in this study (Table 5.7). In order to prepare the stocks, 1 ml of over-grown yeast culture was mixed with 800 µl of sterile glycerol (80% v/v). The mixture was then frozen in liquid nitrogen and incubated at -80°C until further use.

6.2.2   DNA transformation of *S. cerevisiae* cells

Transformation of *S. cerevisiae* cells was performed as described in (Gietz & Schiestl 2007; Schiestl & Gietz 1989). Normally, the strain to be transformed was storage in a glycerol stock at -80°C as indicated in 6.2.1 . A small piece of the frozen glycerol stock was used to inoculate a 10 ml pre-culture. Then, the culture was incubated at 30°C overnight or until stationary growth phase was reached (OD$_{600}$≈3.0). On next day, the over-grown pre-culture was used to inoculate 25 ml of fresh medium to an OD$_{600}$ of 0.1. The culture was then inoculated at 30°C with continuous shaking until the exponential phase growth was reached (OD$_{600}$≈0.8). Cells were pelleted by centrifugation for 5 min at 2,000 rpm and 4°C (swing-out rotor Multifuge 3L-R, Thermo Scientific). The supernatant

was discarded and cells were resuspended in 20 ml of ddH$_2$O. Resuspended cells were centrifuged again and resuspended once more in 20 ml of ddH$_2$O. Cells were then resuspended in 10 ml of LiAc/1x TE buffer and centrifuge for 5 min at 2,000 rpm and 4°C. Cells were finally resuspended in 500 µl of LiAc/1x TE buffer. Typically, 100 µl of resuspended cells were mixed with 1 µg of the plasmid to be transformed, 80 µg of heat-denatured salmon sperm DNA and 500 µl of PEG/LiAc/1x TE buffer. The mixture was incubated for at least 1-3 h at 30°C while shaking. Cells were incubated for 20 min at 42°C. Subsequently, cells were pelleted by centrifugation for 5 min at 2,000 rpm and 25°C using a tabletop centrifuge (Eppendorf, Germany). Cells were resuspended in 300 µl of sterile ddH$_2$O and plated on selective agar. Plates were incubated overnight at 30°C. YPD medium was used for selection based on antibiotic resistance markers and SD-medium for auxotrophic markers. Finally, glycerol stocks of the transformed colonies were created as described in 6.2.1 .

6.2.3   Gene deletion and genomic tagging in *S. cerevisiae* cells

Yeast cells were transformed with linear double stranded DNA fragments to allow homologous recombination as described in (Baudin et al. 1993). This transformation method was used to delete a gene of interest or to introduce a C-terminal tag into a specific gene. For the deletion of a specific gene, an expression cassette that confers resistant to geneticin was introduced to replace the gene of interest (Janke et al. 2004). The replacement was using homologous sequences to the 5'-UTR and 3'-UTR of the targeted gene. In the case of C-terminal tagging of a gene, the yeast codon-optimized GFP tag (yeGFP) was introduced using homologous sequences to C-terminal part and to the 3'-UTR of the targeted gene. In order to select for the integration of yeGFP, hygromycin B phosphotransferase was used as a selection marker (Janke et al. 2004). Hygromycin B phosphatase provides resistance to hygromycin B. PCR was used to amplify the linear DNA used for transformation of yeast cells (6.1.9). The homologous sequences to direct gene deletion or genomic tagging were introduced with the primers used during PCR. The amplified PCR product was purified via MSB Spin PCRapace Kit (Stratec, Germany). Transformation of cells with linear DNA was performed as described in 6.2.2 , except that 2 µg of linear DNA were instead of a circular plasmid. Transformed cells were plated on YDP medium supplemented with geneticin (300 µg/ml) and/or hygromycin B (100 µg/ml).

Plates were incubated for 48h at 30°C for the isolation of transformed colonies. PCR was used with in order to verify the deletion of a gene or the insertion of yeGFP. Finally, glycerol stocks of the selected colonies were created as described in 6.2.1 and storage at -80°C for further use.

6.2.4   Extraction of genomic DNA from *S. cerevisiae* cells

Genomic DNA from yeast cells was prepared using repetitive freeze and thaw cycles as described in (Harju et al. 2004). For the extraction, a yeast colony was inoculated in 2 ml of selective liquid YPD or SD-medium overnight at 30°C with continuous shaking at 90 rpm. 1.5 ml of the saturated yeast culture was centrifuged for 5 min at 2,000 rpm and 25°C using a tabletop centrifuge (Eppendorf, Germany). The cell pellet was resuspended in 200 µl of lysis buffer (2% Triton X-100 (v/v), 1% SDS (w/v), 100 mM NaCl, 10 mM Tris/HCl pH 8.0, 1 mM EDTA, pH 8.0). Resuspended cells were frozen in liquid nitrogen for 2 min and then incubated at 95°C for 1min to thaw the sample quickly. This freeze and thaw cycle was repeated two more times. Cells were then vortexed thoroughly for 1 min. Subsequently, 200 µl of chloroform (100% v/v) were added and then the sample was mixed by vortexing for 2 min. Samples were then centrifuged for 3 min at 14,000 rpm and 25°C. After centrifugation, the aqueous phase was placed into a new 1.5 ml tube containing 400 µl of ice-cold ethanol (100% v/v). Samples were incubated at -20°C for 1-3 h to let the DNA precipitate. DNA was pelleted by centrifugation for 10 min at 14,000 rpm and 25°C. The supernatant was removed carefully by aspiration and the pellets were washed with 500 µl of ethanol (70% v/v). Samples were centrifuged once more for 5 min at 14,000 rpm and 25°C. The pellet was air-dried and resuspended in 20 µl of ddH$_2$O.

6.2.5   Confocal microscopy of living *S. cerevisiae* cells

Confocal fluorescent laser scanning microscopy was used to study the intracellular localization of GFP-tagged proteins. The localization of fluorescent fusion proteins was analyzed in wild type yeast and in mutant cells lacking a specific gene. A yeast codon-optimized GFP (yeGFP) tagged was introduced by homologues recombination as described in 6.2.3 .

Cells expressing the yeGFP-tagged proteins were inoculated in 10 ml YPD or SD-medium supplemented with antibiotic(s). The culture was incubated overnight at 30°C with permanent shaking at 90 rpm in order to obtain a confluent culture. On the next day, the pre-culture was used to inoculate 20 ml of fresh YPD or SD-medium with an initial concentration of $OD_{600}\approx0.1$. The yeast culture was then incubated at 30°C and 90 rpm until exponential growth phase was reached ($OD_{600}\approx0.8-1.0$). An aliquot of 2 ml of culture was centrifuged for 5 min at 3,000 rpm and 25°C using a tabletop centrifuge (Eppendorf, Germany). The supernatant was removed carefully by aspiration and the cell pellet was washed with 1 ml of SD-medium –URA. Cells were centrifuged again with the same conditions as before and the supernatant was also discarded. Centrifuged cells were resuspended with 100 µl of fresh SD-medium –URA. To image the cells, 7 µl of the resuspended cells were directly loaded onto a Superfrost UltraPlus microscope slides (Menzel GmbH, Braunschweig, Germany). Confocal images were acquired with a Leica SP5 confocal laser scanning microscope using a 63x immersion objective (Leica, Germany).

6.3   Protein biochemistry methods

6.3.1   Over-expression of recombinant proteins in *E. coli*

E. coli strains detailed in table 5.6 were used for protein expression in large-scale. First, cells of the preferred *E. coli* strain were transformed by electroporation (6.1.2) with the plasmid encoding the protein of interest. Transformed cells were plated on a selective agar and incubated overnight at 37°C. After incubation, a single colony was picked and inoculated in of 50-100 ml 2YT medium supplemented with the appropriate antibiotic(s). This pre-culture was incubated overnight at 28°C with constant shaking. On the next day, pre-cultures were diluted five times using fresh 2YT or TB medium containing the respective antibiotic(s) and then incubated at 37°C for 1h before starting protein over-expression. Recombinant protein expression was initiated after the addition of 0.05-0.1 mM IPTG to the culture. Cells were then incubated at 18°C or 25°C with continuous agitation at 90 rpm and grown either for 6h or overnight depending on the expression level of the protein of interest. After incubation, 1 of mM PMSF and 10 mM of EDTA were added to prevent protein hydrolysis by unspecific bacterial proteases. Then, cells were harvested by centrifugation for 10 min at 5,000 rpm (RC6 plus centrifuge, F9 rotor;

Sorvall, USA). Supernatant was removed carefully by decantation and cells were resuspended using resuspension buffer (50 mM Tris/HCl pH 7.5, 150-500 mM NaCl, 2-20 mM imidazole, 0-5 mM DTT). Normally, cells were resuspended to a final concentration of OD_{600} of 50-100/ml. Resuspended cells were frozen in liquid nitrogen and stored at -80°C until further used. Samples of 50-200 µl were taken before and after recombinant protein expression in order to be analyzed by SDS-PAGE as described in 6.3.9 .

6.3.2 Protein expression in *S. cerevisiae*

The plasmid(s) encoding the protein(s) of interest were transformed using *S. cerevisiae* strain SFY123 (Table 5.7). Cells were transformed using PEG/LiAC solution as described in 6.2.2 . After selection of transformed cells on selective agar plates, a single colony was picked to inoculate a 50 ml of SD-medium 2% (w/v) of glucose. The pre-culture was incubated overnight at 30°C with constant shaking at 90 rpm. Then, cells were pelleted by centrifugation for 5 min at 2,000 rpm and 25°C and. After centrifugation, the supernatant was discarded and cells were resuspended in fresh medium supplemented with 2% (w/v) of glucose and 2% (w/v) of raffinose. Glucose and raffinose were added from solution stocks at 40% (w/v) and 20% (w/v), respectively. Centrifugation and subsequent resuspension of cells with fresh medium supplemented with 2% (w/v) of glucose and 2% (w/v) of raffinose were repeated two more times. Next, resuspended cells were used to inoculate 100-250 ml of YPD or SD-medium supplemented with antibiotics, glucose and raffinose to a initial concentration of $OD_{600} \approx 0.2$. The culture was incubated at 30°C with constant shaking at 90 rpm until exponential growth phase was reached ($OD_{600} \approx 0.8$-1.0). In order to start protein over-expression, 2% (w/v) of galactose was added to the media. Galactose was added from a sterile solution stock at 20% (w/v). Protein expression was carried for 6-8h at 30°C while shaking the culture at 90 rpm. Cells were then pelleted by centrifugation for 10 min at 2,000 rpm and 25°C (RC6 plus centrifuge, F10 rotor; Sorvall, USA). The yeast pellet was resuspended in resuspension buffer (50 mM Tris/HCl pH 7.5, 150-500 mM NaCl, 2-20 mM imidazole, 0-5 mM DTT) to a final OD_{600} of 20-50/ml. A cocktail of different protease inhibitors was added to the resuspended cells to a final concentration of 1x. The stock (500x) of protease inhibitors contained the following compounds: 5 mg/ml aprotinin, 5 mg/ml leupeptin, 2.5 mg/ml elastatinal, 2.5 mg/ml chymostatin and 0.5 mg/ml pepstatin

A. Cells were frozen in liquid nitrogen and storage at -80°C until further processing. Aliquots of 50-200 µl were taken before and after protein expression to analyze the expression of the protein of interest by SDS-PAGE as described in 6.3.9 .

6.3.3   Sonication of *E. coli* cells

After expression of the protein of interest, cells were resuspended and frozen as described in 6.3.1 . The frozen pellet was rapidly thawed by incubation of the sample in a hot water bath for 10-15 min. After thawing, samples were sonicated in an ice-water bath 4 times for 1 min with 30 seconds of rest in between each sonication cycle. For this, a maximum output sonication power and 45% duty cycle were used (Sonifier 450, Branson, UK). In order to remove the bacterial debris and any insoluble material from the lysate, the sample was cleared by ultracentrifugation at 38,000 rpm and 4°C for 1.5 h (WX Ultracentrifuge, T647.5 rotor, Sorvall). For long-term storage of the cleared lysate at -80°C, sucrose at final concentration of 250 mM was added. Alternatively, the lysate was directly used after ultracentrifugation for protein purification by affinity chromatography (6.3.5).

6.3.4   Disruption of Saccharomyces cerevisiae membrane

As described in 6.3.2 , cells were storage at -80°C after the expression of the protein of interest. Afterwards, frozen cells were frozen in a hot water bath for 10-15 min. Glass beads (diameter of 500 µm) were mixed with the thaw sample in a 1:1 (v/v) ratio. Before use, glass beads were soaked in concentrated HCl for 16h, rinsed thoroughly with ddH$_2$O and finally dried for 20 h at 185°C. In order to inhibit serine proteases, 10 µl of 0.1 mM PMSF were added to the samples. Afterwards, cells were vortexed 5 times for 1 min using the highest power setting of the vortex machine. One minute of resting time was kept in between each vortexing cycle. Yeast lysate was separated from the glass beads by simple decantation. Cell debris and insoluble material was removed from the lysed cells by ultracentrifugation at 38,000 rpm and 4°C for 1.5 h (WX Ultracentrifuge, T-1250 rotor, Sorvall). The cleared lysate was supplemented with 250 mM sucrose and stored at -80°C until further processing.

