

MS-based quantitative analysis of the CRM1 export pathway and
spatial proteomics of the *Xenopus laevis* oocyte

PhD Thesis

in partial fulfillment of the requirements for the degree

“ Doctor rerum naturalium (Dr. rer. nat) ”

in the Molecular Biology Program at the

Georg August University Göttingen,

Faculty of Biology

Submitted by

Samir Karaca

Born in Kardzhali, Bulgaria

Göttingen, September 2014

Thesis committee members:

Prof. Dr. Henning Urlaub (First reviewer)	Bioanalytical Mass Spectrometry Group Max Planck Institute for Biophysical Chemistry , Göttingen Bioanalytics, Department of Clinical Chemistry, University Medical Center, Göttingen
Prof. Dr. Dirk Görlich (Second reviewer)	Department of Cellular Logistics, Max Planck Institute for Biophysical Chemistry, Göttingen
Prof. Dr. Ralph Kehlenbach	Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen

Date of oral examination: 27th October 2014

Affidavit

I hereby declare that the presented thesis entitled "MS-based quantitative analysis of the CRM1 nuclear export pathway and spatial proteomics of the *Xenopus laevis* oocyte" has been written independently with no other sources and aids than quoted.

Göttingen, 29th September 2014

Samir Karaca

Acknowledgments

Foremost, I would like to deeply thank to my mentor, Prof. Dr. Henning Urlaub, for giving me a chance to be part of his group and more importantly, for his support, guidance and advices throughout my work in his lab.

I am grateful to Prof. Dirk Görlich that he gave me a chance to work with him. I am truly motivated by his enthusiasm about science.

I would like to thank Prof. Dr. Ralph Kehlenbach for his guidance, advices, and being part of my committee member. Also, I am grateful to Dr. Ketan Thakar for such productive collaboration and his help for first 2 years of my work.

I want to thank the IMPRS Molecular Biology and GGNB for their financial support and for the scientific environment that they created by providing such an interesting methods courses and lectures. Particularly, I would like to express my gratitude to Dr. Steffen Burkhardt and Kerstin Grüniger for their excellent help, advices from the first day of my arrival to Germany.

I am so grateful to Dr. Koray Kirli for his amazing contribution and help to my work, and his presence whenever I need any help.

I would like to thank my external thesis committee members; Prof. Detlef Doenecke, Dr. Wolfgang Fischle and Dr. Halyna Shcherbata.

I want to thank all members of the Department of Cellular Logistic for the scientific and unscientific environment in the lab. Thanks to Matthias Samwer and Jens Krull for their help and the introduction about *Xenopus*. Additionally, I would like to thank to Volker Cordes for oocyte enucleation and for his help about *Xenopus* work.

Many deepest thanks go to all current and past members of the Bioanalytical Mass Spectrometry Group for creating such a great working environment in the lab and outside the lab. Particularly, I would like to thank; Uwe Plessmann and Monika Raabe for their continuous support, their patience and help even during weekends, Dr. Christof Lenz for his help about mass spectrometric analysis, He-Hsuan for always being available to help and to fix the instruments.

Miro, Romina, Katha, I thank you for your all contributions to my scientific and non-scientific life in the lab and as well as outside the lab.

Particularly, many thanks goes to Miro, firstly for being a supervisor, teaching me so many things about mass spectrometry, SILAC, and even for how to make cool figures, and of course for free gym sessions, amazing cocktails and for the Esprit times. I thank you Ilian, Kundan, Saadia, Sunit, Jasmin, Uzman, Alex creating enjoyable working atmosphere.

I am so happy that I met with amazing group of people, "Leute", in Göttingen. Thank you all for all the memories that we shared together. Göttingen would not be fun and livable place without "Leute".

The "Turkish" community, Koray, Metin, Kevser, and Sinem, deserves my sincere thanks, for their all support, help, contributions to my both scientific and daily-life in Göttingen.

Last, but definitely not least I want to express my sincere gratitude to my family for their continuous, unconditional support and understanding. Particularly, I want to thank my sister, being always available and supportive.

Table of Contents

1	Abstract	1
2	Introduction	2
2.1	Nucleocytoplasmic transport	2
2.1.1	<i>Compartmentation of eukaryotic cells</i>	2
2.1.2	<i>Overview of nucleocytoplasmic trafficking</i>	2
2.1.3	<i>Nuclear export receptors</i>	4
2.1.4	<i>Broad Spectrum Exportins- CRM1 and Exportin 7</i>	6
2.2	<i>Xenopus laevis</i>	9
2.2.1	<i>Xenopus laevis oogenesis</i>	9
2.2.2	<i>Xenopus laevis and proteomics</i>	10
2.3	Mass Spectrometry.....	11
2.3.1	<i>Protein Identification by mass spectrometry</i>	11
2.3.2	<i>Quantitative mass spectrometry</i>	13
3	Material and Methods	17
3.1	Material and reagents.....	17
3.1.1	<i>Laboratory equipment and instruments</i>	17
3.1.2	<i>Chemical and reagents</i>	18
3.1.3	<i>Cell culture media and materials</i>	19
3.1.4	<i>E. coli Strains</i>	19
3.1.5	<i>Commercial kits, buffers and solutions</i>	19
3.1.6	<i>Antibodies</i>	20
3.1.7	<i>Enzymes, proteins and inhibitors</i>	20
3.1.8	<i>Software</i>	20
3.2	Protein Biochemistry.....	21
3.2.1	<i>Protein concentration</i>	21
3.2.2	<i>Ethanol precipitation</i>	21
3.2.3	<i>Denaturing polyacrylamide gel electrophoresis</i>	21
3.2.4	<i>Western blotting and immunodetection</i>	22
3.2.5	<i>Protein expression and purification</i>	22
3.2.6	<i>In vitro binding assays</i>	23
3.3	Cell culture, metabolic labeling and cell-based assays,.....	23
3.3.1	<i>Cell culture and metabolic labeling of HeLa P4 cells</i>	23
3.3.2	<i>Transfection</i>	24

3.3.3	<i>Leptomycin B treatment</i>	24
3.3.4	<i>Subcellular fractionation</i>	24
3.3.5	<i>Indirect Immunofluorescence and confocal microscopy</i>	25
3.4	Xenopus laevis protocols	25
3.4.1	<i>Isolation of Xenopus laevis oocyte nuclei</i>	25
3.4.2	<i>Preparation of Xenopus laevis oocyte extract for CRM1 affinity chromatography</i>	26
3.4.3	<i>CRM1 affinity chromatography with Xenopus laevis oocyte extract</i>	26
3.5	Mass spectrometry based methods	27
3.5.1	<i>Spiking of universal protein standards 2 (UPS2)</i>	27
3.5.2	<i>In gel digestion of proteins</i>	27
3.5.3	<i>Extraction of peptides</i>	28
3.5.4	<i>LC-MS/MS analysis of peptides</i>	28
3.5.5	<i>MS raw data processing</i>	29
3.6	Data analysis and statistics	30
3.6.1	<i>Data analysis and Interpretation of MaxQuant outputs and visualization of data</i>	30
3.6.2	<i>Gene ontology (GO) analysis</i>	31
3.6.3	<i>Protein-protein interaction analysis</i>	31
3.6.4	<i>KEGG orthology assignment and pathway mapping</i>	31
4	Results	32
4.1	Establishment of an approach to detect <i>in vivo</i> localization changes for potential CRM1 cargo identification	32
4.1.1	<i>Preparation of subcellular fractions for MS-based experiment</i>	32
4.1.2	<i>Metabolic labeling of HeLa cells</i>	34
4.1.3	<i>General experimental workflow and overview of the data</i>	35
4.2	Effect of leptomycin B on HeLa proteome	37
4.3	Identification of potential CRM1 cargoes with leptomycin B treatment	40
4.3.1	<i>LC-MS/MS analysis of cytosolic and nuclear fractions</i>	40
4.3.2	<i>Validation of known and novel CRM1 cargoes</i>	45
4.4	Quantitative spatial proteomics and RanGTP dependent CRM1 interactome of <i>Xenopus laevis</i> oocytes	46
4.4.1	<i>Overview of experimental workflow and data analysis</i>	46
4.4.2	<i>In-depth proteome analysis of total oocyte, cytosolic and nuclear fractions</i>	49

4.4.3	<i>CRM1 affinity chromatography and bioinformatics characterization of RanGTP dependent CRM1 interactome</i>	54
4.4.4	<i>Comparison of two orthogonal methods to identify CRM1 export cargoes</i>	59
4.5	Molecular pathway dissection of <i>Xenopus laevis</i> oocyte proteome with respect to RanGTP interactome of CRM1	61
4.5.1	<i>RNA Metabolism</i>	62
4.5.2	<i>Protein Synthesis</i>	65
4.5.3	<i>Ubiquitin Proteasome Pathway</i>	68
5	Discussion	71
5.1	Effects of LMB on the HeLa proteome.....	71
5.2	MS-based quantitative analysis of LMB treated HeLa cell fractions.....	71
5.3	Quantitative proteome analysis of <i>X. laevis</i> oocyte	74
5.4	RanGTP dependent CRM1 interactome with respect to spatial proteome of <i>Xenopus laevis</i>	77
5.4.1	<i>CRM1 affinity chromatography</i>	77
5.4.2	<i>Localization Profile of RanGTP dependent CRM1 interactome</i>	78
5.4.3	<i>Protein Synthesis</i>	79
5.4.4	<i>Vesicle coat proteins and cytoskeleton</i>	80
5.4.5	<i>RNA metabolism and ubiquitin proteasome system</i>	81
5.5	Comparison between LMB treatment in HeLa cells and CRM1 affinity chromatography from <i>X. laevis</i> extract	83
5.6	Conclusion.....	84
6	References	86

List of Figures

Figure 2.1 Schematic representation of nucleocytoplasmic transport.....	5
Figure 2.2 <i>Xenopus laevis</i> oocyte and its giant nucleus.....	9
Figure 2.3 Overview of the MS-based protein identification.....	12
Figure 2.4 MS-based quantitative approaches.....	14
Figure 2.5 Schematic representation of SILAC experimental workflow.....	15
Figure 4.1 Quality control experiments.....	33
Figure 4.2 Metabolic labeling of HeLa cells.....	34
Figure 4.3 Experimental Workflow and general overview.....	36
Figure 4.4 Effect of Leptomycin B on cellular proteome.....	38
Figure 4.5 Comparison of significant lists of different experiments.....	41
Figure 4.6 Selective inhibition of the CRM1 export pathway leads to changes in the subcellular localization of proteins.....	42
Figure 4.7 Validation of several known and novel CRM1 cargoes.....	44
Figure 4.8 Workflow of spatial proteomics and exportome analysis of <i>Xenopus laevis</i> oocytes.....	47
Figure 4.9 Absolute and in-depth proteome of <i>Xenopus laevis</i> oocyte.....	50
Figure 4.10 Spatial proteome of <i>Xenopus laevis</i> oocyte.....	51
Figure 4.11 Top100 spatial proteome of <i>Xenopus laevis</i> oocyte.....	52
Figure 4.12 Quantitative analysis of CRM1 affinity chromatography.....	55
Figure 4.13 Characterization of CRM1 exportome.....	57
Figure 4.14 Comparison of novel CRM1 cargoes between LMB treatment in HeLa cells and CRM1 affinity chromatography with <i>X. laevis</i> oocyte extract.....	60
Figure 4.15 Representation of localization profiling molecular pathways.....	61
Figure 4.16 Localization profile of RNA metabolism.....	63
Figure 4.17 Localization profiling of protein synthesis.....	66
Figure 4.18 Localization profiling of ubiquitin proteasome system.....	69