6.3.5 Protein purification using immobilized metal ion affinity chromatography

If not stated otherwise, purification of recombinant proteins from yeast or bacterial lysates was carried out by Ni^{2+} chelate affinity chromatography. Nickel ions were immobilized on silica beads (prepared by D. Görlich) to isolate recombinant proteins from cellular lysates via a poly histidine tag (His-tag). All proteins expressed in this study contained a N- or C-terminal poly histidine tag (His_6-, His_{10}- or His_{14}-tag). Two different sizes of Ni^{2+} chelate silica beads with different substitution grade were used (6% substituted, 1000 Å and 20% substituted 500Å). The latter was used for the purification of monomeric proteins. In contrast, silica beads 6% substituted and 1000 Å in size were used for the isolation of large protein complexes.

The bacterial and yeast lysates used for protein purification were prepared as indicated in 6.3.1 and 6.3.2 , respectively. The volumen of beads used for the purification of a given protein depended on the amount of soluble protein found in the cleared lysate. An appropriate volume of nickel beads was placed in gravity flow column (Luer lock, Sigma, USA) and further equilibrated with 3 column volumes (CV) of resuspension buffer (45 mM Tris/HCl pH (7.5), 150-500 mM NaCl, 2-20 mM imidazole, 5% (v/v) glycerol, 4.5 mM $MgCl_2$, 0-5 mM DTT). Equilibrated beads were incubated with the lysate for 1 h at 4°C with constant mixing. After incubation, the beads were placed back on the gravity flow column. The flow-through was collected to check for the amount of unbound His-tagged protein in the lysate. Beads were then washed with 1 CV of resuspension buffer and 3 CV of washing buffer (45 mM Tris/HCl pH (7.5), 500-2000 mM NaCl, 4.5 mM $MgCl_2$, 2-20 mM imidazole, 0-5 mM DTT). Whenever needed, the washing buffer wash supplemented with either ATP (5 mM, pH 7.5) to remove bacterial chaperones bound to the recombinant protein or with a high concentration of NaCl (1-2 M) to remove bound nucleic acids. After washing the beads, the protein was eluted by competitive elution using a high concentration of imidazole. Elution of the protein was carried out in 0.5 or 1.0 ml fractions of elution buffer (50 mM Tris/HCl pH 7.5, 150-500 mM NaCl, 400 mM imidazole, 250 mM sucrose, 0-5 mM DTT) using a total volume of 3-4 CV. 1 μl of each collected fraction was soaked on a nitrocellulose membrane. The membrane was then stained with Amido Black Quick Staining solution to check the protein content in each fraction. Fractions with the highest protein content were pooled in a single sample. Protein

absorption at 280 nm was measured using the NanoDrop 2000C spectrophotometer (PeqLab, Germany) according to the manufacturer's instructions. Protein concentration in final sample was calculated using the Beer-Lambert equation (Swinehart 1962). The protein-specific extinction coefficient at 280 nm (ε_{280}) was obtained using the protean software (DNAStar, USA). For long-term storage of the protein, the sample was frozen in liquid nitrogen and placed in a -80°C freezer.

In order to analyze the expression level and the purification efficiency of the protein, protein samples were taken as follows: non-induced cells, induced cells, soluble protein fraction (cleared lysate), insoluble fraction (cellular debris after ultracentrifugation), flow-through and final purified protein. All samples were analyzed by SDS-PAGE and stained with coomassie brilliant blue as indicated in 6.3.9 . For normalization purposes, all samples analysis by SDS-PAGE corresponded to 35 mOD of cells or to 1/1000 of the total purified protein were for analysis.

6.3.6   On-column cleavage protein purification using specific proteases

Proteins expressed in this study often included a protease cleavage site in between the His-tag and the protein of interest. If needed, the His-tag was removed from the recombinant protein by using specific proteases. Different proteases cleavage sites were used during this study such as bdSUMO, SUMOvera, SUMOstar, scSUMO, bdNEDD8, TEV site and many other (0). Removal of the His-tag was performed while the protein was bound to the Ni^{2+}chealate beads (on-column protein cleavage). Therefore, the protein was eluted from the beads after adding resuspension buffer supplemented with the respective protease. Working concentrations of the proteases used in this study were as described in (Frey & Görlich 2014a). On-column protein elution was performed by placing inside the column, 2 CV of resuspension buffer containing the respective protease at right concentration. The buffer was forced to pass carefully through the beads inside column with the help of a syringe. The column was then incubated for 1h at 4°C in order to achieve complete protein cleavage. After incubation, protein was retrieve from the column in 0.5 ml fractions of resuspension buffer containing 250 mM sucrose for a total volume of 1.5 CV. Determination of protein concentration was performed as indicated in 6.3.5 . Finally, the total protein sample was frozen in liquid nitrogen and stored at -80°C for further use.

6.3.7   Purification of binary protein complex in *Saccharomyces cerevisiae*

Purification of binary protein complexes was performed as described in (Frey & Görlich 2014b), except that the recombinant proteins were expressed in *S. cerevisiae* cells. Protein expression was carried out using *S. cerevisiae* strain SFY123 (S288C, Matα, H2B-CFP::TRP1, his3Δ200, leu2Δ0, lys2Δ0, met15Δ0, ura3Δ0) as described in 6.3.2 . Specifically, the binary complex of Citrine•Enhancer was purified in this study (Kirchhofer et al. 2010) was. SFY123 cells were transformed with two plasmids as described in 6.2.2 . One plasmid codified for fusion protein H14-SUMOvera-Enhancer and the second for ZZ-SUMOstar-Citrine. Cells were grown in SD-medium -URA supplemented with 300 μg/ml of hygromycin B to express the recombinant complex as indicated in 6.3.2 . The cleared yeast lysate was incubated with Ni^{2+} chelate beads for 1h at 4°C. Contaminant proteins were removed as indicated in 6.3.5 . Subsequently, the protein complex was eluted after incubation of elution buffer containing 200 nM of SUMOvera protease for 1h at 4°C as in 6.3.6 . For the second purification step, the complex was immobilized via the ZZ-tag (tandem repeat of the Z domain from staphylococcal protein A) using silica beads coupled to an anti Z-domain affibody. The sample was incubated with 2 ml of anti Z-domain beads for 1h at 4°C. After incubation, beads were washed with 3 CV of washing buffer (45 mM Tris/HCl pH 7.5, 500 mM NaCl, 4.5 mM $MgCl_2$, 5 mM DTT). For protein elution, silica beads were incubated with resuspension buffer supplemented with 100 nm of SUMOstar protease for 1 h at 4°C. After elution, samples were pooled and frozen in liquid nitrogen for storage at -80°. Protein samples were taken at different steps of the expression and purification of the complex and analyzed by SDS-PAGE as indicated in 6.3.9 .

6.3.8   Size exclusion chromatography

After protein purification as described in 6.3.5 , the high amount imidazole in the protein preparation was removed by small-scale size exclusion chromatography. For this, Nap5 or PD-10 desalting columns (GE Healthcare, USA) were used according to the manufacturer's instructions. If not stated otherwise, binding buffer (45 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM $MgCl_2$, 5 mM DTT) was used to replace the elution buffer (50 mM Tris/HCl pH 7.5, 150-500 mM NaCl, 400 mM imidazole, 250 mM sucrose, 0-5 mM DTT). After buffer exchange, the protein concentration was measured using the

NanoDrop 2000C spectrophotometer (PeqLab, Germany) according to the manufacturer's instructions. Samples were supplemented with 250 mM sucrose and stored at -80°C until further processing.

Size exclusion chromatography was also used to further purify His-tagged proteins from nucleic acid or protein contaminants. For this purpose, Superdex 200 26/60 column (protein separation range 10-600 kDa) or Superdex 75 16/60 column (protein separation range 3-70 kDa) were used following supplier's instructions. Gel filtration columns were connected to a ÄKTA Purifier or ÄKTA Explorer system (Pharmacia, Sweden) in order to perform and control the purification process. After buffer exchange as described above, the His-tagged protein sample was ultracentrifuged for 30 min at 38,000 rpm and 4°C (Discovery M120 SE ultracentrifuge, Sorvall, S45-A rotor) in order to pelleted any precipitated protein. Before loading the sample, the column was equilibrated with 1.5 CV of binding buffer (45 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM $MgCl_2$, 5 mM DTT). The protein was eluted from the column using 1.2 CV of the same buffer used during equilibration. Fractions of 0.5-1.0 ml were collected as soon as protein leaving the column was detected at 280 nm. All fractions in the elution peak were analyzed by SDS-PAGE as indicated in 6.3.9 . Pure protein samples were pooled and supplemented with 250 mM sucrose for long-term storage at -80°C.

6.3.9   SDS polyacrylamide gel electrophoresis (SDS-PAGE)

The analysis of protein samples was performed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). During this study, SDS-PAGE was based in the protocol described in (Laemmli 1970; Sambrook et al. 1989). Gradient SDS polyacrylamide gels (8-14% w/v) were prepared by Gabriele Hawlitschek and Jürgen Schünemann. All samples corresponding to 35 mOD of cells or 1-2 µg of purified protein were prepared using 5x SDS sample buffer (3 % SDS, 125 mM Tris/HCl (pH 6.8), 50 mM DTT, 1 M sucrose and coomassie brilliant blue G250). Samples were heated for 5 min at 98°C before loading. The electrophoresis was performed using the following conditions: 400 V, 50 mA for 70 min. After the run, the polyacrylamide gels were stained using coomassie staining solution (0.3% (w/v) coomassie brilliant blue G250 in 3% (v/v) acetic acid). Then, the staining solution was discarded and the gel was incubated with dH2O until

background staining disappeared. Finally, gels were scanned using the Epson Perfection V700 Photo scanner.

6.3.10 Protein cleavage assays

Cleavage reactions were carried out using cleavage buffer (45 mM Tris/HCl pH 7.5, 250 mM NaCl, 2 mM $MgCl_2$, 250 mM sucrose, 10 mM DTT) in a total volume of 20 μl. If not stated otherwise, 100 μM of tagged substrate were incubated with various amounts of a given protease for 1-2h at 4°C or 25°C. Prior to the reaction, substrates and proteases were diluted with cleavage buffer to 2-fold of the concentration required for the reaction. Equal volumes of diluted substrate and proteases were mixed in order to start the reaction. Cleavage reaction was then stopped by adding 180 μl of SDS sample buffer (3 % SDS, 125 mM Tris/HCl (pH 6.8), 50 mM DTT, 1 M sucrose and coomassie brilliant blue G250). Samples corresponding to 2.5 μg of substrate were resolved by SDS-PAGE as described in 6.3.9 . Control samples without protease were included in the assays.

6.3.11 On-column protein cleavage assays

On-column protein cleavage assays were performed in mini-spin columns were used (MobiTec, Germany). His-tagged substrates (12 μM) were immobilized onto 50 μl of Ni^{2+} chelate silica beads for 1h at 4°C. After immobilization of the substrate, beads were washed with 500 μl of cleavage buffer (45 mM Tris/HCl pH 7.5, 250 mM NaCl, 2 mM $MgCl_2$, 250 mM sucrose, 10 mM DTT). The buffer was passed through the beads by mild centrifugation of the column for 1 min at 1,000 rpm and 4°C. Washing of the beads was carried out three times as just described. For elution of the protein, 50 μl cleavage buffer supplemented with a certain amount of a given protease were incubated with the pre-loaded beads for 1h at 4°C. The eluted protein was collected by centrifugation of the column for 1 min at 1,000 rpm and 4°C. Next, another 50 μl of cleavage buffer were added to the beads and collected immediately by centrifugation. Elution and washing fractions were pooled before measuring the concentration of eGFP or mCherry at 488 nm and 585 nm, respectively. Also, eluted samples were imaged while illuminated at 366 nm. Buffer sample controls lacking the protease were included in the assays.

6.3.12 Western blot

Western blot was used in this study to analyze the stability of different SUMO-tagged proteins in various eukaryotic cellular extracts. All analyzed samples corresponded to 35 mOD of cells expressing a SUMO fusion protein or to 1 μg of the SUMO-tagged protein. First, protein samples were treated and analyzed by SDS-PAGE as described in 6.3.9 . After SDS-PAGE, the SDS polyacrylamide gel, whatman paper (3.0 mm) and nitrocellulose membranes were equilibrated for 10 min in 1x blotting buffer (100 ml of 10X SDS-PAGE running buffer, 200 ml MetOH, 0.03% w/v SDS). After equilibration, the gel was placed in direct contact with the nitrocellulose membrane followed by a whatman paper. On top of the other side of the gel, another pre-equilibrated piece of whatman paper was placed. The assembled gel was then pressed carefully in order to remove any air bubble formed in between the layers of paper. Proteins were then transfer from the gel to the nitrocellulose membrane using an electrophoretic chamber filled with 1x blotting buffer. Transfer was performed overnight at 4°C and 100 mA with gentile constant shaking of the blotting buffer. Next, the nitrocellulose membrane was removed from the chamber and incubated with blocking buffer (4 % w/v of powder milk in 1x PBS buffer) for 1h at 25°C. The membrane was incubated again fresh blocking buffer supplemented with 0.1% (v/v) of tween-20 and the primary antibody in the appropriated dilution (1:5000) either for 2h at 25°C or overnight at at 4°C. After incubation, the blocking buffer was discarded and the membrane was wash with 1x PBS buffer for 10 min. Washing of the membrane was repeated three times using each time fresh 1x PBS buffer. Subsequently, the membrane was incubated in blocking buffer containing the appropriated secondary antibody (Goat α-rabbit IRdye; Goat α-mouse IRdye, Licor, USA) with the right dilution (1:10000) for 1h at 25°C. Membrane was rinsed 3 times using fresh 1x PBS buffer for 10 min each washing step. Later, the membrane was air-dried before scanning at 700 nm or 800 nm using the Odyssey scanner (Licor, USA) following manufacturer's instructions.

6.3.13 Stability of SUMO-tagged proteins expressed in *Saccharomyces cerevisiae*

Diverse SUMO-tagged proteins were expressed for 6h at 30°C in *S. cerevisiae* strain SFY123 as indicated in 6.3.2 . Proteins were expressed as citrine fusion proteins with a N-terminal ZZ tag (Heikal et al. 2000). In order to analyze the stability of the fusion proteins,

cell lysates were generated by TCA/NaOH extraction as described in (Zuk 1998). Cell lysates corresponding to 35 mOD of cells expressing the citrine fusion proteins were resolved by SDS-PAGE and analyzed by western blot as described in 6.3.12 . An anti-GFP primary antibody was used to detect citrine by western blot. A cell lysate lacking a SUMO-tagged citrine was used as negative control.

6.3.14 Stability of SUMO-tagged substrates in eukaryotic extracts

A eukaryotic cell lysate is a source of endogenous SUMO-specific proteases. Therefore, the stability of different SUMO-tagged MBP fusion proteins was tested in different eukaryotic extracts. 1 μM of SUMO-tagged substrate was incubated with 10 μl of the following extracts: wheat germ extract, *Xenopus laevis* egg extract, rabbit reticulocytes extract, HeLa cell extract, *Drosophila* S2 cell extract, LTE lexsy cell extract. The preparation of the lysates was performed as described in (Mureev et al. 2009; Kovtun et al. 2010; Blow & Laskey 1986; Crevel & Cotterill 1991; Endo et al. 2010; Jackson & Hunt 1983). The reactions were incubated for 2 h at 25°C in the presence and absence of a protease mix containing 1μM of different SUMO-specific proteases (Ulp1, SUMOstar, bdSENP1 and SUMOvera protease) in a total reaction volume of 12.5 μl. The protease mix was used to test for the presence of SUMO-specific protease inhibitory substances present in the lysates. The stability of the substrates was analyzed by western blot as described in 6.3.12 using an anti-MBP primary antibody.

6.3.15 Small-scale binding assays

Recombinant purified proteins were used in this study to analyze protein-protein interactions. In general, binding reactions were performed using a ED-tagged bait protein (1-2 μM) and one or two different prey proteins (3-6 μM). Proteins were mixed in binding buffer (45 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM MgCl$_2$, 5 mM DTT) and incubated for 1-2 h at 4°C to let proteins to interact. Binding reactions were carried out in a total volume reaction of 200 μl. Next, binding reactions were incubated for 1 h at 4°C in mini-spin columns (MobiTec, Germany) containing 25 μl of anti Z-domain affibody beads previously equilibrated with 2x 500 μl of binding buffer. The use of anti Z-domain affibody silica beads allowed the immobilization of ED-tagged bait (Lindborg et al. 2013).

After incubation of the proteins with the beads, the flow-through was collected by gentle centrifugation of the column for 1 min at 1,000 rpm and 4°C. Beads were then washed with 500 µl of binding buffer. The buffer was forced to pass through the beads by gently centrifugation of the column for 1 min at 1,000 rpm and 4°C. Washing of the beads was repeated one more time using again 500 µl of binding buffer. Elution of the interacting proteins was performed by on-column protein cleavage. To do so, 50 µl of binding buffer supplemented with bdSENP1 (50 nM), SUMOvera protease (200 nM) or bdNEDP1 (500nM) were added to the beads and incubated for 1 h at 4°C with constant shaking. The elute was collected by centrifugation of the column for 1 min at 1,000 rpm and 4°C. Beads were washed with another 50 µl of binding buffer and collected again by centrifugation of the column for 1 min at 1,000 rpm and 4°C. The eluate and the washing fractions were combined in a single tube regarded as the protein bound fraction. For this study, protein-protein interaction was performed in the presence and absence of untagged Gsp1Q71L-GTP (3-6 µM). In addition, a sample control lacking the bait protein was included to account unspecific binding of the prey to the anti Z-domain beads. Samples from the input protein fraction, flow-through and the bound fraction corresponding to 1-2 µg of the bait protein were analyzed by SDS-PAGE and stained with coomassie brilliant blue as described in 6.3.9

6.3.16 Nuclear transport receptors binding specificity assays

The interaction between a transport cargo and different nuclear transport receptor (NTRs) from *S. cerevisiae* was analyzed in small-scale binding assays. For the binding reaction, a ED-bdNEDD8-tagged cargo was used as bait and different *E. coli* lysates containing an over-expressed eGFP-tagged NTR were used as preys. The cleared *E. coli* lysates containing the NTR were generated using binding buffer (45 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM MgCl$_2$, 5 mM DTT) as described in 6.3.3 . The concentration of the eGFP-tagged NTRs in the lysates was determined at 488 nm in order to use equal amounts of NTR in all samples. In general, 2 µM of the cargo were incubated in binding buffer with 6 µM of the eGFP-tagged NTR for 1 h at 4°C. Gsp1Q71L-GTP (6 µM) was included in the binding reaction only if the recognition of an export cargo by different NTRs was tested. The final volume of the binding reaction was 200 µl. Samples were then incubated for 1 h at 4°C in mini-spin columns (MobiTec, Germany) containing 25 µl of anti Z-domain

affibody beads previously equilibrated two times with 500 µl of binding buffer. The cargo•NTR complexes were immobilized to the Z-domain affibody matrix via the ED-bdNEDD8-tagged cargo. After immobilization, the flow-though was collected by centrifugation of the column for 1 min at 1,000 rpm and 4°C using a tabletop centrifuge (Eppendorf, Germany). Beads were washed two times with 500 µl of binding buffer in order to remove unbound prey from the beads. Protein elution was performed with 50 µl of binding buffer supplemented with bdNEDP1 (500 nM) for 1h at 4°C. Eluted protein was collected by mild centrifugation of the mini-spin column for 1 min at 1,000 rpm and 4°C. Next, silica beads were washed with another 50 µl of binding buffer in order to retrieve protein left inside the column. Both 50 µl fractions were pooled and regarded as the bound fraction. Last, samples from the input material, flow-through and bound fractions were analyzed by SDS-PAGE as described in 6.3.9 .

6.3.17 Tandem affinity purification

Tandem affinity purification (TAP) was used to identify specific transport cargos for different nuclear transport receptors from *S. cerevisiae*. Purified nuclear transport receptors (Pdr6 and Lph2) with a N-terminal ED-SUMOvera-H12 tag were used as baits and *S. cerevisiae* cellular extract was used as source of transport cargos. The cell extract was prepared in GK75 buffer (20 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 75 mM KCl, 5% glycerol, 0.01% NP40, 0.5 mM DTT) and kindly provided by R. Lührmann. First, 1 ml of the yeast lysate was incubated with 100 µl of low substituted Phenyl-Sheparose matrix (GE healthcare Life Sciences, USA) in order to deplete endogenous nuclear transport receptors. Pre-depleted extract was incubated in GK75 buffer supplemented with 0.5 µM of NTR in a final volume of 1.5 ml. TAP experiments were performed in the presence and absence of Gsp1Q71L-GTP (6 µM) to identify export and import cargos, respectively. Samples were centrifuged for 30 min at 65,000 rpm and 4°C (Discovery M120 SE ultracentrifuge, S100AT4 rotor; Sorvall, USA). Then, the supernatant was incubated for 1h at 4°C to allow the formation of specific protein complexes. Samples were then incubated for 1 h at 4°C with 25 µl of anti Z-domain affibody beads previously equilibrated with 2x 500 µl of GK75 buffer. Beads were pelleted carefully by centrifugation using a tabletop centrifuge for 2 min at 1,000 rpm and 4 °C. The supernatant was removed by aspiration and beads were placed into mini-spin columns (MobiTec, Germany). The flow-through was collected

by centrifugation of the column for 1 min at 4°C and 1,000 rpm. Anti Z-domain beads were then washed three times with 500 µl of GK75 buffer. Bound material was eluted by on-column protein cleavage for 1 h at 4°C with 120 µl of GK75 buffer supplemented with 20 mM imidazole and 250 mM of SUMOvera protease. Elute was collected by centrifugation for 1 min at 1,000 rpm and 4°C. Subsequently, another 120 µl of GK75 buffer supplemented with 20 mM imidazole were added to the beads to recover protein left inside the column. Both fractions were pooled and incubated for 1h at 4°C with 250 µl of Ni^{2+} chelate beads previously equilibrated with buffer GK75. After incubation, beads were centrifuged for 1 min at 1,000 rpm and 4 °C in order to collect the flow-through. Beads were washed twice with 500 µl of GK75 buffer supplement with 20 mM imidazole. For the specific elution of isolated yeast proteins from the nickel beads, 200 µl of Gdn-HCl elution buffer (3M guanidinium chloride, 50 mM Tris/HCl pH 8.0) were incubated with the beads for 5 min at 25°C. The eluate was collected by centrifugation of the samples for 1 min at 1,000 rpm and 4°C. Subsequently, another 200 µl of Gdn-HCl elution buffer were added to the column. Both fractions were pooled and regarded as the protein bound fraction (transport cargos). Guanidinium hydrochloride-containing samples were analyzed by SDS-PAGE for further protein identification by mass spectrometry as described in 6.3.19 . Nickel beads were then washed with 500 µl of GK75 buffer to remove the Gdn-HCl elution buffer left in the column. Last, His-tagged NTR and Gsp1Q17L were eluted from the nickel beads by adding 200 µl of SDS sample buffer supplemented with 400 mM of imidazole. Protein samples at different steps of the TAP assay were taken as follows: (1) Input fraction for the yeast lysate, (2) NTR and (3) Gsp1Q71L-GTP. (4) Elution fraction after on column protein cleavage. (5) Sample after Gdn-HCl buffer elution. (6) Samples after imidazole elution. A sample without NTR was included to analyze the background binding of yeast proteins in the lysate to the affinity matrixes used.

6.3.18 Protein precipitation

Guanidinium hydrochloride was removed from the eluted protein samples obtained during tandem affinity purification in order to analyze them by SDS-PAGE. To do so, selective protein precipitation from guanidinium hydrochloride-containing samples was performed. Typically, 200 µl of the sample were mixed with 1.8 ml of ice-cold propanol (100%, v/v) and incubated overnight at -20°C. Precipitated protein was pelleted by centrifugation for

10 min at 14,000 rpm and 25°C. Supernatant was carefully removed by aspiration without disturbing the small white pellet. Next, 1 ml of ice-cold 70% (v/v) isopropanol was added to wash the pellet. The sample was vortexed briefly and further centrifuged for 2 min at 14,000 rpm and 25°. The washing step was repeated using the same amount of ice-cold 70% (v/v) isopropanol. After centrifugation, the pellet was air-dried for 5 min to 1 h at 25°C. Finally, the dried pellet was dissolved in 35 μl of SDS sample buffer (3 % SDS, 125 mM Tris/HCl (pH 6.8), 50 mM DTT, 1 M sucrose and coomassie brilliant blue G250) in order to be analyzed by SDS-PAGE as described in 6.3.9 .