Abbreviations

%	percent	M	molar
1D	one dimensional	<i>m/z</i>	mass-to-charge ratio
2D	two dimensional	MALDI	matrix-assisted laser desorption ionization
aa	amino acid	min	minute
ACN	acetonitrile	MNase	Micrococcal nuclease
AFM	atomic force microscopy	MS	mass spectrometry
APS	ammonium peroxodisulfate	MS/MS	tandem mass spectrometry
ATP	adenosine triphosphate	Mw	molecular weight
bp	base pair	n	nano
BSA	bovine serum albumin	NE	nuclear extract
<i>C.elegans</i>	<i>Caenorhabditis elegans</i>	OD	optical density
C°	degree centigrade	PAGE	polyacrylamide gel electrophoresis
cm	centimeter	PBS	phosphate buffered saline
Da	Dalton (g/mol)	PCI	phenol/chlorophorm/isoamyl alcohol
DDA	data-dependent acquisition	PEG	polyethylene glycol
ddH ₂ O	double distilled water (sterile)	PMSF	phenylmethylsulfonyl fluoride
DMEM	Dulbecco's Modified Eagle's medium	ppb	parts per billion
DNA	deoxyribonucleic acid	ppm	parts per million
DNase	deoxyribonuclease	PTM	posttranslational modification
DTT	dithiothreitol	RNA	ribonucleic acid
<i>E.coli</i>	<i>Escherichia coli</i>	RP	reversed phase
e.g.	for example, <i>exempli gratia</i>	rpm	revolutions per minute
EDTA	ethylendiaminetetraacetic acid	RT	retention time or room temperature

ESI	electrospray ionization	s	second
<i>et al.</i>	and others, <i>et alii</i>	SCX	strong cation-exchange
FA	formic acid	SDS	sodium dodecyl sulfate
FBS	fetal bovine serum	SILAC	stable isotope labeling with amino acids in cell culture
FW HM	full width at half maximum	snRNA	small-nuclear RNA
g	gram or gravity force	TB	tris/borate
h	hour	TBE	tris/borate/EDTA
H	heavy (SILAC)	Tris	tris-(hydroxymethyl) aminoethane
His-tag	histidine tag	TSS	transcription start site
HPLC	high pressure liquid chromatography	U	unit
i.e.	that is, <i>id est</i>	V	volt
IPI	International Protein Index	v/v	volume per volume
k	kilo	vol.	volume
KDM	lysine demethylase	w/v	weight per volume
KMT	lysine methyltransferase	WT	wild type
l	liter	α	anti-/antibody
L	light (SILAC)	D	delta (deletion)
LC	liquid chromatography	m	micro
m	milli or meter	M	molar

1 Abstract

Nuclear export is a vital cellular process, which supplies the cytoplasm with essential biological macromolecules, such as matured tRNAs and ribosomal subunits. This process is carried out by a set of proteins called exportins. In higher eukaryotes, up to now eight RanGTPase-dependent exportins have been identified. Among these exportins, CRM1 mediates one of the major nuclear export pathways with the broadest range of cargoes. So far, more than 100 structurally and functionally diverse CRM1 cargoes have been described. The CRM1-cargo interaction occurs through recognition of short peptide sequences, which are called nuclear export signal (NES). Although, this consensus sequence is commonly present in the primary structure of many proteins, it might be inaccessible to CRM1. This poses challenges for *in silico* identification of CRM1 cargoes. Therefore, this study focused on experimental identification of CRM1 cargoes with two orthogonal approaches. First, a novel MS-based approach was established to capture *in vivo* localization changes between the nucleus and the cytosol upon inhibition of the CRM1 export pathway with Leptomycin B treatment in HeLa cells. This led to identification of many known and novel CRM1 cargoes (totally ~140). Some of these were verified by microscopic analysis. Second, the CRM1 affinity chromatography was employed to selectively enrich RanGTP dependent interaction partners from *X. laevis* oocyte extract, which resulted in identification of large number of CRM1 binders (~640). These two approaches provided a comprehensive catalog of candidate CRM1 cargoes, which most of them belong to cytoplasmic activities, such as translation, intracellular membrane trafficking and cytoskeleton based processes. Complementary to CRM1 cargo identification, nucleocytoplasmic distribution of *X. laevis* proteome was investigated and this resulted in quantitative mapping of ~6300 proteins. This offered valuable insights into degree of compartmentation of a eukaryotic cell, and the spatial distribution of distinct molecular activities, such as RNA metabolism, protein production and degradation. The nucleocytoplasmic distribution of candidate CRM1 cargoes revealed that ~17% exclusively cytosolic proteins were RanGTP dependent CRM1 binders. This observation highlights that the permeability barrier of the nuclear pore complex is alone insufficient to keep cytosolic proteins out of the nucleus and CRM1 safeguards the nucleus by counteracting leakage of the many cytosolic proteins.

2 Introduction

2.1 Nucleocytoplasmic transport

2.1.1 Compartmentation of eukaryotic cells

Among the most important features of eukaryotic evolution is the emergence of membrane bound compartments within the cell. Each of these compartments contains a distinct repertoire of proteins that mediate specific molecular functions. As a result, a eukaryotic cell is capable of performing a wide range of molecular activities simultaneously with minimal interference. The definitive example of this emergent compartmentation is the spatial separation of the eukaryotic genome by the nuclear envelope (NE). This sequestration is accompanied by numerous benefits; firstly, eukaryotes are able to accommodate and process substantially higher amounts of genetic material and contribute genomic stability (e.g. protection from reactive oxygen species). Secondly, the presence of the NE prevents uncontrolled access of transcription factors to genome, a regulatory process for gene expression unavailable to prokaryotes. Third, the NE uncouples transcription and RNA processing from translation and prevents protein production of unspliced or incompletely spliced transcripts. Hence, eukaryotes are able to handle intron-containing genes. While the benefits are numerous, such subcellular compartmentation imposes the necessity of the controlled exchange of proteins between the cytosol and the nucleus. With protein translation occurring in the cytosol, proteins involved in processes such as DNA replication, transcription, and chromosomal organization must be selectively imported to the nucleus. On the other hand, mRNAs, matured tRNAs and assembled ribosomes have to be exported to the cytosol.

2.1.2 Overview of nucleocytoplasmic trafficking

Kinetic analysis of translocation between the cytosol and the nucleus has shown that a single growing mammalian cell can exchange ~1 million macromolecules per second (Ribbeck and Görlich, 2001). This astonishing bidirectional trafficking is accomplished at the pores embedded into the NE called nuclear pore complexes (NPC). This enormous trafficking is the collective work of ~ 80 individual proteins (Güttler and Görlich, 2011). These proteins can be classified into three functional categories: i) proteins that constitute NPC, namely the nucleoporins, ii) soluble cargo ferrying proteins, nuclear transport receptors (NTRs), iii) components that feed metabolic energy into these systems, namely the RanGTPase system. NPCs are proteinaceous gates embedded into

NE (Watson, 1954), where macromolecule exchange takes place between the cytosol and the nucleus.

The NPC is composed of ~ 30 different nucleoporins, and is 55-66 MDa in yeast (Rout and Blobel, 1993; Yang et al., 1998), and 125 MDa (Reichelt et al., 1990) in vertebrates. However, the architecture of it is quite conserved from yeast to humans. It has characteristic 8-fold symmetry, which was well documented already more than a 40 years ago (Gall, 1967). The central channel of the NPC is the site where active transport occurs. Its diameter is about 40 nm (Panté and Kann, 2002) and it behaves like a sieve, where small molecules, such as metabolites, small proteins, can passively move through the central channel. The free passage of the molecules with a diameter larger than 5nm, equivalent spherical 30 kDa protein, is quite restricted and needs assistance of the NTRs (Mohr et al., 2009). The NTRs not only eliminates size restriction but also accelerates NPC passage of molecules with factors of 100 to 1000 fold over passive diffusion (Ribbeck and Görlich, 2001). This is accomplished with interaction between the NTR and FG (phenylalanine/glycine) repeat domains of the nucleoporins. NTRs constantly make transport cycle between the nucleus and the cytoplasm with forming NTR-cargo complex (directly or via adapter) in one side of the NE, and release cargoes on the other side.

The majority of the transport between the cytosol and the nucleus is facilitated via the *Importin* β (*Imp* β) superfamily of proteins. This superfamily is composed of 21 members in mammals and 14 members in yeast, which are sequence-related to importin β . Common features of the *Imp* β superfamily are an N-terminal RanGTP binding domain and ability to interact with FG-nucleoporins (Ribbeck and Görlich, 2002). Based on their functions, they can be subdivided into three categories importins, exportins and carriers that perform both import and export cycles, which is depicted at Figure 2.1 (selective references; Görlich et al. 1994; Wen et al. 1995; Mingot 2001).

Directionally of the transport is provided by the RanGTPase system. Ran is a member of small Ras related GTPases (Bischoff and Ponstingl, 1991; Drivas et al., 1990; Melchior et al., 1993; Moore and Blobel, 1992). Accordingly, it switches between either in GTP or GDP bound states. Chromatin bound, Ran guanine nucleotide exchange factor (RanGEF) RCC1 induces GDP to GTP exchange. On the other hand, GTPase-activating protein (RanGAP) promotes hydrolysis of Ran bound GTP (Bischoff et al., 1994; Klebe et al., 1995) with help of RanBP1 or Ran binding domain of RanBP2/Nup358. The exclusive nuclear localization of RCC1 and exclusive cytoplasmic localization of RanGAP, RanBP1 and RanBP2/Nup358 generates steep RanGTP concentration across NE, with

~1000 fold higher RanGTP concentration in the nucleus (Görlich et al., 1996; Izaurralde et al., 1997; Kalab et al., 2002; Ribbeck and Görlich, 2002; Richards et al., 1997; Smith et al., 2002).

One major distinction between exportins and importins is the way they harness RanGTP while loading the cargoes (See Figure 2.1). An exportin loads its cargoes in the presence of high RanGTP concentration (in the nucleus). The resulting ternary export complex translocates through the NPC. Upon GTP hydrolysis, at the cytoplasmic site, cargo is released and exportin returns to nucleus for another round of an export. On the contrary, importins operate in exactly the opposite manner, where they load their cargoes at the low RanGTP concentration (in the cytosol). Dimeric import-cargo complexes pass through the NPC and enter the nucleus. Subsequent binding of RanGTP to importin dissociates the cargo, and then resulting importin-RanGTP returns into the cytosol where RanGAP and RanBP1 hydrolyzes GTP. Thus, importin becomes available for another import cycle.

During one full cycle of an import and export, NTRs bring a RanGTP to cytoplasm. Even though, Ran is small enough to passively pass through NPC, constant depletion of RanGTP molecules would deplete the steep RanGTP gradient across NE (Görlich et al., 2003). NTF2 is dedicated for RanGDP import into nucleoplasm, where RCC1 is required for disassembly of NTF2-RanGDP complex.

2.1.3 Nuclear export receptors

In higher eukaryotes, eight RanGTP-dependent exportins have currently been characterized (see Table 1 with selective cargoes). Among 8 exportins, up to now 6 exportins have only been documented in nuclear export and the remaining 2, Imp13 and Xpo4; have shown to be involved both in export and import of proteins. One of the striking differences between exportins is in their cargo spectrum. Two exportins, CAS (Xpo2) and exportin 6 (Xpo6), transport only a single but highly abundant cargo. The former, CAS, transports highly abundant importin α , which is an import adapter of importin β (Kutay et al., 1997). The latter, Xpo6, recognizes single and highly abundant cargo as well, which is actin in complex with profilin. Profilin-actin-ATP complex controls actin filament (F-actin) elongation. This complex can slowly leak into nucleus, and leads to form actin paracrystals in the nucleus. Hence export of this complex is necessary to preserve integrity of the nucleus (Stüven et al., 2003). Xpo-t and exportin 5 (Xpo5) recognize RNA based export signals. Xpo-t directly recognizes matured tRNAs and translocates them to the cytosol in RanGTP dependent manner (Hellmuth et al., 1998; Kutay et al., 1998). On the other hand, Exp5 has broader RNA based export repertoire. Exp5 works in parallel with Xpo-t for tRNA export within a slightly different context.

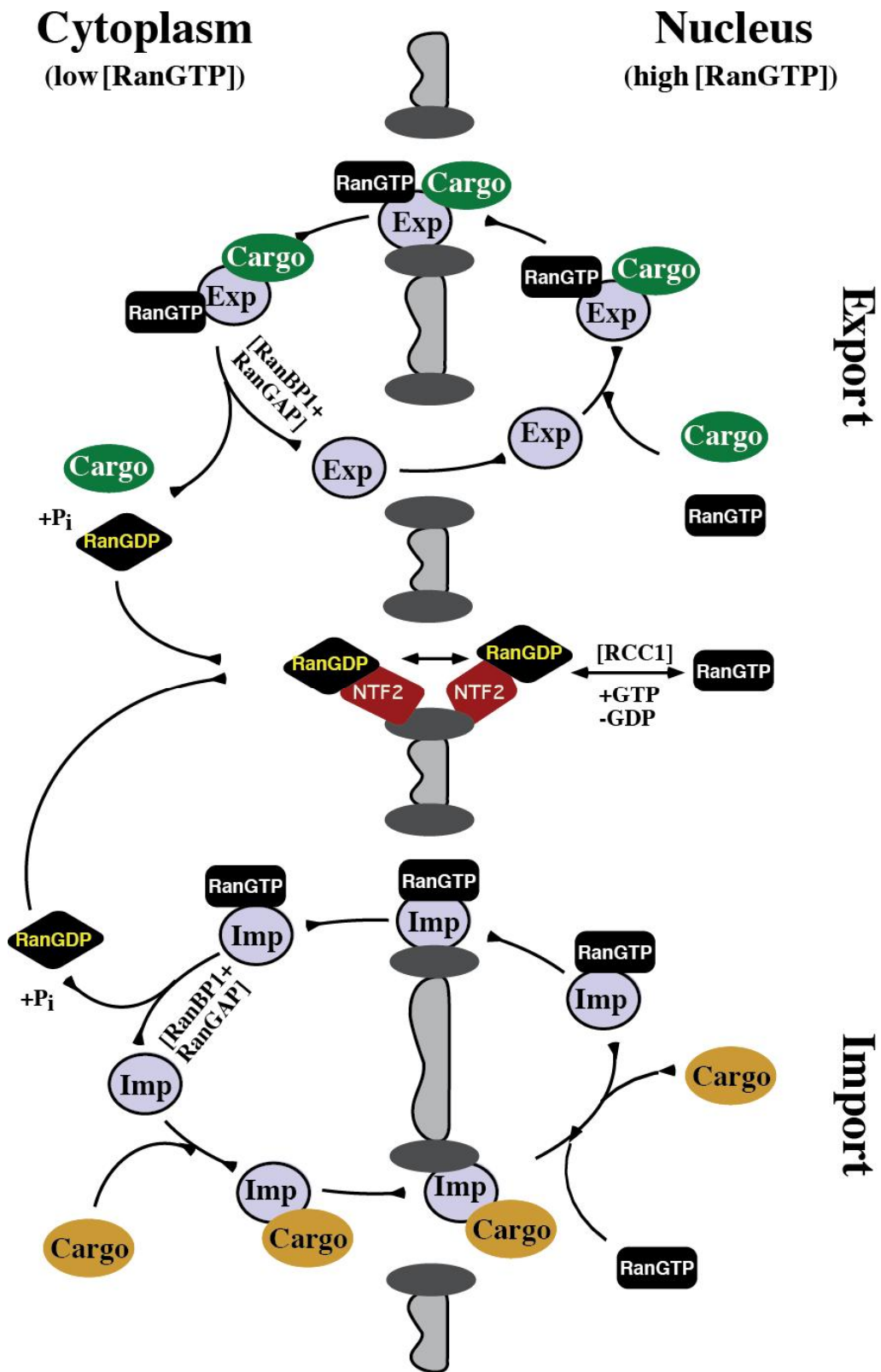


Figure 2.1 Schematic representation of nucleocytoplasmic transport.

This figure has been adapted from (Görllich and Kutay, 1999) for details see text.

The complex of amino acid loaded tRNAs together with eEF1A is a substrate of Xpo5 with tRNA as the direct interaction partner (Bohnsack et al., 2002). Besides, Xpo5 recognizes dsRNAs in sequence independent manner, and transports pre-miRNAs, adenoviral VA1 RNA and human Y1 RNA (Bohnsack et al., 2004; Gwizdek et al., 2003). More recently, it has been shown to assist CRM1 in 60S ribosomal export in higher eukaryotes (Wild et al., 2010). Unlike the exportins mentioned so far, in higher eukaryotes, two members of importin b superfamily works as a bidirectional carriers. Exportin 4 excludes eIF5A and Smad3 from the nucleus, and imports Sox-type transcription factors (Gontan et al., 2009; Kurisaki et al., 2006; Lipowsky et al., 2000). The other carrier, Importin 13, ferries Mago-Y14 complex into nucleus, and eIF1A out of nucleus (Mingot et al., 2001).

2.1.4 Broad Spectrum Exportins- CRM1 and Exportin 7

CRM1 (chromosome region maintenance 1) mediates the export of numerous functionally and structurally unrelated cargoes. More than 100 cargoes were described in the literature. Identification of the CRM1 goes back almost a decade before its role in nucleocytoplasmic transport was documented. Originally, a mutational screen in *Schizosaccharomyces pombe* identified mutations in CRM1 that caused deformed nuclear chromosome domains (Adachi and Yanagida, 1989). Later, in late 1990s, studies showed its essential role in nucleocytoplasmic transport (Fornerod et al., 1997; Stade et al., 1997). *Streptomyces sp.* produce antibiotic called Lemptomycin B (LMB), which is a branched-chain fatty acid. Later, LMB was found as a specific inhibitor of the CRM1 mediated export pathway at low nanomolar concentration (Kudo et al., 1998; Neville and Rosbash, 1999). LMB specifically modifies cysteine 528 in the NES binding region of human CRM1, thereby abolishing cargo CRM1 binding (Monecke et al., 2009). Hence, LMB has been extensively used in to identify potential CRM1 cargoes. Additionally, Xu et al. compiled manually curated CRM1 database, which contains ~ 250 cargoes across all eukaryotes (Xu et al., 2012). Export functions of CRM1 can be categorized into four cargo-based groups: First, it feeds the cytoplasm with macromolecules, which are needed for essential biosynthetic processes, particularly ribonucleoprotein (RNP) complexes. Foremost, it exports 60S ribosomal subunit with adapter protein NMD3 (Gadal et al., 2001; Ho et al., 2000; Thomas and Kutay, 2003). In *Xenopus* oocytes, 5S ribosomal RNA in complex with proteins is stored in the cytosol, and later imported into nucleus. LMB treatment hampers efficient export of 5S RNA indicating the CRM1 mediated export (Murdoch et al., 2002). Another RNP, signal recognition particle (SRP) needs the CRM1 for its export (Ciuffo and Brown, 2000; Grosshans et al., 2001). Furthermore, CRM1 has a role during maturation of spliceosomal U snRNPs. The complex of UsnRNA cap structure, CBC and

PHAX, is recognized by CRM1. Pre U snRNPS are exported for further maturation steps into the cytosol (Izaurralde et al., 1995; Ohno et al., 2000) before they are being imported into nucleus via importin β as Snurportin 1 is an adapter. Taken all together, CRM1 is heavily involved in nuclear export of RNP complexes. Second, CRM1 preserves identity of the nuclei by preventing leakage of RanBP1, RanGAP, and many translation factors into the nucleus (Bohnsack et al., 2002; Maurer et al., 2001; Richards et al., 1997).

Table 1. Functionally characterized vertebrate Importin- β like exportins.

NTR	Selected cargoes	Adapter	Selected references
<i>Exportins</i>			
CRM1 (Exportin 1)	Leu-rich NES containing Cargoes		(Wen et al., 1995)
	HIV genomic RNA	HIV Rev	(Fischer et al., 1995)
	m ⁷ G-capped UsnRNAs	PHAX and CBC	(Izaurralde et al., 1995)
	60S pre-ribosomal subunit	NMD3	(Ho et al., 2000)
	Snurportin 1 (SPN1)		(Paraskeva et al., 1999)
CAS (Exportin 2)	Importin α s		(Kutay et al., 1997)
Exp-t (Xpot)	tRNA		(Kutay et al., 1998)
Exportin 5 (Xpo5)	tRNA, eEF1A (via-tRNA)		(Bohnsack et al., 2002)
	dsRNA-binding proteins (via dsRNA)		(Brownawell and Macara, 2002)
	Pre-miRNAs		
	60S pre-ribosomal subunits		(Bohnsack et al., 2004) (Wild et al., 2010)
Exportin 6 (Xpo6)	Actin-profilin complexes		(Stüven et al., 2003)
Exportin 7 (Xpo7)	P50RhoGAP, 14-3-3s		(Mingot et al., 2004)
<i>Bidirectional NTRs</i>			
Importin 13	<i>Import:</i> Mago-Y14, Ubc9 <i>Export:</i> eIF1A		(Mingot et al., 2001)
Exportin 4 (Xpo4)	<i>Import:</i> Sox2, SRY		(Gontan et al., 2009)
	<i>Export:</i> eIF5A, Smad3		(Lipowsky et al., 2000) (Kurisaki et al., 2006)

Third, CRM1 acts a regulatory element by modulating the spatial distribution of several kinases, transcription factors and cyclins until signal dependent promotion or inhibition of the export. Prominent examples include PKI-mediated nuclear depletion of PKI-PKA (cAMP-dependent protein kinase) NF-AT, Cyclin D1, c-Fos (Benzeno et al., 2006a; Kehlenbach et al., 1998; Sasaki et al., 2006; Wen et al., 1995). Lastly, many viruses exploit CRM1 for their infection cycles. The most prominent example is Human Immunodeficiency Virus (HIV), where HIV-1 Rev protein acts as an adapter protein for CRM1 to export its RNA genome, thereby providing assembly of new viral particles (Fischer et al., 1995; Malim et al., 1989, 1991). Cargo and RanGTP binding to CRM1 operates in positive cooperativity; binding of one increases the affinity of the other. Subsequent to the formation of CRM1 export complex, CRM1 translocates through the NPC with interacting FG-nucleoporins. At the cytoplasmic side, concerted action of RanGAP and RanBP1 or RanBP1 domain of Nup358/RanBP2 hydrolyzes GTP, promoting disassembly of the export complexes.