6.3.19 Mass spectrometry analysis

Protein identification in samples eluted with Gdn-HCl elution buffer was performed by liquid chromatography-mass spectrometry analysis (LS-MS). All experiments were carried out by Uwe Plessman. Prior LS-MS, guanidinium hydrochloride was removed from the samples as described in 6.3.18 . After dissolving the pellet in SDS sample buffer, samples were loaded into a 10% (w/v) polyacrylamide gel (NuPAGE, Life Technologies) and stained with colloidal coomassie blue staining solution. Specific protein bands were excised from the gel and washed with 30 μl of ddH$_2$O during 15 min at 25°C. Water was then removed and the gel piece was incubated with 20 μl of 50% (v/v) acetonitrile for 15 min at 25°C. The supernatant was discarded and 20 μl of 100% (v/v) acetonitrile for 15 min at 25°C. Next, acetonitrile was removed and the sample was incubated with 20 μl of 100 mM ammonium bicarbonate solution for 5 min at 25°C. Supernatant was removed and another 20 μl of 100% (v/v) acetonitrile were added. After removal of the supernatant, the gel piece was air-dried for 5 min to 1 h. The dried gel was then incubated with 50 μl of 100 mM ammonium bicarbonate solution supplemented with 10 μM DTT and further incubated at 56°C for 45 min. Then, the supernatant was removed and 30 μl of 55 mM iodoacetamide were added to the gel. The iodoacetamide solution was kept in the dark for 30 min at 37°C. Iodoacetamide was removed and 20 μl of 0.1% (w/v) RapiGest buffer (Waters, USA) complemented with 50mM ammonium bicarbonate solution were added to the sample. The resuspended gel was incubated at 37°C for 10 min. After incubation, RapiGest buffer was removed and trypsin (stock solution 0.1 mg/ml. Promega, USA) was added in a enzyme to protein ratio of 1:50. The sample was incubated overnight at 37°C for optimum enzymatic digestion. The peptide fragments were precipitated with

trifluoriacetic acid to a final concentration of 0.5% (v/v). Samples were incubated for 10 min at 37°C. The pellet was then collected by centrifugation of the sample for 10 min at 14,000 rpm at 25°C. The protein pellet was dissolved in a solvent system containing 2 % acetonitrile (v/v) and 0.05 % formic acid (v/v) to a final volume of 20 µl. For LS-MS analysis, a sample corresponding to 1µg of protein was injected into a nano liquid chromatographic system (UltiMate^{TM} 3000 RSLCnano system; Thermo Scientific, USA). Salt was removed from the sample using a trapping column for 3 min at a flow rate of 10 µl/min in 95 % of mobile phase A (0.1 % FA in H$_2$O, v/v) and 5 % of mobile phase B (80 % ACN and 0.05 % FA in H$_2$O, v/v). Elution of the samples from the trapping column was performed during 43 min with a linear gradient of 15-46 % mobile phase B at a flow rate of 300 nL/min. Eluted peptides were analyzed on a Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific, USA). Resulting data was analyzed using MaxQuant 1.5.5.1 and Perseus 1.5.2.6 software (Cox & Mann 2012; Tyanova et al. 2016).

6.4 Methods for the selection of bdSUMO and bdSENP1 mutants

6.4.1 *In vivo* selection system for proteases with orthogonal specifies.

SUMO-specific protease 1 *(*bdSENP1) and bdSUMO protein from *Brachypodium distachyon* were used as model protease/substrate pair to test the functionality of the *in vivo* selection system. Wild type bdSUMO was used to account for efficient protein cleavage by bdSNEP1. In contrast, a cleavage resistant mutant of bdSUMO, named here as bdSUMO*, was engineered to account for the absence of substrate recognition by bdSENP1. bdSUMO* has two point mutations (G96A; G97A) in positions P2 and P1 of bdSUMO. To test the performance of the selection system, four different plasmids were produced by placing bdSUMO or bdSUMO* in different combinations at each terminus of hygromycin B phosphotransferase. These plasmids are under the control of the *GAPDH* promoter and the *t1* terminator. To see more detailed information about the plasmid (Table 6.1).

Expression of bdSENP1 is under control of a *T5* IPTG-inducible promoter and a *lamda t0 trx* terminator. Additionally, the ribosomal binding site was modified from TAAAGAGGAGA to CAAAACAAGT to reduce bdSENP1 expression level as described in (Registry of Standard Biological Parts, www.partsregistry.org).

Table 6.1 Plasmids used the performance of the *in vivo* selection system in *E. coli*

Plasmid number	Protein codified	Promoter/terminator	Selection marker(s)
pAV0159	bdSUMO-degron^{NER}-Hph-bdSUMO-ssrA	*GAPDH* promoter; *lambda t1* terminator	Spectinomycin; Hygromycin B
pAV0160	bdSUMO*-degron^{NER}-Hph-bdSUMO-ssrA	*GAPDH* promoter; *lambda t1* terminator	Spectinomycin; Hygromycin B
pAV0161	bdSUMO-degron^{NER} Hph-bdSUMO*-ssrA	*GAPDH* promoter; *lambda t1* terminator	Spectinomycin; Hygromycin B
pAV0162	bdSUMO*-degron^{NER}-Hph-bdSUMO*-ssrA	*GAPDH* promoter; *lambda t1* terminator	Spectinomycin; Hygromycin B
pAV0076	bdSUMO-Hph-bdSUMO	*GAPDH* promoter; *lambda t1* terminator	Spectinomycin; Hygromycin B
pAV0031	bsSENP1	T5 IPTG inducible promoter	Kanamycin

E. coli NEB Express F' cells (*fhuA2 [lon] ompT gal sulA11 R(mcr-73::miniTn10--*Tet^S*)2 [dcm] R(zgb-210::Tn10--*Tet^S*) endA1 Δ(mcrC-mrr)114::IS10*) were used to perform the *in vivo* selection system for proteases with orthogonal substrate specificities. Electro-competent cells containing plasmid pAV0031 were transformed by electroporation with plasmids pAV0159, pAV0160, pAV0161, and pAV0162. Cells were grown overnight at 37°C on plates supplemented with the appropriated antibiotics for selection of transformants (50 µg/ml of kanamycin and 50 µg/ml of spectinomycin). Single colonies were picked and incubated in selective 2YT medium for 6h at 37°C. After incubation, a small aliquot of each bacterial culture was diluted to a final OD$_{600}$ of ≈ 1.0 using fresh 2YT medium to achieve a final cellular concentration of ≈ 8.8x10^8 cells/ml (Vincke et al. 2012). The diluted aliquot was further subjected to a series of 10-fold dilutions (1:50 to 1:500,000) using fresh 2YT medium. Subsequently, 5 µl of each dilution were spotted onto 2YT agar plates supplemented with hygromycin B (600 µg/ml) in the presence and absence of 0.1 mM IPTG. Plates were incubated for 18 h at 37°C. After incubation plates were scanned using an Epson Perfection V700 Photo scanner. Controls plates lacking hygromycin B were also included in the analysis. Also, a hygromycin B phosphotransferase fusion protein without protein degradation signals (plasmid pAV0076) was used during the test as a positive control.

6.4.2 Library construction and selection of bdSUMO mutants

In order to create the library of bdSUMO mutants, three codons corresponding to positions T60, D67 and Q75 in bdSUMO were randomized. The library was generated by PCR using randomized 5' phosphorylated primers AV0351 and AV0353 (Table 6.2). For the library construction, plasmid pAV0171 was used as DNA template for PCR. pAV0171 encodes for the fusion protein bdSUMO-degron^{NER}-Hph-SUMOstar-ssrA. PCR yielded a single DNA product of 4024 bp checked by agarose gel electrophoresis as in 6.1.4 . The amplified library was purified using the MSB spin PCRapace kit (Stratatec, Germany) and eluted in 50 µl of ddH$_2$O. Then, 10 µg of the purified library were treated with BsaI at 37°C for 2 h in order to generate sticky ends for further ligation of the library. After digestion, the library was purified again using the MSB spin PCRapace kit (Stratatec, Germany) and eluted in 30 µl of ddH$_2$O. Ligation of the library was performed as described in 6.1.8 using 2 µg (0.75 pmoles) of the bdSUMO library in a final reaction volume of 10 µl.

In order to select bdSUMO mutants that were not recognized by SUMOstar protease, the ligated library was introduce into electro-competent *E. coli* NEB Express F' cells previously transformed with plasmid pAV0031. Specifically, 50 µl of competent cells were transformed with 200 ng of the ligated library. Transformed cells were incubated in 1 ml 2YT medium supplemented 0.5% (w/v) of glucose for 2 h and 37°C. After recovery, cells were inoculated on several 2YT medium plates supplemented with hygromycin B (400 µg/ml) and IPTG (0.1mM) for 16h at 37°C. Single colonies were picked and incubated for 16h at 37°C in fresh 2YT medium for the analysis of individual bdSUMO mutants as described in 6.4.1 . To do so, 5 µl of each dilution were pipetted onto 2YT agar plates supplemented with hygromycin B (750 µg/ml) in the presence and absence of 0.2 mM IPTG. Plates were incubated for 18 h at 37°C and further scanned using an Epson Perfection V700 Photo scanner.

Table 6.2 Oligonucleotides used for bdSUMO and bdSENP1 mutant library construction

Primer number	Sequence (5'-3')	Mutation introduced
AV0351	GAT<u>GGTCTC</u>TCGCCTACGTGCGGAA*NNK*A CCCCGGATGAACTGGAAATGGAAGATGG	Q75X
AV353	TCC<u>GGTCTC</u>AGGCGACGACC*MNN*AAAGA GGAAAGCAATTGC*MNN*CATGTCTACAGAC TGAC	T60X ; D67X
AV0359	AAC<u>GGTCTC</u>AGTGCACGAAGCGTCA*NNK*A TTGTGATCACGCGCGA	R269X
AV0358	ACA<u>GGTCTC</u>ACACGGACAGGGTTTCGGA*M NN*TTTGCGGCCACCCAGTGC	N280X
AV0361	GAA<u>GGTCTC</u>ATTACAAAAGCGTTCGT*NNK* TGGACTACG*NNK*CGCAAACTGGGTTACA	R346X ; K350X
AV0360	GAA<u>GGTCTC</u>TGTAATCATAGCCGCCGTTA ATCAGCTTTT	none

A = adenosine; G = guanosine; C = cytosine; T = thymine; N = A/G/C/T; K = G/T; M = C/C
BsaI recognition site is underlined. Randomized codons in bdSUMO and SENP1 are highlighted in red.

6.4.3 Library assembly and selection of bdSENP1 mutants

Four different codons corresponding to positions R269, N280, R346 and K350 in the C-terminal catalytic domain of bdSENP1 were randomized for the construction of the mutant library. Because the four randomized codons are spread over the entire coding sequence of bdSENP1, mutations were introduced using two different subsequent PCR reactions. Positions R269 and N280 were first randomized by an initial PCR reaction using degenerated primers AV0358 and AV0359 (Table 6.2). Plasmid pAV0031 encoding bdSENP1 was used as template for creation of the library. The amplified library was then purified using the MSB spin PCRapace kit (Stratatec, Germany) and eluted in 50 µl of ddH$_2$O. Forward and reverse primers used for the creation of the library contained a 5' BsaI site. Therefore, 15 µg of purified library were treated with BsaI at 37°C for 2 h in order to generate sticky ends for further ligation of the library. The restricted PCR product was then purified again using the MSB spin PCRapace kit (Stratatec, Germany) and eluted in 30 µl of ddH$_2$O. 5 µg (1.9 pmol) of the library were ligated as in 6.1.8 Subsequently, 1 µg of the ligated library was transformed in 100 µl of *E. coli* 5-alpha cells. Transformed cells were resuspended in 1ml of 2YT medium supplemented with 0.5% (w/v) of glucose and incubated for 1h and 37°C. After incubation, the whole 1 ml sample was inoculated in 250 ml of selective 2YT medium for further large-scale DNA purification as in 6.1.3 . The purified library DNA was used for further randomization of positions R346 and K350 of bdSENP1by PCR using randomized primers AV0360 and AV0361 (Table 6.2). Again, a single DNA product of 4621 bp was amplified after PCR. The amplified library was purified using the MSB spin PCRapace kit (Stratatec, Germany) and eluted in 50 µl

ddH$_2$O. Then, 10 μg of purified library were treated with BsaI at 37°C for 2 h. Next, the ligation of the library was performed as described in 6.1.8 using 2 μg (0.75 pmoles) of the digested library in a final volume of 10 μl.

The assembled library was used to select for bdSENP1 that recognized SUMOvera as a cognate substrate. For this, NEB Express F' cells expressing bdSUMOstar-degron^{NER}-Hph-SUMOvera-ssrA were transformed with 200 ng of the ligated library. The cells were resuspended in 1 ml of 2YT buffer supplemented with 0.5% (w/v) of glucose and incubated for 3 h at 37°C. Cells were inoculated on several 2YT agar plates supplemented with hygromycin B (400 μg/ml) and IPTG (0.1mM). Plates were incubated for 16 h at 37°C. The resulting colonies were picked and inoculated in different in 5 ml of fresh 2YT for further analysis. The analysis of individual bdSENP1 mutants was performed as described in 6.4.3 . To do so, 5 μl of each diluted sample were placed onto 2YT agar plates supplemented with hygromycin B (750 μg/ml) in the presence and absence of 0.2 mM IPTG. Plates were incubated for 18 h at 37°C. Finally, plates were scanned using an Epson Perfection V700 Photo scanner.