The molecular basis of CRM1 cargo interaction relies on recognition of linear nuclear export signals (NESs) in RanGTP dependent manner. Initially, classical NESs were named as leucine rich NES. The discovery of other cargoes, randomization and selection studies showed other hydrophobic amino acids, like isoleucine, valine, methionine or phenylalanine could be found in the NES. First, it was proposed that NES follows a consensus sequence of four spaced hydrophobic amino acids (denoted Φ^1 - Φ^4) and it follows this pattern, Φ^1 -(x)₂₋₃- Φ^2 -(x)₂₋₃- Φ^3 -x- Φ^4 where x are preferentially, charged, polar or small amino acids (Kutay and Güttinger, 2005). Recently, structural analysis of different NES peptides revised this consensus sequence such that there is additional, upstream Φ residue (Φ^0), that participates in CRM1 binding (Dong et al., 2009b; Güttler et al., 2010). Besides the linear NESs, three-dimensional structure of protein or domain takes part in CRM1 interaction, for example in SPN1 (Monecke et al., 2009).

Apart from the well-established role of the CRM1 during interphase, after loss of nuclear compartmentation, fraction of CRM1 localizes to kinetochores and centrosomes (Arnaoutov et al., 2005; Wang et al., 2005) ; where it might participate in mitotic specific events, for more details see review Hutten and Kehlenbach, 2006.

Similar to CRM1, Xpo7 can be regarded as a broad spectrum exportin, However, until now, three export cargoes were documented in the literature; LKB1 kinase via adapter STRAD α (Dorfman and Macara, 2008), RhoGAP1 and 14-3-3 σ (Mingot et al, 2004). Mutational analysis of RanGTP dependent interaction partner of Xpo7, eIF1, revealed that this interaction is essentially different than CRM1 cargo recognition. Principally, folded domains and basic residues are important for Xpo7 cargo recognition (Mingot et al., 2004).

2.2 *Xenopus laevis*

The African clawed frog, *Xenopus laevis*, is a pseudotetraploid vertebrate which is native to Sub-Saharan Africa. It is one of the model organisms with important contribution to medical science, due to the several reasons. First, the key cellular and molecular mechanisms are quite conserved.

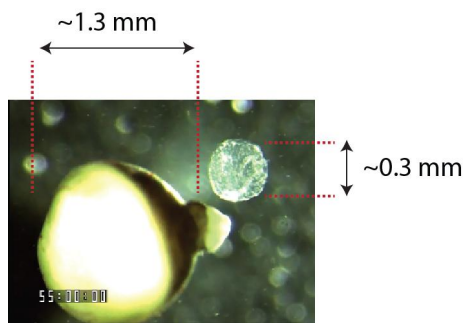


Figure 2.2 *Xenopus laevis* oocyte and its giant nucleus.

Next, by simply rupturing the eggs, large quantities of versatile cell-free extracts can be obtained for use in variety of biochemical assays (Harland and Grainger, 2011). Another feature that lends towards the utility of *X. laevis* as a model organism is that its oocytes, eggs, and embryos can be readily injected with variety of biological molecules, nucleic acids, proteins, or even complete nuclei, to test function of a given gene (Gurdon et al., 1958). Furthermore, by virtue of its gigantic size, *X. laevis* oocyte can be manually dissected to obtain intact nucleus (Figure 2.2).

Scientific studies with *X. laevis* as a model organism have resulted in major achievements, such as nuclear transplantation (Gurdon et al., 1958), discovery of nucleolus for rRNA production (Brown and Dawid, 1968; Brown and Gurdon, 1964), identification key cell cycle components, including discovery of maturation promoting factor (MPF) (Murray and Kirschner, 1989). Despite the numerous practical advantages of *X. laevis* system, it has limitations when it comes to genetic research. First, it has long generation time over a year. Second, it is challenging to study genetics due to pseudotetraploid genome which yielding gene duplicates.

2.2.1 *Xenopus laevis* oogenesis

A *Xenopus laevis* oocyte enters oogenesis slightly bigger than a somatic cell. First, the oocyte replicates its DNA at S-phase and then it spends several months in G2-like growth state during which its nuclear envelope is intact, transcription is active. During this growth, the oocytes can be

divided into 6 stages based on the size and the appearance (Dumont, 1972). In the first two stages, oocytes are transparent and free of yolk protein, vitellogenin. The beginning of vitellogenesis marks stage III. Vitellogenin is synthesized in liver and is taken up by oocytes via micropinocytosis (Romano et al., 2004). Stage IV and V are marked by pigmentation, where animal (dark pigmented area) and vegetal (light pigmented area) pole is formed. During this transformation, the oocyte stockpiles, enormous amount of proteins, mRNAs and tRNAs for later usage. Once oocyte is fully grown, that is Stage VI, it is arrested at prophase meiosis I. It is essentially at G2-like arrest state, which is maintained by inactive state of maturation promoting factors (MPF) (Philpott and Yew, 2008). A stage VI oocyte is an enormously big cell with a diameter about 1.3 mm, and it has volume of 1 μ l, that is roughly 1,000,000 times of a somatic cell (Liu, 2006).

Meiotic maturation is initiated by progesterone, which causes a decrease in cyclic adenosine monophosphate (cAMP) within minutes. The instant decrease in cAMP is reflected by a decrease in protein kinase A (PKA) activity (Tunquist and Maller, 2003).. Withdrawal of the inhibitory effect of PKA's results in activation of MAP kinase cascade (Tunquist and Maller, 2003). This results in germinal vesicle breakdown (GCVD), chromosome condensation and extrusion of the first polar body. Simultaneously, new translation is required for few number of proteins, a prominent example is Mos (Gebauer et al., 1994; Sheets et al., 1995). In the absence of transcription, this rapid protein expression is achieved by lengthening of poly-A tail of Mos mRNA. After maturation, oocytes are arrested again at meiotic metaphase II and wait for the fertilization (Philpott and Yew, 2005).

2.2.2 *Xenopus laevis* and proteomics

Despite being such an important model system, *Xenopus laevis* is underexploited in terms of proteomics. Proteomics experiments rely on well-annotated genome information for faithful protein identification. Even though the *X. laevis* genome sequencing has already initiated, it has proceeded at slower rate due to the pseudotetraploid genome. Such a drawback can be minimized by relying on genome of close relative *X.tropicalis*, which has small diploid genome, and its genome has been recently sequenced (Hellsten et al., 2010). Recent improvements in RNA sequencing technologies make it possible genome-free construction of protein sequence databases (Wühr et al., 2014). During the writing of this thesis, a study compiled mRNA derived reference database for *X. laevis* which outperforms current available protein databases in terms of protein identification (Wühr et al., 2014).

2.3 Mass Spectrometry

2.3.1 Protein Identification by mass spectrometry

The term, protein derives from Greek word *proteios* meaning “ primary” or ‘ of first rank” and was first time used by the Dutch chemist Jöns Berzelius in 1838 to underline the significance of this class of molecules. Indeed, J.J. Berzelious postulated proteins are one of the most important biological molecules that carry out a vast and diverse number of functions inside the cell. Therefore, knowing the composition of proteins in sample, cell extract, or tissue is crucial information. An unbiased protein identification method is Edman degradation (Edman, 1949).It relies on stepwise identification of amino acids from N-terminus via chemical cleavage and subsequent derivatization. Even though, it had significant impact on protein-research, it requires large amount of homogenous samples, free N-termini and its workflow is time and labor consuming. The discovery of soft ionization techniques brought the mass-spectrometry into center of protein-research. Basically, soft ionization techniques allow production of charged ion species without decomposing their chemical structure in the gas phase. The first discovered technique was matrix assisted laser desorption /ionization (MALDI) (Karas and Hillenkamp, 1988; Tanaka et al., 1988) where molecules are co- precipitated with large excess of ultraviolet absorbing matrix. Subsequently, irradiation of matrix with laser pulses transfers energy to the molecules resulting in ionization of peptides. The second technique is electrospray soft ionization (ESI) discovered by John Fenn and his colleagues (Fenn et al., 1989) . A continuous flow of the liquid containing analytes sprayed through a needle at high voltage results in ionization of peptides. Following ionization, molecular ions are channeled to mass analyzer for mass-to charge (m/z) determination. The discovery of these techniques has allowed scientist to measure (m/z) ratios of various molecules e.g. peptides, proteins and even large particles like viruses (Utrecht et al., 2011).

Although m/z measurement of intact proteins is possible, it is much more preferable to analyze peptides for routine protein identification. Proteins might be soluble in MS-compatible buffer. In addition, sensitivity of the MS is much higher for peptides than proteins. More importantly, it is much more difficult to deduce sequence information from peptides. Therefore, prior to MS detection proteins were cut into manageable sizes with endopeptidases.

1D-PAGE not only provides reasonable fractionation but also clears up the sample from digestion and ionization interfering substances (e.g. detergents). Afterwards, proteins are digested with choice of endopeptidase. In general, trypsin is the most preferred endopeptidase since it

produces widest range of MS-compatible peptides through cleaved at the carboxyl site of arginine and lysine. Alternatively, separation can be performed at the peptide level after in-solution or filter-aided sample preparation (FASP) digestion (Wiśniewski et al., 2009). For instance, peptides can be fractionated based via isoelectric focusing, size exclusion or ion exchange chromatography. Prior to MS analysis, resulting peptides are further separated by different approaches; the most routine one is reversed-phase (RP) C18 liquid chromatography (LC) coupled to ESI-MS.

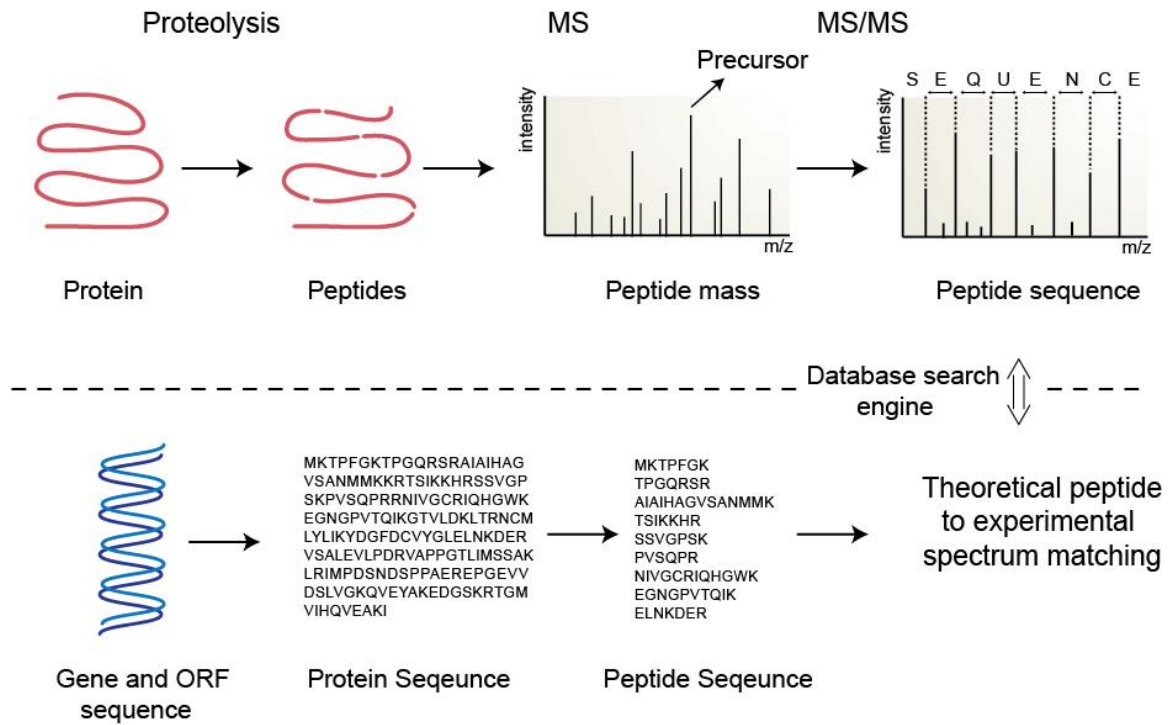


Figure 2.3 Overview of the MS-based protein identification.