7 ABBREVIATIONS

aa	Amino acid
Ac	Acetate
Amp	Ampicillin
ATP	Adenosine 5'-triphosphate
bd	Brachypodium distachyon
BSA	Bovine serum albumin
bp	Base pairs
C-terminus (CT)	Carboxy-terminus
COOH	Carboxyl group
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
eEF2	Eukaryotic elongation factor 2
eGFP	Enhanced green fluorescent protein
eIF4A	Eukaryotic initiation factor 4 A
eIF5A	Eukaryotic initiation factor 5 A
FG	Phenylalanine-glycine dipeptide
FW	Framework
GA	Gibson assembly
Gnd-HCl	Guanidinium hydrochloride
GFP	Green fluorescent protein
GTP	Guanosine 5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HygB	Hygromycin B
Hph	Hygromycin B phosphotransferase
IBB	Importin β-binding domain of Importin α
IF2	*E. coli* translation initation factor 2
Impβ	Importin β
IPTG	Isopropyl-β-D-thiogalactopyranosid
Kap	Karyopherin
kDa	Kilodalton
Lph2	Karyopherin 122
LC	Liquid chromatography
MBP	Maltose binding protein
MDa	Megadalton
mCherry	monomeric cherry protein
mRNA	Messenger RNA
MS	Mass spectrometry
NEDD8	Neural precursor cell expressed, developmentally down-regulated 8
NEDP1	NEDD8-specific protease
NES	Nuclear export signal
NLS	Nuclear localization signal
NPC	Nuclear pore complex
NH2	Amino group
N-terminus (NT)	Amino-terminus
NTR	Nuclear transport receptor
Nup	Nucleoporin
OD_{600}	Optical density measured at a wavelength of 600 nm
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
Pdr6	Pleiotropic drug resistance protein 6
PDB	Protein data bank
PMSF	Phenylmethanesulfonylfluoride
ProtA	Protein A from Staphylococcus aureus
Ran	Ras-related nuclear protein

RanGAP	Ran GTPase-activating protein
Ran-GTP	GTP-bound Ran
RanBP2	Ran-binding protein 2
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
rpL12A	Ribosomal protein L12A
rpm	Revolutions per minute
RT	Room temperature
sc	Saccharomyces cerevisiae
ssDNA	Single-stranded desoxynucleic acid
ssrA	10Sa RNA
ssRNA	Single-stranded ribonucleic acid
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SENP1	Sentrin-specific protease 1 / SUMO protease
SUMO	Small ubiquitin-related modifier
TAP	Tandem affinity purification
TB	Terrific broth medium
TCEP	Tris(2-carboxyethyl)phosphin
tCherry	monomeric cherry protein
TEV	Tobacco etch virus
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
Ubc9	E2 SUMO conjugating enzyme
Ulp1	Ubiquitin-like protease
URA	Uracil
v/v	Volume/volume
w/v	Weight/volume
yeGFP	Yeast codon-optimized green fluorescent protein
YT	Yeast extract tryptone medium
ZZ	IgG-binding domain of *Staphylococcus* protein A (in tandem)

8 ACKNOWLEDGMENTS

esta hermosa ciudad alemana. Quiero agradecer a José Flores Uribe porque a pesar de la distancia siempre fuiste un gran apoyo para mí y nunca me dejaste solo, te quiero amigo, tsssss. Por último, el más grande de mis agradecimientos va para mi familia por nunca dejarme solo y apoyarme en cada una de mis decisiones. Sin su apoyo, esta tesis nunca se hubiera escrito, los amo. Madre y padre, les dedico esta tesis.

9 REFERENCES

Adachi, Y. & Yanagida, M., (1989). Higher order chromosome structure is affected by cold-sensitive mutations in a Schizosaccharomyces pombe gene crm1+ which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. *Journal of Cell Biology*, **108(4)**, 1195–1207.

Aitchison, J.D., Blobel, G. & Rout, M.P., (1996). Kap104p: A Karyopherin Involved in the Nuclear Transport of Messenger RNA Binding Proteins. *Science*, **274(5287)**, 624–627.

Aksu, M., Trakhanov, S. & Görlich, D., (2016). Structure of the exportin Xpo4 in complex with RanGTP and the hypusine-containing translation factor eIF5A. *Nature Communications*, **7(5)**, 11952.

Albertini, M. et al., (1998). A novel nuclear import pathway for the transcription factor TFIIS. *Journal of Cell Biology*, **143(6)**, 1447–1455.

Allen, N.P.C. et al., (2001). Proteomic Analysis of Nucleoporin Interacting Proteins. *Journal of Biological Chemistry*, **276(31)**, 29268–29274.

Amor-Mahjoub, M. et al., (2006). The effect of the hexahistidine-tag in the oligomerization of HSC70 constructs. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, **844(2)**, 328–334.

von Appen, A. et al., (2015). In situ structural analysis of the human nuclear pore complex. *Nature*, **526(7571)**, 140–143.

Arnold, M. et al., (2006). Transportin is a major nuclear import receptor for c-Fos: A novel mode of cargo interaction. *Journal of Biological Chemistry*, **281(9)**, 5492–5499.

Arts, G.-J., Fornerod, M. & Mattaj, Iain W., (1998). Identification of a nuclear export receptor for tRNA. *Current Biology*, **8(6)**, 305–314.

Bachmair, a, Finley, D. & Varshavsky, a, (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, **234(4773)**, 179–186.

Bakhrat, A. et al., (2006). Nuclear import of Ho endonuclease utilizes two nuclear localization signals and four importins of the ribosomal import system. *Journal of Biological Chemistry*, **281(18)**, 12218–12226.

Balzi, E. et al., (1987). The multidrug resistance gene PDR1 from Saccharomyces cerevisiae. *The Journal of biological chemistry*, **262(35)**, 16871–9.

Baudin, A. et al., (1993). A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic Acids Research*, **21(14)**, 3329–3330.

Binz, H.K. et al., (2004). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nature biotechnology*, **22(5)**, 575–582.

Bischoff, F.R. et al., (1994). RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proceedings of the National Academy of Sciences of the United States of America*, **91(7)**, 2587–91.

Bischoff, F.R. & Görlich, D., (1997). RanBP1 is crucial for the release of RanGTP from importin β-related nuclear transport factors. *FEBS Letters*, **419(2–3)**, 249–254.

Bischoff, F.R. & Ponstingl, H., (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature*, **354(6348)**, 80–82.

Blow, J.J. & Laskey, R.A., (1986). Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs. *Cell*, **47(4)**, 577–587.

Bohnsack, M.T. et al., (2002). Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO Journal*, **21(22)**, 6205–6215.

Bohnsack, M.T., (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, **10(2)**, 185–191.

Bonner, W.M., (1975). Protein migration into nuclei. I. Frog oocyte nuclei in vivo accumulate microinjected histones, allow entry to small proteins, and exclude large proteins. *Journal of Cell Biology*, **64(2)**, 421–430.

Brachmann, C.B. et al., (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: A

useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, **14(2)**, 115–132.

Brownawell, A.M. & Macara, I.G., (2002). Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *Journal of Cell Biology*, **156(1)**, 53–64.

Brune, C., (2005). Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA*, **11(4)**, 517–531.

Bullock, T.L. et al., (1996). The 1.6 angstroms resolution crystal structure of nuclear transport factor 2 (NTF2). *Journal of molecular biology*, **260(3)**, 422–31.

Bürckstümmer, T. et al., (2006). An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nature methods*, **3(12)**, 1013–1019.

Butt, T., Malakhova, O. & Malakhov, M., (2010). Methods and compositions for enhanced protein expression and purification. *US Patent PCT/US2004/020778*.

Butt, T.R. et al., (2005). SUMO fusion technology for difficult-to-express proteins. *Protein Expression and Purification*, **43(1)**, 1–9.

Caesar, S., Greiner, M. & Schlenstedt, G., (2006). Kap120 functions as a nuclear import receptor for ribosome assembly factor Rpf1 in yeast. *Molecular and cellular biology*, **26(8)**, 3170–3180.

Calado, A. et al., (2002). Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *EMBO Journal*, **21(22)**, 6216–6224.

Cavalier-Smith, T., (1988). Origin of the cell nucleus. *BioEssays*, **9(2–3)**, 72–78.

Chan, P. et al., (2011). Purification of Heterotrimeric G Protein Subunits by GST-Ric-8. *Journal of Biological Chemistry*, **286(4)**, 2625–2635.

Chant, A. et al., (2005). Attachment of a histidine tag to the minimal zinc finger protein of the Aspergillus nidulans gene regulatory protein AreA causes a conformational change at the DNA-binding site. *Protein Expression and Purification*, **39(2)**, 152–159.

Chaves, S.R. & Blobel, G., (2001). Nuclear Import of Spo12p, a Protein Essential for Meiosis. *Journal of Biological Chemistry*, **276(21)**, 17712–17717.

Chaves, S.R. & Rosenblum, J.S., (2011). Efficient nuclear transport of structurally disturbed cargo: Mutations in a cargo protein switch its cognate Karyopherin. *PLoS ONE*, **6(2)**, 1–7.

Cheeseman, I.M. & Desai, A., (2005). A Combined Approach for the Localization and Tandem Affinity Purification of Protein Complexes from Metazoans. *Science Signaling*, **2005(266)**, l1.

Chen, W. et al., (1991). VII. Yeast sequencing reports. The DNA sequencing of the 17 kb HindIII fragment spanning the LEU1 and ATE1 loci on chromosome VII from Saccharomyces cerevisiae reveals the PDR6 gene, a new member of the genetic network controlling pleiotropic drug resistance. *Yeast*, **7(3)**, 287–299.

Chen, X., Pham, E. & Truong, K., (2010). TEV protease-facilitated stoichiometric delivery of multiple genes using a single expression vector. *Protein Science*, **19(12)**, 2379–2388.

Christie, M. et al., (2016). Structural Biology and Regulation of Protein Import into the Nucleus. *Journal of Molecular Biology*, **428(10)**, 2060–2090.

Chuderland, D., Konson, A. & Seger, R., (2008). Identification and Characterization of a General Nuclear Translocation Signal in Signaling Proteins. *Molecular Cell*, **31(6)**, 850–861.

Claustre, H. & Maritorena, S., (2003). The Many Shades of Ocean Blue. *Science*, **302(11)**, 1514–1515.

Cordingley, M.G. et al., (1989). Cleavage of small peptides in vitro by human rhinovirus 14 3C protease expressed in Escherichia coli. *Journal of virology*, **63(12)**, 5037–45.

Cordingleys, G., Colonno, J. & Callahan, P.L., (1990). Substrate Requirements of Human Rhinovirus 3C Protease for PeptideCleavage in Vitro. *Journal of Biological Chemistry*, **265(16)**, 9062–9065.

Coutavas, E. et al., (1993). Characterization of proteins that interact with the cell-cycle regulatory protein Ran/TC4. *Nature*, **366(6455)**, 585–587.

Cox, J. & Mann, M., (2012). 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics*, **13(Suppl. 16)**, S12.

Crevel, G. & Cotterill, S., (1991). DNA replication in cell-free extracts from Drosophila melanogaster. *The*

EMBO Journal, **10(13)**, 4361–4369.

Van Criekinge, W. & Beyaert, R., (1999). Yeast Two-Hybrid: State of the Art. *Biological procedures online*, **2(1)**, 1–38.

Cusick, M.E. et al., (2005). Interactome: Gateway into systems biology. *Human Molecular Genetics*, **14 (Suppl. 2)**, 171–181.

Dean, K. a et al., (2001). Signal recognition particle protein 19 is imported into the nucleus by importin 8 (RanBP8) and transportin. *Journal of cell science*, **114**, 3479–3485.

Dieckhoff, P. et al., (2004). Smt3/SUMO and Ubc9 are required for efficient APC/C-mediated proteolysis in budding yeast. *Molecular Microbiology*, **51(5)**, 1375–1387.

Dower, W.J., Miller, J.F. & Ragsdale, C.W., (1988). High efficiency transformation of E. coli by high voltage electroporation. *Nucleic acids research*, **16(13)**, 6127–45.

Eibauer, M. et al., (2015). Structure and gating of the nuclear pore complex - Supplementary. *Nature Communications*, **6(5)**, 7532.

Endo, Y., Takai, K. & Ueda, T., (2010). *Cell-Free Protein Production* Y. Endo, K. Takai, & T. Ueda, eds., Totowa, NJ: Humana Press.

Enenkel, C., Blobel, G. & Rexach, M., (1995). Identification of a yeast karyopherin heterodimer that targets import substrate to mammalian nuclear pore complexes. *Journal of Biological Chemistry*, **270(28)**, 16499–16502.

Englmeier, L., Olivo, J.C. & Mattaj, I.W., (1999). Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis. *Current Biology*, **9(1)**, 30–41.

Ferrigno, P. et al., (1998). Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin β homologs NMD5 and XPO1. *EMBO Journal*, **17(19)**, 5606–5614.

Fischer, U. et al., (2015). A non-canonical mechanism for Crm1-export cargo complex assembly. *eLife*, **2015(4)**, 1–20.

Fischer, U. et al., (1995). The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell*, **82(3)**, 475–483.

Fornerod, M. et al., (1997). CRM1 Is an Export Receptor for Leucine-Rich Nuclear Export Signals. *Cell*, **90(6)**, 1051–1060.

Frey, S. & Görlich, D., (2014a). A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *Journal of Chromatography A*, **1337**, 95–105.