This figure has been adapted from (Schmidt, 2010), for details see text.

More complex separation techniques can be applied in the single LC-set up where strong cation exchanged (SCX) is followed by RP-LC. Regardless of which separation technique, at which level is preferred, generated peptides ionized and analyzed by mass spectrometry. Compared to MALDI, ESI-MS is more suitable for wider range of applications, from routine protein identification to complicated proteins quantification purposes. Since the direct coupling of LC to MS provides continuous supply of peptides during MS run. Even though, MALDI-MS is less favored, more specialized applications of it are attracting attention, such as imaging mass spectrometry. In this

approach, MALDI-MS generates spatial map of molecular weight of biomolecules distributed through a sample, such as a tissue sections (Chaurand et al., 2006).

Subsequent to electrospray ionization, peptides move into vacuum system of the instrument and are channeled into mass analyzer for m/z determination. There is variety of mass analyzers with different working principles. The most commonly used mass analyzer for identification and quantification purposes is Orbitrap. It is a barrel-like shape electrode that traps ions that make harmonic motions around it depending on their masses. Generated image current of ions is converted to mass information via Fourier transformation. A single MS cycle is composed of two steps: i) MS spectrum: detections of all ions over defined mass range (e.g. 350-1600 m/z) at a particular instant, ii) repetitive MS/MS scans where selected precursor ion is initially fragmented, then generated fragment ions are measured to record MS/MS spectrum. The MS spectrum contains peptide mass information, and the MS/MS spectrum has peptide sequence information. For routine applications, most mass spectrometers are operated in data-dependent-acquisition (DDA) mode. This means that most abundant precursor ions are sequentially isolated, fragmented and then fragment ions are recorded. Number of MS/MS scans per MS cycle depends on speed of the mass analyzer; state of art instruments can go up to 25 scans per cycle. In an optimized workflow, the speed of the mass analyzer is main limiting factor to increase protein identification (Michalski et al., 2011). Additional features of a mass analyzer are the resolution and accuracy. Low to sub ppm (parts per million) m/z accuracy provides higher confidence in peptide identification, and high resolution provides greater accuracy for quantification and contributes increased identification confidence.

After the generation of raw MS data, a database search is performed (Figure 2.3). A reference protein database digested with respective endopeptidase generates theoretical peptide library containing theoretical masses and fragmentation pattern. Then, experimentally recorded peptide masses and spectrums are statistically matched theoretical peptides. Eventually, identified peptides lead to identification of proteins.

2.3.2 Quantitative mass spectrometry

Only two types information are gained from a MS experiment, mass and intensity. Mass information of precursor and fragment ions is used for peptide identification. Additionally, m/z of ions provides to map post translational modifications (PTM), and helps to elucidate of three-dimensional structure of proteins. The intensity information enables researchers to gain more insights about the dynamic nature of the proteome from relative abundance of proteins to absolute protein copy numbers. However, MS is not quantitative *per se*. Every peptide has

different physicochemical properties. In other words, the MS response of ions is unique and cannot be directly correlated with abundance. Therefore, numerous experimental approaches have been developed to make MS experiment quantitative. Figure 2.4 summarizes current approaches of quantitative mass spectrometry (qMS). Overall, qMS approaches can be separated into two broad categories, relative and absolute quantification. Depending on purpose of experiment, a label-based or label-free approach can be followed. In label-based approaches, chemically identical peptide species are compared at different experimental states. In order to differentiate peptides by MS, a mass tag is introduced into peptides. This is achieved by modifying peptides with stable isotopes, which differs in only m/z .

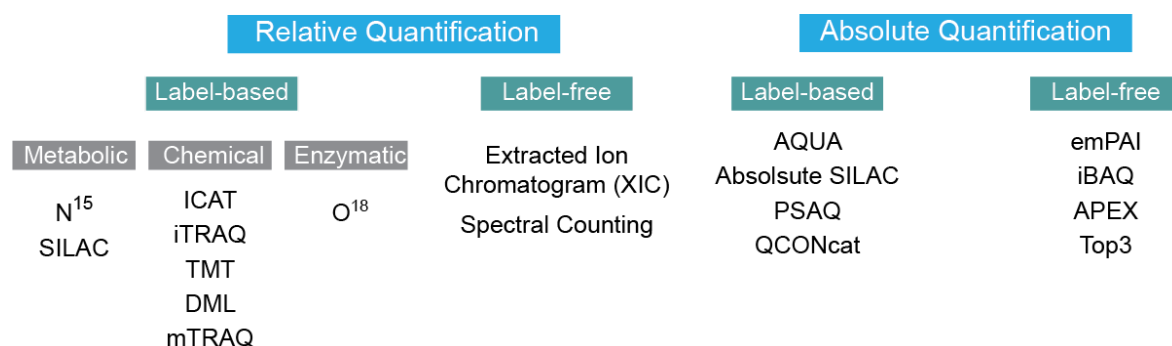


Figure 2.4 MS-based quantitative approaches.

This figure has been adapted from (Nikolov et al., 2012), see text for details.

Subsequent to MS analysis, relative intensities of the same peptide species can directly be compared. A great advantage is that differential labeled samples are analyzed simultaneously; which prevents any biases coming from sample handling and MS analysis. Stable isotopes can be introduced either chemically or enzymatically after protein digestion (Boersema et al., 2009; Mirgorodskaya et al., 2000) , or more preferentially, labeled amino acids can be metabolically incorporated into proteins (Ong et al., 2002). In general, label-based approaches provide high quantification accuracy, but requires an additionally step for labeling. Furthermore, only defined number of samples can be compared within the same workflow. Recent developments in sample preparation have significantly increased this number up to 18 with combination of chemical and metabolic labeling (McAlister et al., 2012). Alternatively, without any prior labeling, label-free approaches can be achieved to make MS experiment quantitative. This has the advantage of allowing the comparison of unlimited samples without any labeling step beforehand. Label-free quantification can be based on peptide spectral counts or MS1 intensity. However, it brings

variation that might arise from sample handling and MS analysis. Therefore, quantification accuracy is relatively moderate compare to label-based approaches. Label-free approaches can be modified to obtain absolute abundance information with spiking already known amounts of standard proteins, e.g. iBAQ (Schwanhäusser et al., 2011).

Stable isotope labeling by amino acids in cell culture

One of the most widely used MS-based relative quantification approaches is stable isotope labeling by amino acids in cell culture (SILAC) (Ong et al., 2002). Heavy isotope labeled essential amino acids are incorporated into the proteome metabolically. The principle advantage of SILAC is the earliest incorporation of the label and ability to mix samples earlier than any other labeling method. Therefore, SILAC reduces biases that might arise from any steps of sample handling and MS analysis. A typical workflow of SILAC-based experiment is depicted at Figure 2.5.

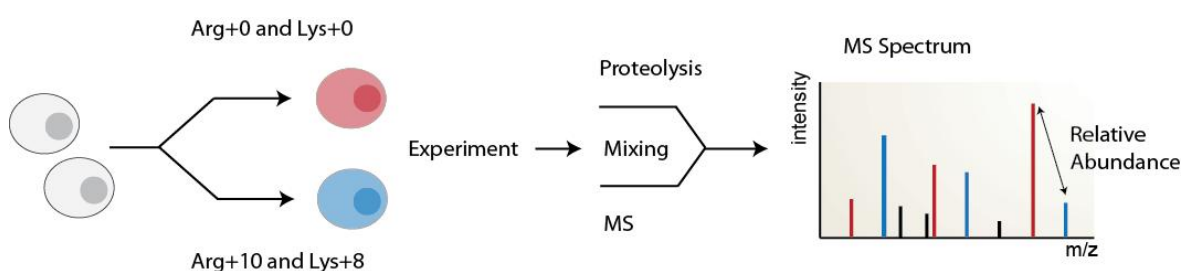


Figure 2.5 Schematic representation of SILAC experimental workflow.

See text for details.

Initially, SILAC was performed only in the cell culture, but has now extended to label whole organisms as well, such as *E.coli*, worm, fly, and mouse (Krüger et al., 2008; Larance et al., 2011; Soufi and Macek, 2014; Sury et al., 2010). In a standard practice, a population of cell is grown in a medium containing natural light isotope (^{12}C , ^{14}N , ^1H) containing amino acids and other population is cultured with "heavy" amino acids (^{13}C , ^{15}N , ^2H). Complete incorporation (> 99%) is achieved after 5-6 generations. Afterwards, experimental set up is separately performed, e.g. differential treatment, overexpression, knockouts, or pull down. Light and heavy samples are mixed, and processed for MS analysis. The most commonly used endopeptidase is trypsin, since it generates widest range of MS-compatible peptides. Trypsin cleaves carboxyl site of arginine and lysine; hereby those amino acids are the most preferred for metabolic labeling. After trypsinization all

but extreme C-terminal peptide is unlabeled unless the last residue is lysine or arginine. MS analysis of an SILAC experiment is carried out in a similar manner as routine protein identification. During computational analysis by using defined mass shift between light and heavy peptides, SILAC pairs are detected, and the direct comparison of SILAC pair intensities provides relative ratios of peptides, and subsequently of proteins.

Intensity based absolute quantification (iBAQ)

Aside from the relative quantification of proteins, another common aim of MS-based approaches is to measure absolute amount of proteins. Most commonly used method for this task is Absolute Quantification of Protein (AQUA) (Gerber et al., 2003). This is achieved with spiking of synthetic heavy-isotope labeled reference peptides into sample mixture with known amounts. The combination of AQUA with "selected reaction monitoring (SRM) " advanced the approach and it became quite successful in measuring the absolute amount of proteins with high accuracy (Picotti et al., 2010). Unfortunately, designing the appropriate references peptides requires substantial time and effort. Therefore this method is not suitable for high throughput applications. Alternatively, high throughput, more cost effective approaches was developed, such as emPAI, APEX and iBAQ (Ishihama et al., 2005; Lu et al., 2007; Schwanhäusser et al., 2011). Intensity based absolute quantification (iBAQ) relies on the sum of the intensities of all detected peptides of a protein normalized by the number of theoretically observable peptides. iBAQ intensities provide the most accurate absolute abundance of all identified protein in sample (Wilhelm et al., 2014b). Additionally, to determine absolute amounts of proteins, a reference protein mixture can be spiked into sample. iBAQ intensities of the reference proteins is plotted against known amounts of reference proteins to generate regression curve to estimate absolute abundance of proteins in the sample. This approach has been successfully employed in different studies (Schwanhäusser et al. 2011; Wilhelm et al. 2014). Applicability of any MS-based approaches strongly relies on the availability of computational platform to analyze the generated raw data. One of the main reasons why SILAC and iBAQ are currently method of choice is the availability of complete computational analysis platform, MaxQuant (Cox and Mann, 2008). MaxQuant together with "sister" software, "Perseus", provides complete data analysis platform for raw data processing, database search, peptide and protein quantification, statistical evaluation and data visualization.

3 Material and Methods

3.1 Material and reagents

3.1.1 Laboratory equipment and instruments

ÄKTA Explorer/Purifier	Pharmacia, Upsala (SWE)
Balances	Metler-Toledo, Giesen (DE)
BBD 6220 CO2 incubator	Heraeus, Hanau (DE)
Centrifuge Cryofuge 6000i	Heraeus, Hanau (DE)
Centrifuge Sorvall Evolution RC Centrifuges	Heraeus, Hanau (DE)
Dumont #5 Forceps	Fine Science Tools, Foster City(USA)
Centrifuges table top 5415R/5810R	Eppendorf, Hamburg (DE)
Electrophoresis power supplies	Bio-Rad, München (DE)
EASY nLC-1000	Thermo Scientific, Dreieich (DE)
Gel Doc 2000 gel documentation system	Bio-Rad, München (DE)
Gene Pulser	BioRad, Burlington (USA)
HP1100 and HP1200 LC systems	Agilent, Santa Clara (USA)
Laminar flow clean bench	Heraeus, Nahau (DE)
LTO Orbital XL	Thermo Fischer Scientific, Bremen (DE)
LTO Orbitrao Velos	Thermo Fischer Scientific, Bremen (DE)
Orbitrap Q Exactive	Thermo Fischer Scientific, Bremen (DE)
NanoDrop ND-1000	Peqlab, Erlangen (DE)
Perfection V700 PRO Scanner	Epson, Nagano (JP)
pH meter	Metler-Toledo, Giesen (DE)
Sonication vath SONOREX Super	BANDELIN Electronic, Berlin (DE)
Sroval SA600 rotor	Thermo Scientific, Braunschweig (DE)
Sorval SS34 rotor	Thermo Scientific, Braunschweig (DE)
SpeedVac Savant SPD121P	Thermo Scientific, Braunschweig (DE)
SensoQuest lab cyclers	SensoQuest, Göttingen (DE)
Thermomixer Comfort	Eppendorf, Hamburg (DE)
UV Transimulator	Bio-Rad, München (DE)
Xcell Sure Lock Mini NuPAGE cell	Invitrogen, Karlsruhe (DE)

3.1.2 Chemical and reagents

4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	VWR, Poole (DE)
Acetic acid	Merck, Darmstadt (DE)
Acetonitrile, LiChrosolv	Merck, Darmstadt (DE)
Acrylamide/Bisacrylamide (37.5:1)	Merck, Darmstadt (DE)
Agarose	Serva, Heidelberg, (DE)
Ammonium hydrogen carbonate	Fluka, Buchs (CH)
Ammonium peroxodisulfate	AppliChem, Darmstadt (DE)
Boric acid	Merck, Darmstadt (DE)
Bovine serum albumin (BSA)	Sigma-Aldrich, Steinheim (DE)
Bromophenol blue	Serva, Heidelberg (DE)
Cleland's reagent (DTT, for MS analysis)	Calbiochem, Darmstadt (DE)
Coomassie Brilliant Blue G-250	Fluka, Buchs (CH)
Deoxynucleotide-5'-phosphate (dATP, dGTP)	Roth, Karlsruhe (DE)
Dithiothreitol (DTT)	AlexisBiochemicals, Farmingdale (USA)
Ethanol	Merck, Darmstadt (DE)
Ethidium bromide (solution 10ml/ml)	Roth, Karlsruhe (DE)
Ethylendiamine tetraacetate (EDTA)	Roth, Karlsruhe (DE)
Formic acid (FA)	Fluka, Buchs (CH)
Glycerol	Merck, Darmstadt (DE)
Hydrochloric acid (37% HCl)	Merck, Darmstadt (DE)
Iodacetamide (IAA)	Sigma-Aldrich, Steinheim (DE)
Magnesium chloride (MgCl ₂)	Merck, Darmstadt (DE)
Methanol, LiChrosolv	Merck, Darmstadt (DE)
N,N,N',N'- Tetramethylethylendiamid (TEMED)	Merck, Darmstadt (DE)
Phenol:Chlorophorm:Isoamil alcohol (PCI) [25:24:1]	Roth, Karlsruhe (DE)
Phenyl-methylsulfonyl fluoride (PMSF)	Roche, Mannheim (DE)
Potassium chloride (KCl)	Merck, Darmstadt (DE)
Potassium dihydrogen phosphate (KH ₂ PO ₄)	Roth, Karlsruhe (DE)
Sodium acetate	Roth, Karlsruhe (DE)

Sodium chloride (NaCl)	Merck, Darmstadt (DE)
Sodium dodecyl sulfate	VWR, Poole (DE)
Sodium hydrogen phosphate (Na ₂ HPO ₄)	Merck, Darmstadt (DE)
Sodium hydroxide (NaOH)	Merck, Darmstadt (DE)
Tricaine	Sigma-Aldrich, Steinheim (DE)
Urea	Merck, Darmstadt (DE)
Universal Protein Standard 2	Sigma-Aldrich, Steinheim (DE)

3.1.3 Cell culture media and materials

DMEM High Glucose (w/o Lys, w/o Arg)	PAA Laboratories, Colbe (DE)
Fetal bovine serum	PAA Laboratories, Colbe (DE)
Dialyzed L-Arginine (Arg0)	Sigma-Aldrich, Steinheim (DE)
L-Arginine, 13C6 (Arg6)	Euriso-top, Saarbrücken (DE)
L-Lysine (Lys0)	Sigma-Aldrich, Steinheim (DE)
L-Lysine, 13C6, 15N2 (Lys8)	Euriso-top, Saarbrücken (DE)
Penicillin/Streptomycin 100x	PAA Laboratories, Colbe (DE)

3.1.4 *E. coli* Strains

BLR	Novagen
NEB Express lq	New England Biolabs
NEB10-beta	New England Biolabs
BL21 CodonPlus	Agilent Technologies
BL21-DE3	Agilent Technologies

3.1.5 Commercial kits, buffers and solutions

Bratford Reagent	Bio-Rad, München (DE)
Mini-PROTEAN 4-12% TGX gels	Bio-Rad, München (DE)
NuPAGE Antioxidant	Invitrogen, Karlsruhe (DE)
NuPAGE LDS Sample Buffer (4x)	Invitrogen, Karlsruhe (DE)
NuPAGE MOPS SDS Running Buffer (20x)	Invitrogen, Karlsruhe (DE)
NuPAGE Novex 4-12% Bis-Tris gels, 1mm	Invitrogen, Karlsruhe (DE)
NuPAGE Sample Reducing Agent (10x)	Invitrogen, Karlsruhe (DE)
SeeBlue Plus2 Protein Standard	Invitrogen, Karlsruhe (DE)

3.1.6 Antibodies

Antibodies	Host	Supplier
α -HA-epitope	mouse	Covance (USA)
α -SP1	rabbit	Thermo Scientific (USA)
α -tubulin	rabbit	Proteintech (USA)
α -GAPDH	rabbit	Proteintech(USA)
α -lamin	mouse	BD Biosciences
α -GST	goat	Amersham Biosciences
α -penta-His	mouse	Qiagen (DE)
α -mouse Alexa-Flour-488	donkey	Molecular Probes (DE)

3.1.7 Enzymes, proteins and inhibitors

Benzonase	Calbiochem, Darmstadt (DE)
Proteinase inhibitor cocktail EDTA free	Roche, Mannheim (DE)
Restriction enzymes	New England Biolabs, Ipswich (USA)
Trypsin	Roche, Mannheim (DE)
Trypsin	Promega, Mannheim (DE)

3.1.8 Software

Adobe Creative Suite 4	Adobe Systems, San Hose (USA)
Cytoscape	Cytoscape Consortium
Lasergene 7	DNASTAR, Madison (USA)
MaxQuant	Max Planck Institute for Biochemistry(DE)
Perseus	Max Planck Institute for Biochemistry(DE)
Microsoft Office	Microsoft Corporation, Redmont (USA)
R language for statistical computing	R Foundation for Statistical Computing
Phyton	SCRI, Invergowrie, Dundee(UK)

3.2 Protein Biochemistry

3.2.1 Protein concentration

Protein concentration was determined by Bradford assay according to the manufacturer's protocol against a BSA standard.

3.2.2 Ethanol precipitation

Protein extracts were precipitated with adding of three volumes of 100 % ethanol (Merck) and one-tenth volume of 3M sodium acetate. Samples were stored -20 °C at least 2hr., then centrifuged at 17.000 g at 4 °C for 15 min. Supernatant was removed and pellet was washed with 80 % ethanol (Merck) and centrifuged again as above and the pellet was air dried for couple of minutes.

3.2.3 Denaturing polyacrylamide gel electrophoresis

Following mass spectrometric analysis

Precipitated proteins were diluted in 1x NuPAGE LDS sample buffer and 1x NuPAGE reducing agent and heated at 70 °C for 10 min. Samples were separated on 4-12 % gradient NuPAGE Novex Bis-Tris gels using 1 x MOPS SDS running buffer supplemented with NuPAGE antioxidant for 50 min at constant 200 V according to instructions provided by the manufacturer SeeBlue Plus2 pre-stained protein standard was used a protein molecular weight reference. SDS-PAGE gels were stained using Coomassie Brilliant Blue G250 staining solution (Neuhoff et al., 1988) overnight and de-stained by rinsing with water for at least 2 h with several washings.

Documentation purposes

The discontinuous SDS-PAGE (sodiumdodecylsulfate polyacrylamide gel electrophoresis) was done according to standard protocols (Sambrook and Russell, 2001) that provide an up to date version of the original description(Laemmli, 1970). Gabriele Kopp prepared gels and Jürgen Schünemann with composition described below. Equipment such as glass plate sets, combs, electrophoresis chambers were custom made by workshop of the MPI-BPC. Running current was constant, at 50 mA, until the bromophenol blue dye present in the sample buffer reached the bottom of the gel. Afterwards, proteins were fixed and stained by heating the gel in 3% acetic acid and 1:100 dilution of the Coomassie stock solution (2 % (w/v) Coomassie Brilliant Blue G250 in 50 % Ethanol). Gels were destained in water and were documented.