Frey, S. & Görlich, D., (2007). A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes. *Cell*, **130(3)**, 512–523.

Frey, S. & Görlich, D., (2014b). Purification of protein complexes of defined subunit stoichiometry using a set of orthogonal, tag-cleaving proteases. *Journal of Chromatography A*, **1337**, 106–115.

Frey, S. & Görlich, D., (2015). The Xenopus laevis Atg4B protease: Insights into substrate recognition and application for tag removal from proteins expressed in pro- and eukaryotic hosts. *PLoS ONE*, **10(4)**, 1–25.

Friesen, H. et al., (2005). Interaction of the Saccharomyces cerevisiae cortical actin patch protein Rvs167p with proteins involved in ER to Golgi vesicle trafficking. *Genetics*, **170(2)**, 555–568.

Gadal, O. et al., (2001). Nuclear export of 60s ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. *Molecular and cellular biology*, **21(10)**, 3405–15.

Gant, T.M. & Wilson, K.L., (1997). Nuclear Assembly. *Annual Review of Cell and Developmental Biology*, **13(1)**, 669–695.

Gavin, A.-C. et al., (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature*, **440(7084)**, 631–6.

Giannone, R.J. et al., (2007). Dual-tagging system for the affinity purification of mammalian protein complexes. *BioTechniques*, **43(3)**, 296–302.

Gibson, D.G. et al., (2008). Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science*, **319(5867)**, 1215–1220.

Gibson, D.G. et al., (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, **6(5)**, 343–5.

Gietz, R.D. & Schiestl, R.H., (2007). Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, **2(1)**, 35–37.

Gingras, A., Raught, B. & Sonenberg, N., (1999). eIF4 Initiation Factors: Effectors of mRNA Recruitment to Ribosomes and Regulators of Translation. *Annual Review of Biochemistry*, **68(1)**, 913–963.

Glatter, T. et al., (2009). An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Molecular systems biology*, **5(237)**, 237.

Gloeckner, C.J. et al., (2007). A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics*, **7(23)**, 4228–4234.

Gontan, C. et al., (2009). Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *Journal of Cell Biology*, **185(1)**, 27–34.

Görlich, D. et al., (1997). A novel class of RanGTP binding proteins. *Journal of Cell Biology*, **138(1)**, 65–80.

Görlich, D. et al., (1996). Identification of different roles for RanGDP and RanGTP in nuclear protein import. *The EMBO journal*, **15(20)**, 5584–94.

Görlich, D. et al., (1995). Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Current Biology*, **5(4)**, 383–392.

Görlich, D. & Kutay, U., (1999). Transport Between the Cell Nucleus and the Cytoplasm. *Annual Review of Cell and Developmental Biology*, **15(1)**, 607–660.

Görlich, D., Seewald, M.J. & Ribbeck, K., (2003). Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO Journal*, **22(5)**, 1088–1100.

Gregio, A.P.B. et al., (2009). eIF5A has a function in the elongation step of translation in yeast. *Biochemical and Biophysical Research Communications*, **380(4)**, 785–790.

Greiner, M., Caesar, S. & Schlenstedt, G., (2004). The histones H2A/H2B and H3/H4 are imported into the yeast nucleus by different mechanisms. *European Journal of Cell Biology*, **83(10)**, 511–20.

Grosshans, H. et al., (2001). Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export. *Journal of Cell Biology*, **153(4)**, 745–761.

Gruber, S., Haering, C.H. & Nasmyth, K., (2003). Chromosomal cohesin forms a ring. *Cell*, **112(6)**, 765–777.

Grünwald, M. & Bono, F., (2011). Structure of Importin13-Ubc9 complex: nuclear import and release of a key regulator of sumoylation. *The EMBO Journal*, **30(2)**, 427–438.

Güttler, T. et al., (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nature structural & molecular biology*, **17(11)**, 1367–76.

Güttler, T. & Görlich, D., (2011). Ran-dependent nuclear export mediators: a structural perspective. *The EMBO journal*, **30(17)**, 3457–74.

Hahn, S. & Schlenstedt, G., (2011). Importin β-type nuclear transport receptors have distinct binding affinities for Ran-GTP. *Biochemical and Biophysical Research Communications*, **406(3)**, 383–388.

Harder, B. et al., (2008). TEV protease-mediated cleavage in Drosophila as a tool to analyze protein functions in living organisms. *BioTechniques*, **44(6)**, 765–772.

Harju, S., Fedosyuk, H. & Peterson, K.R., (2004). Rapid isolation of yeast genomic DNA: Bust n' Grab. *BMC Biotechnology*, **4(1)**, 1.

Harper, S. & Speicher, D.W., (2011). Purification of Proteins Fused to Glutathione S-Transferase. In *Methods in molecular biology*, **(681)**, 259–280.

Heikal, A.A. et al., (2000). Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: Coral red (dsRed) and yellow (Citrine). *Proceedings of the National Academy of Sciences*, **97(22)**, 11996–12001.

Hellmuth, K. et al., (1998). Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Molecular and cellular biology*, **18(11)**, 6374–86.

Hendriks, I.A. & Vertegaal, A.C.O., (2016). A comprehensive compilation of SUMO proteomics. *Nature reviews. Molecular cell biology*, **17(9)**, 581–95.

Herrmann, J., Lerman, L.O. & Lerman, A., (2007). Ubiquitin and ubiquitin-like proteins in protein regulation. *Circulation Research*, **100(9)**, 1276–1291.

Hickey, C.M., Wilson, N.R. & Hochstrasser, M., (2012). Function and regulation of SUMO proteases. *Nature Reviews Molecular Cell Biology*, **13(12)**, 755–766.

Higuchi, T. & Uhlmann, F., (2005). Stabilization of microtubule dynamics at anaphase onset promotes chromosome segregation. *Nature*, **433(7022)**, 171–176.

Himeno, H., Kurita, D. & Muto, A., (2014). TmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell. *Frontiers in Genetics*, **5(4)**, 1–13.

Ho, J.H., Kallstrom, G. & Johnson, A.W., (2000). Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *The Journal of cell biology*, **151(5)**, 1057–66.

Hochstrasser, M., (2001). SP-RING for SUMO: New functions bloom for a ubiquitin-like protein. *Cell*, **107(1)**, 5–8.

Hodges, J.L., (2005). Nuclear Import of TFIIB Is Mediated by Kap114p, a Karyopherin with Multiple Cargo-binding Domains. *Molecular Biology of the Cell*, **16(7)**, 3200–3210.

Hood, J.K. & Silver, P.A., (1998). Cse1p is required for export of Srp1p/importin-α from the nucleus in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, **273(52)**, 35142–35146.

Hopper, A.K., Traglia, H.M. & Dunst, R.W., (1990). The yeast RNA1 gene product necessary for RNA processing is located in the cytosol and apparently excluded from the nucleus. *Journal of Cell Biology*, **111(2)**, 309–321.

Huber, J. et al., (1998). Snurportin 1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO Journal*, **17(14)**, 4114–4126.

Huber, L. A, (2003). Is proteomics heading in the wrong direction? *Nature reviews. Molecular cell biology*, **4(1)**, 74–80.

Hyun Ah Kang & Hershey, J.W.B., (1994). Effect of initiation factor eIF-5A depletion on protein synthesis and proliferation of Saccharomyces cerevisiae. *Journal of Biological Chemistry*, **269(6)**, 3934–3940.

Isoyama, T. et al., (2001). Nuclear Import of the Yeast AP-1-like Transcription Factor Yap1p Is Mediated by Transport Receptor Pse1p, and This Import Step Is Not Affected by Oxidative Stress. *Journal of Biological Chemistry*, **276(24)**, 21863–21869.

Izaurralde, E. et al., (1995). A cap-binding protein complex mediating U snRNA export. *Nature*, **376(6542)**, 709–12.

Izaurralde, E. et al., (1997). The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO Journal*, **16(21)**, 6535–6547.

Jackson, R.J. & Hunt, T., (1983). Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods in enzymology*, **96**, 50–74.

Jäkel, S. et al., (2002). Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO Journal*, **21(3)**, 377–386.

Jäkel, S. et al., (1999). The importin β/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *EMBO Journal*, **18(9)**, 2411–2423.

Jäkel, S. & Görlich, D., (1998). Importin β, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO Journal*, **17(15)**, 4491–4502.

Janke, C. et al., (2004). A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, **21(11)**, 947–962.

Jaramillo, M. et al., (1991). RNA unwinding in translation: assembly of helicase complex intermediates comprising eukaryotic initiation factors eIF-4F and eIF-4B. *Molecular and cellular biology*, **11(12)**, 5992–7.

Johnson, E.S. & Blobel, G., (1997). Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *Journal of Biological Chemistry*, **272(43)**, 26799–26802.

Kaffman, A., Rank, N.M., O'Neill, E.M., et al., (1998). The receptor Msn5 exports the phosphorylated

transcription factor Pho4 out of the nucleus. *Nature*, **396(6710)**, 482–6.

Kaffman, A., Rank, N.M. & O'Shea, E.K., (1998). Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes and Development*, **12(17)**, 2673–2683.

Kagey, M.H., Melhuish, T.A. & Wotton, D., (2003). The polycomb protein Pc2 is a SUMO E3. *Cell*, **113(1)**, 127–137.

Kaiser, P. et al., (2008). Tandem Affinity Purification Combined with Mass Spectrometry to Identify Components of Protein Complexes. In M. Starkey & R. Elaswarapu, eds. *Genomics Protocols*. Totowa, NJ: Humana Press, 309–326

Kamitani, T., Nguyen, H.P. & Yeh, E.T.H., (1997). Preferential Modification of Nuclear Proteins by a Novel Ubiquitin-like Molecule. *Journal of Biological Chemistry*, **272(22)**, 14001–14004.

Kataoka, N., Bachorik, J.L. & Dreyfuss, G., (1999). Transportin-SR, a nuclear import receptor for SR proteins. *Journal of Cell Biology*, **145(6)**, 1145–1152.

Katzmann, D.J., Babst, M. & Emr, S.D., (2001). Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell*, **106(2)**, 145–155.

Keiler, K.C., (2008). Biology of *trans* -Translation. *Annual Review of Microbiology*, **62(1)**, 133–151.

Kelley, K. et al., (2015). Atomic structure of the Y complex of the nuclear pore. *Nature structural & molecular biology*, **22(5)**, 425–31.

Kerscher, O., Felberbaum, R. & Hochstrasser, M., (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annual review of cell and developmental biology*, **22**, 159–80.

Kessler, M.M. et al., (1997). Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes and Development*, **11(19)**, 2545–2556.

Khorasanizadeh, S., Peters, I.D. & Roder, H., (1996). Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nature structural biology*, **3(2)**, 193–205.

Kim, K.Y. et al., (2010). Yeast Mpk1 Cell Wall Integrity Mitogen-activated Protein Kinase Regulates Nucleocytoplasmic Shuttling of the Swi6 Transcriptional Regulator. *Molecular Biology of the Cell*, **21(9)**, 1609–1619.

Kimple, M.E., Brill, A.L. & Pasker, R.L., (2013). Overview of affinity tags for protein purification. *Current Protocols in Protein Science*, **(Suppl.73)**, 608–616.

Kirchhofer, A. et al., (2010). Modulation of protein properties in living cells using nanobodies. *Nature Structural & Molecular Biology*, **17(1)**, 133–8.

Kirli, K. et al., (2015). A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *eLife*, **4(12)**, 1–28.

Klebe, C. et al., (1995). Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. *Biochemistry*, **34(2)**, 639–647.

Knott, J.A. et al., (1989). The expression and purification of human rhinovirus protease 3C. *European Journal of Biochemistry*, **182(3)**, 547–555.

Kosobokova, E.N., Skrypnik, K.A. & Kosorukov, V.S., (2016). Overview of fusion tags for recombinant proteins. *Biochemistry*, **81(3)**, 187–200.

Kostallas, G., Löfdahl, P.A. & Samuelson, P., (2011). Substrate profiling of tobacco Etch virus protease using a novel Fluorescence-Assisted whole-cell assay. *PLoS ONE*, **6(1)**.

Kostallas, G. & Samuelson, P., (2010). Novel fluorescence-assisted whole-cell assay for engineering and characterization of proteases and their substrates. *Applied and Environmental Microbiology*, **76(22)**, 7500–7508.

Kostelansky, M.S. et al., (2007). Molecular Architecture and Functional Model of the Complete Yeast ESCRT-I Heterotetramer. *Cell*, **129(3)**, 485–498.

Kovtun, O. et al., (2010). Towards the construction of expressed proteomes using a Leishmania tarentolae based cell-free expression system. *PLoS ONE*, **5(12)**.

Kramer, E.B. & Hopper, A.K., (2013). Retrograde transfer RNA nuclear import provides a new level of tRNA quality control in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, **110(52)**, 21042–7.

Kumar, J.K., Tabor, S. & Richardson, C.C., (2004). Proteomic analysis of thioredoxin-targeted proteins in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **101(11)**, 3759–64.

Kurisaki, A. et al., (2006). The mechanism of nuclear export of Smad3 involves exportin 4 and Ran. *Molecular and cellular biology*, **26(4)**, 1318–32.