Gradient gel ingredients

	"Heavy" Gel 16%	"Light" Gel 7.5%	Stacking Gel 4.5%
2M Tris pH 8,8	40 ml	40 ml	----
0,5 M Tris pH 6,8 H ₂ O	----	----	15 ml
H ₂ O	32 ml	107ml	68 ml
2M Sucrose	10 ml	----	----
Glycerol (87%)	8 ml	----	----
10 % SDS	2 ml	2ml	2ml
Rotiphorese Gel 30	108 ml	51 ml	15 ml
TEMED	120 µl	120 µl	120 µl
APS 10%	2 x 580 µl	2 x 580 µl	100 µl
Total Volume	200 ml	200 ml	100ml

3.2.4 Western blotting and immunodetection

Antibodies and suppliers were listed at 3.1.6. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

3.2.5 Protein expression and purification

The expression of His-CRM1 (Guan et al., 2000) and RanQ69L (Melchior et al., 1995) were done as described before and dialyzed against transport buffer (TPB). Overnight expression of GST-SPN1 in BL21 CodonPlus was done at 16 °C with 0.5 mM IPTG induction. Bacterial lysis was done in buffer C. Overnight expression of GST-CCP1 aa1–120 in BL21-DE3 was done at 16 °C with 0.1 mM IPTG induction. Bacterial lysis was done in buffer C. Bacterial lysis was done in buffer A containing 1% Triton X-100, 4 mM β-mercaptoethanol, 0.4 µg/ml lysozyme and protease inhibitors. Insoluble components were pelleted with centrifugation at 100.000g for 45 min. Glutathione Sepharose beads (GE Healthcare) were with cleared supernatant for 1.5 h at 4 ° C. Beads were washed three times and elution of bound proteins were done with 15 mM glutathione in the appropriate buffer and dialyzed against TPB. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

<u>Transport Buffer (TPB)</u>	<u>Buffer C</u>	<u>Buffer A</u>
20 mM HEPES-KOH pH 7.3	50mM Tris pH 6.8	50mM Tris pH 8
110 mM KOAc	300mM NaCl	250 mM NaCl
2mM Mg(OAc) ₂	1 mM MgCl ₂	2 mM MgCl ₂
2 mM DTT	0.25 mM EDTA	10% glycerol
Protease inhibitors	1mM DTT	Protease inhibitors
	Protease inhibitors	

3.2.6 In vitro binding assays

Before immobilization of 5 µg of GST fusion proteins onto glutathione Sepharose beads (GE Healthcare), beads were incubated with 10 mg/ml BSA in buffer B (50 mM Tris pH 7.4, 200 mM NaCl, 1 mM MgCl₂, 5 % glycerol). After several washings, beads were incubated with 5 µg of CRM1 alone or with RanQ69L (GTP) in 300 µl of buffer containing 2 mg/ml BSA. Binding was done at 4°C for 1.5 hr, then beads were washed three times with buffer B. Elution of bound proteins were done with SDS-sample buffer, then they are subjected SDS-PAGE for immunoblotting. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

3.3 Cell culture, metabolic labeling and cell-based assays,

3.3.1 Cell culture and metabolic labeling of HeLa P4 cells

HeLa P4 cells (Charneau et al., 1994) were grown in custom High Glucose Dulbecco' s Modified Eagle' s Medium (DMEM) medium deficient in the amino acids arginine and lysine supplemented with either light (Arg0 and Lys0) or heavy (Lys8 and Arg6) amino acids (3.1.3) with final concentration of 50 mg/ml, 10 % fetal bovine serum (FBS) and 1 X Penicillin/Streptomycin (100 IU/ml and 100 µg/ml, respectively), were added before culturing of the cells. Medium was changed every 2-3 days and cell adapted to the appropriate SILAC medium for at least five cell doublings. The cells were cultivated at 37° C, 5 % CO₂ and 95 % relative humidity. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

3.3.2 Transfection

Transfection of HeLa P4 cells (2×10^6 cells) with plasmids containing GFP-Sequestosome, HA-DDX3, YFP-CCP1, FLAG_CIP2A, HA_GNL3L, GFP-NMD3, HA_RanBP1, GFP-PDCD2L and Rev₍₄₈₋₁₁₆₎-GFP₂-M9, as a positive control, were done by calcium phosphate method as described before (Ausubel et al., 1994). This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

3.3.3 Leptomycin B treatment

After 24 h of transfection of HeLa P4 cells with Rev₍₄₈₋₁₁₆₎-(GFP)₂-M9 containing plasmid, positive control for LMB treatment, were treated with 10 nM LMB for three hours. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

3.3.4 Subcellular fractionation

Two modes of experimental workflow were done for SILAC analysis of HeLa cells. Forward mode: equal number of light" and LMB treated "heavy" cells and LMB-treated light" and "heavy" cells were mixed. Cells were trypsinized, washed with in cold medium and spun down at 4 °C for 5 min at 100 *g. Pellet were resuspended in PBS and centrifuged as above. 10% of the cells were collected and directly boiled in SDS, spun down at 14.000g. The supernatant was collected as total cell lysate to check effect of the LMB on total cellular proteome. The rest of the cells initially were incubated with ice-cold buffer 1, on ice for 10 min and centrifuged at 2000g. The supernatant was collected and named as cytosol enriched fraction. The pellet incubated with 400 µl ice-cold buffer 2 on ice for 30 min, and then centrifuged at 7000g. The supernatant contained organaller enriched fractions, such as membrane bound organalles, endoplasmic reticulum, Golgi, mitochondria, and some nuclear luminal proteins (M) were removed. The pellet was resuspended and incubated with 400 µl of ice-cold buffer 3, at 4 ° C for 1 h, and centrifuged at 7000g for 10 min. The supernatant contained extracted soluble nuclear proteins, and pellet contained insoluble fraction and membrane bound fraction. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

<u>Buffer 1</u>	<u>Buffer 2</u>	<u>Buffer 3</u>
50 mM HEPES pH 7.4	50 mM HEPES pH 7.4	50 mM HEPES pH 7.4
150 mM NaCl	150 mM NaCl	150 mM NaCl
1% digitonin (1 μ l/10 ⁶ cells)	1% Nonidet P-40	0.5% sodium deoxycholate
Protease inhibitors	Protease inhibitors	0.1% SDS
		Benzonase (1 U/ml)
		Protease inhibitors

3.3.5 Indirect Immunofluorescence and confocal microscopy

Immunofluorescence staining was done as described previously (Hutten and Kehlenbach, 2006) after 24h of transfection using Hoechst 33258 as a DNA-stain. Images were taken with using a LSM 510-Meta confocal microscope and processed using Axio Vision Rel. 4.8 LE, and Adobe Photoshop 6.0. This protocol was performed by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

3.4 *Xenopus laevis* protocols

3.4.1 Isolation of *Xenopus laevis* oocyte nuclei

Anesthetization of female *X. laevis* in 3g/L Tricaine (Sigma) according to animal rights regulations and lobes of ovary were surgically removed by a skilled technician (Liu, 2006). Following all steps were carried out at 18°C. Approximately 10 ml of ovary was cut into smaller pieces, put into 50ml Falcon tube and treated with 50 mg of collagenase (Worthington Biochemical Corp., Lakewood, NJ) and volume completed to 50 mL and incubated for 1 hr on a shaker. Following two washing step in Ca²⁺-free MBM for 20 min inactivated the collagenase. After visual inspection, every time 2 healthy stage VI oocytes were transferred into 35 mm glass Petri dish containing 5:1/HEPES buffer (10 mM pH 7.5, 83 mM, 17 mM NaCl) for enucleation under a microscope (Leica). Enucleation was two pairs of No.5 forceps (DuMont). One pair has bent and blunt tip, this was used to hold and lock the oocyte. Other one has straight and sharp ends. This one was used to make small hole on the animal pole. After making small hole, it was widened slightly to make opening for nucleus to squeeze out. A slight pressure is used to gently squeeze out nucleus from oocyte. Then immediately, intact nucleus was transferred into new petri dish containing 5:1/HEPES buffer. Gently but repetitive-washings with pipetting to clean the nucleus from yolk protein, and cytosolic contaminants were done. As the nucleus, gets cleaner, it was getting more

transparent. Then it was directly transferred into Eppendorf tubes containing absolute ethanol for proteins precipitation. The remaining enucleated oocyte was transferred new Eppendorf tube containing 5:1/HEPES buffer. Totally, 60 oocytes were enucleated. For total oocyte analysis, 60 healthy stage VI oocytes were selected. Total oocyte and enucleated oocytes were diluted with 5:1/HEPES buffer (total volume of 1.2 mL) containing protease inhibitors, and homogenized small pestle in ice. Successive two low speed centrifugations (S45A rotor; 30,000 g; 10 min; 4°C), remaining aggregates, membrane fragments and yolk particles were removed from the extract and supernatant were collected. In order to directly compare the protein concentration between cytosol and nucleus (assuming yolk free-cytosol 500 nl and nucleus 50 nl), 1-to-10 volume normalization was done for total and enucleated oocytes. 10 % of the starting volume was precipitated with ethanol precipitation.

3.4.2 Preparation of *Xenopus laevis* oocyte extract for CRM1 affinity chromatography

Oocytes were obtained from skilled technician and collagenase (Sigma) treated as previously described at 3.4.1. Oocytes were crashed with dounce homogenizer and extract was directly stored at -80 °C until usage. Koray Kirli from Department of Logistics, MPI- Biophysical Chemistry, Goettinge performed the rest of the protocol. The extract is diluted to 1:5 volume in binding buffer (20mM HEPES pH 7.5, 90mM KAc, 2mM MgOAc, 250mM Sucrose, 5mM DTT, 2.5uMCytochalasinB), including 1X sigma protease inhibitor. Series of centrifugation was performed. First centrifugation was done for 30 min 4°C 17000g to get rid of yolk, lipids and insoluble particles. The oocyte extract was collected with syringe. Second centrifugation was done for 30min with S50A rotor at 4°C 100000g, and lastly extract was centrifugated for 1h 4°C S50A rotor at 100000g. Afterwards, the extract was incubated with half volume of phenyl sepharose matrix (low substitution) for 30 min in at 4°C with slow rotation. Beads were settled, supernatant were collected. Phenyl sepharose depleted extract was aliquoted and frozen in liquid nitrogen, and stored at -80 °C until usage.

3.4.3 CRM1 affinity chromatography with *Xenopus laevis* oocyte extract

The following protocol was performed by Koray Kirli from Department of Logistics, MPI- Biophysical Chemistry, Goettingen. The expression of CRM1 was done as described (Kirli). Each binding assay was performed with 0.5 nmole CRM1 that was immobilized onto 20 µl of streptavidin-agarose beads in the Mobicols (MoBiTec, Göttingen). Volume was completed to 500 ul with binding buffer (20mM HEPES pH 7.5, 90mM KAc, 2mM MgOAc, 250mM Sucrose, 5mM DTT) and immobilization was performed for 1hr at 4°C on SB3 rotator (Bibby Scientific, France) at 10rpm. Afterwards, beads were washed 5 times with 500 µl binding buffer. Beads were 5 min

incubated with excess free biotin on ice and then washed 3 times again. 500 μ l of buffer added and reaction mixture was transferred into 1.5 ml Eppendorf tubes, spanned down, carefully buffer was removed. Pelleted beads were diluted 1:1 binding buffer. Phenyl sepharose depleted extracts were quickly thawed, and centrifuged for 15min at 4°C in S45A rotor at 37000 rpm. CRM1 affinity chromatography was performed in 1.5 mL Eppendorf tubes with 20 μ l of CRM1 immobilized streptavidin agarose beads μ l in total volume of 1 mL *Xenopus laevis* oocyte extract. To mimic nuclear environment, 5 μ M Ran₅₋₁₈₀ Q69L (GTP loaded) was added. Samples without RanGTP same volume of RanGTP buffer was added with same NaCl concentration. Reaction mixtures were incubated in cold room at 4°C on SB3 rotator at 10 rpm for 3 hours. Afterwards, samples were centrifuged at 1000 rpm at 4°C in a refrigerated tabletop centrifuge. Flow-through carefully discarded and beads were resuspended with 500 μ l binding buffer and transferred into Mobicols. Beads were washed three times binding buffer, and then Mobicols were placed into 1.5 ml Eppendorf tubes. Bound proteins were eluted with 2 times 30 μ l SDS sample buffer and at each time, samples were waited 5 min at 40 °C. Mobicols in 1.5 ml Eppendorf tubes were centrifuged at 1000 rpm for 1 min at room temperature tabletop centrifuge. Totally 60 μ l elution was collected, and 15 μ l of each elution was analyzed on SDS-PAGE for visualization and rest was used for MS analysis.

3.5 Mass spectrometry based methods

3.5.1 Spiking of universal protein standards 2 (UPS2)

In-depth characterization of *X. laevis* nuclear, cytosolic and total oocyte extract, 10.6 ug of UPS2 (Sigma) containing bacterially expressed 48 human proteins were spiked into samples before running SDS gel.