Kutay, U. et al., (1997). Export of Importin α from the Nucleus Is Mediated by a Specific Nuclear Transport Factor. *Cell*, **90(6)**, 1061–1071.

Kutay, U. et al., (1998). Identification of a tRNA-specific nuclear export receptor. *Molecular cell*, **1(3)**, 359–69.

Kuwata, T. & Nakamura, T., (2008). BCL11A is a SUMOylated protein and recruits SUMO-conjugation enzymes in its nuclear body. *Genes to Cells*, **13(9)**, 931–940.

Laemmli, U.K., (1970). Cleavage of Structural Proteins during Assembly of Head of Bacteriophage-T4. *Nature*, **227**, 680–685.

Lange, A. et al., (2015). Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin α. *The Journal of Biological Chemistry*, **282(8)**, 5101–5105.

Lee, C.-D. et al., (2008). An improved SUMO fusion protein system for effective production of native proteins. *Protein science : a publication of the Protein Society*, **17(7)**, 1241–8.

van Leene, J. et al., (2007). A tandem affinity purification-based technology platform to study the cell cycle interactome in Arabidopsis thaliana. *Mol Cell Proteomics*, **6(7)**, 1226–1238.

Leslie, D.M. et al., (2002). Kap121p-Mediated Nuclear Import Is Required for Mating and Cellular Differentiation in Yeast. *Molecular and Cellular Biology*, **22(8)**, 2544–2555.

Li, S.J. & Hochstrasser, M., (1999). A new protease required for cell-cycle progression in yeast. *Nature*, **398(6724)**, 246–251.

Li, S.J. & Hochstrasser, M., (2003). The Ulp1 SUMO isopeptidase: Distinct domains required for viability, nuclear envelope localization, and substrate specificity. *Journal of Cell Biology*, **160(7)**, 1069–1081.

Li, S.J. & Hochstrasser, M., (2000). The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. *Mol Cell Biol*, **20(7)**, 2367–2377.

Li, Y., (2010). Commonly used tag combinations for tandem affinity purification. *Biotechnology and applied biochemistry*, **55(2)**, 73–83.

Li, Y. et al., (2011). Highly efficient purification of protein complexes from mammalian cells using a novel streptavidin-binding peptide and hexahistidine tandem tag system: Application to Bruton's tyrosine kinase. *Protein Science*, **20(1)**, 140–149.

Li, Y., (2011). The tandem affinity purification technology: An overview. *Biotechnology Letters*, **33(8)**, 1487–1499.

Liang, Y. et al., (2007). The Role of Trs65 in the Ypt/Rab Guanine Nucleotide Exchange Factor Function of the TRAPP II Complex. *Molecular Biology of the Cell*, **18(7)**, 2533–2541.

Lindborg, M. et al., (2013). High-affinity binding to staphylococcal protein A by an engineered dimeric Affibody molecule. *Protein Engineering, Design and Selection*, **26(10)**, 635–644.

Lipowsky, G. et al., (2000). Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *The EMBO journal*, **19(16)**, 4362–71.

Liu, L. et al., (2008). Enhanced protein expression in the baculovirus/insect cell system using engineered SUMO fusions. *Protein Expression and Purification*, **62(1)**, 21–28.

Liu, Y. et al., (2004). Molecular Chaperone Hsp90 Associates with Resistance Protein N and Its Signaling Proteins SGT1 and Rar1 to Modulate an Innate Immune Response in Plants. *Journal of Biological Chemistry*, **279(3)**, 2101–2108.

Lo, K.Y. et al., (2009). Ribosome stalk assembly requires the dual-specificity phosphatase Yvh1 for the exchange of Mrt4 with P0. *Journal of Cell Biology*, **186(6)**, 849–862.

Lu, Q. et al., (2009). Optimized procedures for producing biologically active chemokines. *Protein Expression and Purification*, **65(2)**, 251–260.

Lund, E. et al., (2004). Nuclear export of microRNA precursors. *Science*, **303(5654)**, 95–98.

Mahajan, R. et al., (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, **88(1)**, 97–107.

Malakhov, M.P. et al., (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *Journal of Structural and Functional Genomics*, **5(1–2)**, 75–86.

Malik, H.S., Eickbush, T.H. & Goldfarb, D.S., (1997). Evolutionary specialization of the nuclear targeting apparatus. *Proceedings of the National Academy of Sciences of the United States of America*, **94(12)**, 13738–13742.

Marblestone, J.G. et al., (2006). Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein science : a publication of the Protein Society*, **15(1)**, 182–9.

Matsuura, Y., (2016). Mechanistic Insights from Structural Analyses of Ran-GTPase-Driven Nuclear Export of Proteins and RNAs. *Journal of Molecular Biology*, **428(10)**, 2025–2039.

Matunis, M.J., Coutavas, E. & Blobel, G., (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *Journal of Cell Biology*, **135(6)**, 1457–1470.

Maurer, P. et al., (2001). The nuclear export receptor Xpo1p forms distinct complexes with NES transport substrates and the yeast Ran binding protein 1 (Yrb1p). *Molecular biology of the cell*, **12(3)**, 539–49.

McCluskey, A.J., Poon, G.M.K. & Gariépy, J., (2007). A rapid and universal tandem-purification strategy for recombinant proteins. *Protein Science*, **16(12)**, 2726–2732.

McCoy, J. & La Ville, E., (1997). Expression and Purification of Thioredoxin Fusion Proteins. In *Current Protocols in Protein Science*. Hoboken, NJ, USA: John Wiley & Sons. 6.7.1-6.7.14.

Mingot, J.-M. et al., (2004). Exportin 7 defines a novel general nuclear export pathway. *The EMBO journal*, **23(16)**, 3227–3236.

Mingot, J.M. et al., (2001). Importin 13: A novel mediator of nuclear import and export. *EMBO Journal*, **20(14)**, 3685–3694.

Mohr, D. et al., (2009). Characterisation of the passive permeability barrier of nuclear pore complexes. *The EMBO journal*, 28(17), 2541–2553.

Monecke, T. et al., (2009). Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP. *Science*, 324(5930), 1087–1091.

Moore, M.S. & Blobel, G., (1994). Purification of a Ran-interacting protein that is required for protein import into the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, **91(12)**, 10212–10216.

Morehouse, H. et al., (1999). The importin/karyopherin Kap114 mediates the nuclear import of TATA-binding protein. *Proceedings of the National Academy of Sciences of the United States of America*, **96(22)**, 12542–7.

Morris, J.H. et al., (2014). Affinity purification–mass spectrometry and network analysis to understand protein-protein interactions. *Nature Protocols*, **9(11)**, 2539–2554.

Mosammaparast, N. et al., (2001). Nuclear import of histone H2A and H2B is mediated by a network of karyopherins. *Journal of Cell Biology*, **153(2)**, 251–262.

Mühlhäusser, P. et al., (2001). Multiple pathways contribute to nuclear import of core histones. *EMBO Reports*, **2(8)**, 690–696.

Müller, S. et al., (2001). SUMO, ubiquitin's mysterious cousin. *Nature reviews. Molecular cell biology*, **2(3)**, 202–210.

Mureev, S. et al., (2009). Species-independent translational leaders facilitate cell-free expression. *Nature biotechnology*, **27(8)**, 747–752.

Nacerddine, K. et al., (2005). The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Developmental Cell*, **9(6)**, 769–779.

Neff, C. & Sachs, A., (1999). Eukaryotic translation initiation factors 4G and 4A from Saccharomyces cerevisiae interact physically and functionally. *Molecular and cellular biology*, **19(8)**, 5557–5564.

Nilsson, B. et al., (1987). A synthetic IgG-binding domain based on staphylococcal protein A. *Protein Engineering*, **1(2)**, 107–113.

Nilsson, J., Weis, K. & Kjems, J., (2002). The C-terminal extension of the small GTPase Ran is essential for defining the GDP-bound form. *Journal of Molecular Biology*, **318(2)**, 583–593.

Ohno, M. et al., (2000). PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell*, **101(2)**, 187–198.

Okada, C. et al., (2009). A High-Resolution Structure of the Pre-microRNA Nuclear Export Machinery. *Science*, **326(5957)**, 1275–1279.

Packer, M.S. & Liu, D.R., (2015). Methods for the directed evolution of proteins. *Nature reviews. Genetics*, **16(7)**, 379–394.

Paine, P.L., Moore, L.C. & Horowitz, S.B., (1975). Nuclear envelope permeability. *Nature*, **254(5496)**, 109–14.

Paraskeva, E. et al., (1999). CRM1-mediated recycling of snurportin 1 to the cytoplasm. *The Journal of cell biology*, **145(2)**, 255–64.

Pattenden, L.K. & Thomas, W.G., (2008). Amylose Affinity Chromatography of Maltose-Binding Protein. *Affinity Chromatography*. Totowa, NJ: Humana Press, 169–190.

Perentesis, J.P. et al., (1992). Saccharomyces cerevisiae elongation factor 2: Genetic cloning, characterization of expression, and G-domain modeling. *Journal of Biological Chemistry*, **267(2)**, 1190–1197.

Peroutka, R.J. et al., (2008). Enhanced protein expression in mammalian cells using engineered SUMO fusions: secreted phospholipase A2. *Protein science : a publication of the Protein Society*, **17(9)**, 1586–1595.

Peroutka III, R.J. et al., (2011). SUMO Fusion Technology for Enhanced Protein Expression and Purification in Prokaryotes and Eukaryotes. *Heterologous Gene Expression in E.coli: Methods and Protocols*. Totowa, NJ: Humana Press, 15–30.

Petrassi, H.M. et al., (2005). A strategy to profile prime and non-prime proteolytic substrate specificity. *Bioorganic and Medicinal Chemistry Letters*, **15(12)**, 3162–3166.

Pichler, A. et al., (2004). The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. *Nature Structural & Molecular Biology*, **11(10)**, 984–991.

Plafker, S.M. et al., (2000). Importin-11, a nuclear import receptor for the ubiquitin-conjugating enzyme, UbcM2. *The EMBO journal*, **19(20)**, 5502–13.

Plafker, S.M. & Macara, I.G., (2002). Ribosomal protein L12 uses a distinct nuclear import pathway mediated by importin 11. *Molecular and cellular biology*, **22(4)**, 1266–75.

Pollard, V.W. et al., (1996). A novel receptor-mediated nuclear protein import pathway. *Cell*, **86(6)**, 985–994.

Poulsen, J.W. et al., (2013). Using Guanidine-Hydrochloride for Fast and Efficient Protein Digestion and Single-step Affinity-purification Mass Spectrometry. *Journal of Proteome Research*, **12(2)**, 1020–1030.

Proc, J.L. et al., (2010). A Quantitative Study of the Effects of Chaotropic Agents, Surfactants, and Solvents on the Digestion Efficiency of Human Plasma Proteins by Trypsin. *Journal of Proteome Research*, **9(10)**, 5422–5437.

Quan, Y. et al., (2008). Evolutionary and transcriptional analysis of karyopherin beta superfamily proteins. *Molecular & cellular proteomics : MCP*, **7(7)**, 1254–69.

Rackham, O. & Chin, J.W., (2005). A network of orthogonal ribosome·mRNA pairs. *Nature Chemical Biology*, **1(3)**, 159–166.

Rao, R.N., Allen, N.E. & Hobbs, J.N., (1983). Genetic and enzymatic basis of hygromycin B resistance in Escherichia coli . Genetic and Enzymatic Basis of Hygromycin B Resistance in Escherichia coli. *Antimicrobial Agents and Chemotherapy*, **24(5)**, 689–695.

Reichelt, R. et al., (1990). Correlation between structure and mass distribution of the nuclear pore complex

and of distinct pore complex components. *Journal of Cell Biology*, **110(4)**, 883–894.

Reverter, D. & Lima, C.D., (2004). A basis for SUMO protease specificity provided by analysis of human Senp2 and a Senp2-SUMO complex. *Structure*, **12(8)**, 1519–1531.

Reverter, D. & Lima, C.D., (2006). Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates. *Nature Structural & Molecular Biology*, **13(12)**, 1060–1068.

Rexach, M. & Blobel, G., (1995). Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell*, **83(5)**, 683–692.

Ribbeck, K. et al., (1998). NTF2 mediates nuclear import of Ran. *The EMBO journal*, **17(22)**, 6587–98.

Ribbeck, K. et al., (1999). The translocation of transportin-cargo complexes through nuclear pores is independent of both Ran and energy. *Current Biology*, **9(1)**, 47–50.

Ribbeck, K. & Görlich, D., (2001). Kinetic analysis of translocation through nuclear pore complexes. *The EMBO journal*, **20(6)**, 1320–30.

Ribbeck, K. & Görlich, D., (2002). The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO Journal*, **21(11)**, 2664–2671.

Rigaut, G. et al., (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology*, **17(10)**, 1030–1032.

Robert, X. & Gouet, P., (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, **42(W1)**, 320–324.

Rogers, G.W., Richter, N.J. & Merrick, W.C., (1999). Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *Journal of Biological Chemistry*, **274(18)**, 12236–12244.

Rothenbusch, U. et al., (2012). Sumoylation regulates Kap114-mediated nuclear transport. *The EMBO Journal*, **31(11)**, 2461–2472.

Rout, M.P. et al., (2000). The yeast nuclear pore complex: Composition, architecture, transport mechanism. *Journal of Cell Biology*, **148(4)**, 635–651.

Rout, M.P. & Blobel, G., (1993). Isolation of the yeast nuclear pore complex. *Journal of Cell Biology*, **123(4)**, 771–783.

Rout, M.P., Blobel, G. & Aitchison, J.D., (1997). A distinct nuclear import pathway used by ribosomal proteins. *Cell*, **89(5),** 715–725.

Rozen, F. et al., (1990). Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Molecular and cellular biology*, **10(3)**, 1134–44.

Rubio, V. et al., (2005). An alternative tandem affinity purification strategy applied to Arabidopsis protein complex isolation. *Plant Journal*, **41(5)**, 767–778.

Sacher, M. et al., (2000). Identification and characterization of five new subunits of TRA *European journal of cell biology*, **79(2)**, 71–80.

Saitoh, H. & Hinchey, J., (2000). Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *Journal of Biological Chemistry*, **275(9)**, 6252–6258.

Sambrook, J., Fritsch, E.F. & Maniatis, T., (1989). *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor laboratory press.* Cold Spring Harbor. NY, USA.

Sandersjöö, L. et al., (2014). A protease substrate profiling method that links site-specific proteolysis with antibiotic resistance. *Biotechnology Journal*, **9(1)**, 155–162.

Sarkar, S. & Hopper, A.K., (1998). tRNA nuclear export in saccharomyces cerevisiae: in situ hybridization analysis. *Molecular biology of the cell*, **9(11)**, 3041–55.

Sato, M. & Toda, T., (2007). Alp7/TACC is a crucial target in Ran-GTPase-dependent spindle formation in fission yeast. *Nature*, **447(7142)**, 334–337.

Saveliev, S. & Ph, D., (2013). Enhanced Protein Mass Spectrometry Analysis with Trypsin / Lys-C Mix Trypsin / Lys-C Mix for Protein Mass Spec Analysis. *Nature methods*, **10(11)**, I–II.

Schiestl, R.H.R. & Gietz, R.D.R., (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Current Genetics*, **16(5–6)**, 339–346.

Schimanski, B. et al., (2005). Highly Efficient Tandem Affinity Purification of Trypanosome Protein

Complexes Based on a Novel Epitope Combination Highly Efficient Tandem Affinity Purification of Trypanosome Protein Complexes Based on a Novel Epitope Combination. *Eukaryotic Cell*, **4(11)**, 1942–1950.

Schlenstedt, G. et al., (1997). Yrb4p, a yeast Ran-GTP-binding protein involved in import of ribosomal protein L25 into the nucleus. *EMBO Journal*, **16(20)**, 6237–6249.

Schmidt, H.B. & Görlich, D., (2016). Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends in Biochemical Sciences*, **41(1)**, 46–61.

Schwartz, T.U., (2016). The Structure Inventory of the Nuclear Pore Complex. *Journal of Molecular Biology*, **428(10)**, 1986–2000.

Schwoebel, E.D. et al., (1998). Ran-dependent signal-mediated nuclear import does not require GTP hydrolysis by Ran. *Journal of Biological Chemistry*, **273(52)**, 35170–35175.

Seeler, J.-S. & Dejean, A., (2003). Nuclear and unclear functions of SUMO. *Nature reviews. Molecular cell biology*, **4(9)**, 690–699.

Seiser, R.M. et al., (2006). Ltv1 is required for efficient nuclear export of the ribosomal small subunit in Saccharomyces cerevisiae. *Genetics*, **174(2)**, 679–691.

Semon, D. et al., (1987). Plasmid-determined bleomycin resistance in Staphylococcus aureus. *Plasmid*, **17(1)**, 46–53.

Senger, B. et al., (1998). Mtr10p functions as a nuclear import receptor for the mRNA-binding protein Np13p. *EMBO Journal*, **17(8)**, 2196–2207.

Seufert, W., Futcher, B. & Jentsch, S., (1995). Role of a ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature*, **373(6509)**, 78–81.

Shen, L. et al., (2006). SUMO protease SENP1 induces isomerization of the scissile peptide bond. *Nature Structural & Molecular Biology*, **13(12)**, 1069–1077.

Sloan, K.E., Gleizes, P.E. & Bohnsack, M.T., (2016). Nucleocytoplasmic Transport of RNAs and RNA-Protein Complexes. *Journal of Molecular Biology*, **428(10)**, 2040–2059.

Stade, K. et al., (1997). Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell*, **90(6)**, 1041–1050.

Straube, K., Blackwell, J.S. & Pemberton, L.F., (2010). Nap1 and Chz1 have Separate Htz1 Nuclear Import and Assembly Functions. *Traffic*, **11(2)**, 185–197.

Stüven, T., Hartmann, E. & Görlich, D., (2003). Exportin 6: A novel nuclear export receptor that is specific for profilin-actin complexes. *EMBO Journal*, **22(21)**, 5928–5940.

Suh-Lailam, B.B. & Hevel, J.M., (2009). Efficient cleavage of problematic tobacco etch virus (TEV)-protein arginine methyltransferase constructs. *Analytical Biochemistry*, **387(1)**, 130–132.

Swinehart, D.F., (1962). The Beer-Lambert Law. *Journal of Chemical Education*, **39(7)**, p.333.

Tagwerker, C., (2006). A Tandem Affinity Tag for Two-step Purification under Fully Denaturing Conditions: Application in Ubiquitin Profiling and Protein Complex Identification Combined with in vivoCross-Linking. *Molecular & Cellular Proteomics*, **5(4)**, 737–748.

Taxis, C. & Knop, M., (2012). TIPI: TEV Protease-Mediated Induction of Protein Instability. In *Methods in Molecular Biology*, **(832)**, 611–626.

Thomas, F. & Kutay, U., (2003). Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *Journal of cell science*, 1**16(Pt. 12)**, 2409–19.

Thomsen, M.C.F. & Nielsen, M., (2012). Seq2Logo: A method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Research*, **40(W1)**, 281–287.

Titov, A.A. & Blobel, G., (1999). The karyopherin Kap122p/Pdr6p imports both subunits of the transcription factor IIA into the nucleus. *Journal of Cell Biology*, **147(2)**, 235–245.

Truant, R. & Cullen, B.R., (1999). The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. *Molecular and cellular biology*, **19(2)**, 1210–7.

Tsai, A. & Carstens, R.P., (2006). An optimized protocol for protein purification in cultured mammalian cells using a tandem affinity purification approach. *Nature protocols*, **1(6)**, 2820–2827.

Turanli-Yildiz, B., Alkim, C. & Petek, Z., (2012). Protein Engineering Methods and Applications. In *Protein Engineering*. 33–58.

Tyanova, S. et al., (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature methods*, **13(9)**, 731–40.

Uetz, P. et al., (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature*, **403(6770)**, 623–627.

Ullah, R. et al., (2016). Correction: Activity of the human rhinovirus 3C protease studied in various buffers, additives and detergents solutions for recombinant protein production. *PLoS ONE*, **11(7)**, 1–13.

Valentini, S.R. et al., (2002). Genetic Interactions of Yeast Eukaryotic Translation Initiation Factor 5A (eIF5A) Reveal Connections to Poly(A)-Binding Protein and Protein Kinase C Signaling. *Genetics*, **160(2)**, 393–405.

Veraksa, A., Bauer, A. & Artavanis-Tsakonas, S., (2005). Analyzing protein complexes in Drosophila with tandem affinity purification-mass spectrometry. *Developmental Dynamics*, **232(3)**, 827–834.

Vertegaal, A.C.O. et al., (2004). A proteomic study of SUMO-2 target proteins. *Journal of Biological Chemistry*, **279(32)**, 33791–33798.

Vetter, I.R. et al., (1999). Structural View of the Ran–Importin β Interaction at 2.3 Å Resolution. *Cell*, **97(5)**, 635–646.

Vincke, C. et al., (2012). Generation of Single Domain Antibody Fragments Derived from Camelids and Generation of Manifold Constructs. In P. Chames, ed. *Methods in molecular biology*. Methods in Molecular Biology. Totowa, NJ: Humana Press, 145–176.

Walls, D. & Loughran, S.T., (2011). Tagging Recombinant Proteins to Enhance Solubility and Aid Purification. In *Methods in molecular biology*. 151–175.

Wang, K.H. et al., (2008). Tuning the strength of a bacterial N-end rule degradation signal. *Journal of Biological Chemistry*, **283(36)**, 24600–24607.

Watson, M.L., (1959). Further observations on the nuclear envelope of the animal cell. *The Journal of biophysical and biochemical cytology*, **6(2)**, 147–156.

Weinmann, L. et al., (2009). Importin 8 Is a Gene Silencing Factor that Targets Argonaute Proteins to Distinct mRNAs. *Cell*, **136(3)**, 496–507.

Wen, W. et al., (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell*, **82(3)**, 463–473.

West, M. et al., (2007). Novel interaction of the 60S ribosomal subunit export adapter Nmd3 at the nuclear pore complex. *Journal of Biological Chemistry*, **282(19)**, 14028–14037.

Winzeler, E.A. et al., (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science*, **285(5429)**, 901–906.

Woestenenk, E.A. et al., (2004). His tag effect on solubility of human proteins produced in Escherichia coli: A comparison between four expression vectors. *Journal of Structural and Functional Genomics*, **5(3)**, 217–229.

Xu, D. et al., (2012). Sequence and structural analyses of nuclear export signals in the NESdb database. *Molecular Biology of the Cell*, **23(18)**, 3677–3693.

Xu, X. et al., (2010). The tandem affinity purification method: An efficient system for protein complex purification and protein interaction identification. *Protein Expression and Purification*, **72(2)**, 149–156.

Xu, Z. et al., (2006). Crystal structure of the SENP1 mutant C603S-SUMO complex reveals the hydrolytic mechanism of SUMO-specific protease. *The Biochemical journal*, **398(3)**, 345–352.

Yan, Y., Orcutt, S.J. & Stricmler, J.E., 2009. The use of SUMO as a fusion system for protein expression and purification. *Chimica oggi*, **27(6)**.

Yang, P., Sampson, H.M. & Krause, H.M., (2006). A modified tandem affinity purification strategy identifies cofactors of the Drosophila nuclear receptor dHNF4. *Proteomics*, **6(3)**, 927–935.

Yi, R. et al., (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes and Development*, **17(24)**, 3011–3016.

Yoshida, K. & Blobel, G., (2001). The karyopherin Kap142p/Msn5p mediates nuclear import and nuclear

export of different cargo proteins. *Journal of Cell Biology*, **152(4)**, 729–739.

Yu, H. et al., (2008). High-Quality Binary Protein Interaction Map of the Yeast Interactome Network. *Science*, **322(5898)**, 104–110.

Zenser, N., Bettinger, K. & Song, K., (2008). Tandem affinity purification systems and methods utilizing such systems. *US Patent 11/537,064*.

Zhang, Z. et al., (2006). Nuclear localization of the Saccharomyces cerevisiae ribonucleotide reductase small subunit requires a karyopherin and a WD40 repeat protein. *Proceedings of the National Academy of Sciences of the United States of America*, **103(5)**, 1422–7.

Zhu, H. et al., (2001). Global analysis of protein activities using proteome chips. *Science*, **293(5537)**, 2101–2105.

Zuk, D., (1998). A single amino acid substitution in yeast eIF-5A results in mRNA stabilization. *The EMBO Journal*, **17(10)**, 2914–2925.

Zuo, X., Li, S., et al., (2005). Enhanced expression and purification of membrane proteins by SUMO fusion in escherichia coli. *Journal of Structural and Functional Genomics*, **6(2–3)**, 103–111.

Zuo, X., Mattern, M.R., et al., (2005). Expression and purification of SARS coronavirus proteins using SUMO-fusions. *Protein Expression and Purification*, **42(1)**, 100–110.

10   CURRICULUM VITAE

Name:	**Arturo Vera Rodríguez**
Date of birth:	**14.06.1988**
Place of birth:	**Mexico City**
Email:	averaro@gwdg.de

EDUCATION

(2013–2017)

Ph.D. Department of Cellular Logistics under the Supervision of Dr. Dirk Görlich at the Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.

Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences (GGNB) and Molecular Biology program at the Georg August University Göttingen, Germany.

(2011–2013)

M.Sc. in Molecular biology. Department of Cellular Logistics under the Supervision of Dr. Dirk Görlich and Dr. Steffen Frey at the Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.

International Max Planck Research School (IMPRS) for Molecular Biology at the Georg August University Göttingen

(2006–2010)

B.Sc. in Biotechnological engineering. Department of Molecular Biology under the supervision of Prof. Dr. Jesus Badillo at the National Polytechnic Institute of Mexico.