3.5.2 In gel digestion of proteins

In-gel digestion was performed as described (Shevchenko et al., 2006) with minor modifications. Unless otherwise stated, all incubation steps were performed at 26°C in thermomixer (Eppendorf, Hamburg, DE) at 1050 rpm for 15 min. All solutions were prepared with LiChrosolv H₂O (Merck, Darmstadt, DE). After Coomassie Brilliant Blue staining, each lane was cut into 23 equally sized pieces with using in-house built device. Briefly, gel slices first washed with 150 μ l H₂O and dehydrated with 150 μ l acetonitrile (ACN). Dried gel pieces were incubated with 150 μ l of 10 mM DTT reducing solution for 50 min, then alkylated with 55 mM iodoacetamide for 20 min at 26 °C in dark. After another round of hydration of gel pieces with ACN, they were rehydrated with 15-20 μ l of digestion buffer containing trypsin (Roche) for 30 min on ice. Afterwards appropriate

amount of digestion buffer without trypsin were added to completely cover the gel pieces. They were let overnight digestion with trypsin at 37 °C.

Digestion buffer

15 µl Trypsin (0.1 µg/µl)

50 µl Ammonium bicarbonate (50 mM, pH 8.0)

50 µl H₂O (LiChrosolv)

3.5.3 Extraction of peptides

Extraction of peptides was done as described before (Shevchenko et al., 2006). Briefly, all incubation steps were performed using a thermomixer (1050 rpm) at 37°C for 15 min. First round of incubation was done with adding of 30 µl of water, and it is followed by 100 µl ACN. The supernatant was transferred into new tubes, and the dehydrated gel pieces were incubated with 50 µl 5 % [v/v] formic acid (FA), followed by addition of 50 µl ACN. Supernatant was pooled with previous step supernatant, and gel pieces incubated with another 50 µl ACN. Supernatant from all extraction steps were pooled together and dried in vacuum centrifuge (Thermo Scientific, Braunschweig, DE). Dried peptides were stored at -20 C until submitted to LC-MS.

3.5.4 LC-MS/MS analysis of peptides

LC-MS/MS analysis HeLa P4 SILAC samples

The extracted peptides initially were dissolved in 20 µl 3 % ACN/ 1% [v/v] FA by vortexing and brief sonication on water bath. Each MS run one-fourth of the sample were loaded into an in-house packed C18 trap column (1.5 cm, 360 µm o.d. 150 µm inner diameter, Reprosil-Pur 120 Å, 5 µm outer diameter, C18-AQ, Dr. Maisch GmbH, Germany) at a flow rate of 10 µl/min. Retained peptides were eluted and separated on an analytical C18 capillary column (15 cm, 360 µm outer diameter, 75 µm inner diameter, Reprosil-Pur 120 Å, 5 µm, C18-AQ, Dr. Maisch GmbH, Germany) at a flow rate of 300 nl/min with a gradient from 5% to 37% acetonitrile in 0.1% formic acid for 50 min including column equilibrium and wash by using an Agilent 1100 nano-flow LC system (Agilent Technologies, Santa Clara, CA). Agilent 1100 nano-flow LC was coupled to LTQ-Orbitrap Velos (Thermo Electron, Bremen, Germany) and it was operated in data-dependent mode. The survey scans were acquired in the Orbitrap (m/z 350–1600) with a resolution of 30,000 at m/z 400 with a target value of 1×10^6 . For Up to 15 of the most intense ions with charges ≤ 2 from the survey scan were sequentially selected for collision-induced dissociation (CID) in the LTQ linear ion trap

with a normalized collision energy of 35 %, activation $q=0.25$. In order to avoid repeating the sequencing of peptides dynamic exclusion was set to 60 seconds.

LC-MS/MS analysis *Xenopus Laevis* samples

The extracted peptides initially were dissolved in 20 μ l 3 % ACN/ 1% [v/v] FA by vortexing and brief sonication on water bath. Samples were analyzed triplicates and for each MS run one-fourth of the sample were concentrated on a Reversed Phase-C18 precolumn (0.15 mm ID x 20 mm self-packed with Reprosil-Pur 120 C18-AQ 3 μ m material) and separated by reversed phase-C18 nanoflow chromatography (0.075 mm ID x 200 mm Picofrit column (New Objective, Woburn, MA/USA) self-packed with Reprosil-Pur 120 C18-AQ 3 μ m material) using a 120 min linear gradient (5-35% acetonitrile vs. 0.1% formic acid, 300 nl/min) on an EASY nLC-1000 system (Thermo Scientific, Dreieich). The eluent was analyzed using a Top15 method in Data Dependent Acquisition mode on a Q Exactive high-resolution mass spectrometry system (Thermo Scientific, Dreieich) operated under Tune 2.2 using HCD fragmentation, with a normalized Collision Energy of 25%. In order to avoid repeating the sequencing of peptides dynamic exclusion was set to 60 seconds.

3.5.5 MS raw data processing

Raw data processing of HeLa P4 SILAC samples

MaxQuant software (version 1.0.13.13) and the Mascot search engine (version 2.3.2) were used for analysis of raw MS files from the LTQ- Orbitrap Velos. Quant.exe module of MaxQuant generated the peak lists were searched against the International Protein Index human protein database (version 3.86, containing 91,695 entries) supplemented with 179 common contaminants (e.g. keratins, serum albumin) and concatenated with the reverse sequences of all entries. Database (Mascot) search parameters were set as: cysteine carbamidomethylation was as a fixed modification, whereas methionine oxidation and N-terminal protein acetylation were as variable modifications; tryptic specificity with no proline restriction and up to two missed cleavages was set. The MS survey scans and MS/MS mass tolerance was set 7 ppm and 0.5 Da respectively. A minimal length of five amino acids was considered for identification. The false discovery rate was set to 1% at both the peptide and the protein level. For identification and quantification a posterior error probability (PEP) of peptides was required to be at maximum 0.05. Re-quantify was enabled, and " keep low scoring versions of identified peptides" was disabled. A minimum ratio count of two for each protein was required for Quantification of SILAC pairs by considering unique and razor peptides

Raw data processing of *Xenopus Laevis* samples

Andromeda incorporated MaxQuant software (version 1.3.0.5) were used for analysis of raw files of *X. laevis* MS experiments (Cox and Mann, 2008). Data were searched against three databases at the same time, *X. laevis* NCBI non-redundant (download date: Nov 2013), *X. tropicalis* (download date: Nov 2013) and Uniprot-reviewed *X. laevis* (download date: Sep 2013) to maximize the protein identification. Protein databases were supplemented with UPS2 standard protein sequences and with 179 common contaminants (e.g. keratins, serum albumin) and concatenated with the reverse sequences of all entries. Database search parameters were same as SILAC sample analysis as above. For all *X. laevis* experiments a minimal length of six amino acids was considered for identification. For in-depth proteomics, MS survey scans and MS/MS mass tolerance was set 7 ppm and 20 ppm respectively. For quantification, iBAQ algorithms were enabled. For exportin affinity chromatography experiments, MS survey scans and MS/MS mass tolerance was set 7 ppm and 20 ppm respectively. For quantification, label-free quantification algorithms were enabled.

3.6 Data analysis and statistics

3.6.1 Data analysis and Interpretation of MaxQuant outputs and visualization of data

Data analysis of HeLa P4 SILAC samples

MaxQuant output files, tab-delimited text files, were further processed with Excel and R with in-house written scripts (Team, 2010). All " Reverse" and " Contaminant" entries were excluded from further analysis. SILAC experiments were performed in two mode, forward and reverse, where at forward experiment heavy cells were treated with LMB, and reverse experiment light cell treated with LMB. This provides high level of confidence and avoids false positives due to the experimental workflow. Totally, four biological replicates of cytosolic and nuclear fractions were performed and each biological replicate was analyzed twice. For total HeLa cell extract only two biological replicates were done with two technical replicates each. Proteins having adversely in forward and reverse labeling experiments in terms of quantification were excluded from the further analysis. P-value (Significance B, calculated by MaxQuant) was set as a main criterion for data interpretation. This prevents the arbitrary or empirical setting of fold-change cut-off values. For proteins with p values ≤ 0.01 , changes were considered significant. Results were plotted using R (Team, 2010) and editing of figures were done Adobe Creative Suite CS5.

Raw data processing of *X. laevis* samples

As for HeLa SILAC samples, MaxQuant output files, tab-delimited text files, were further processed with Excel and R (Team, 2010) with in-house written scripts. All “Reverse” and “Contaminant” entries were excluded from further analysis. iBAQ intensities were used for calculation of absolute protein amounts. In order to avoid quantification biases coming from common peptides of UPS2 standard and *X. laevis* proteins, only unique peptides of UPS2 were considered for quantification. A linear regression curve of absolute amounts of UPS2 standard were plotted against iBAQ intensities. By using regression curve equation, measured iBAQ intensities of *X. laevis* proteins are converted to absolute protein amounts. Results were plotted using R (Team, 2010) and editing of figures were done Adobe Creative Suite CS5.

3.6.2 Gene ontology (GO) analysis

Gene Ontology enrichment analysis was done using DAVID (Huang et al., 2009).

3.6.3 Protein-protein interaction analysis

Protein interaction network analysis was performed using interaction data from the STRING database (Szklarczyk et al., 2011) in which only high-confidence (score < 0.7) interactions were represented in the network. Protein-protein interaction figures were plotted by using Cytoscape (Shannon et al., 2003), editing of figures were done Adobe Creative Suite CS5.

3.6.4 KEGG orthology assignment and pathway mapping

For bioinformatics analysis, identified *Xenopus* proteins were initially mapped human orthologs with KEGG orthology with in-house written Python script (Cock et al., 2009). Koray Kirli wrote Python script from the Department of Cellular Logistics, Max Planck Institute of Biophysical Chemistry. Human orthologs of *X. laevis* proteins then were mapped KEGG pathway and BRITe. Extensive manual interpretation was done to remove allelic variants of proteins and partial sequenced protein variant.

4 Results

4.1 Establishment of an approach to detect *in vivo* localization changes for potential CRM1 cargo identification

Up to now, more than 100 structurally and functionally diverse CRM1 cargoes have been documented in the literature. The majority of such cargoes are recognized via short peptides stretches, which are called nuclear export signal (NES) (Fischer et al., 1995; Wen et al., 1995). Even though many proteins bear such consensus sequence, it might be buried inside of the hydrophobic core of the protein, therefore it might not be functional NES. Hence, *in silico* identification of CRM1 cargoes is still difficult. An antifungal antibiotic, Leptomycin B (LMB), was frequently used to identify potential CRM1 cargoes. LMB covalently modifies the cysteine residue in the NES binding pocket of the CRM1 (Monecke et al., 2009). This modification prevents formation of export complexes and leads to nuclear accumulation of potential CRM1 cargoes. Using this antibiotic, we set out to establish an approach to detect *in vivo* localization changes between the nucleus and the cytoplasm with quantitative mass spectrometry and aim to identify novel CRM1 cargoes.

4.1.1 Preparation of subcellular fractions for MS-based experiment

The initial step for establishment of the approach was to obtain pure cytosolic and nuclear fraction. Commonly used strategies to isolate subcellular compartments start with physical disruption of cell membrane either with homogenization or with detergent treatment. It is then followed by centrifugation based approaches and/or affinity enrichment to obtain desired compartment. Many different protocols are described in the literature for variety of purposes (for review see Dreger 2003). Each protocol serves for different aims and experimental designs; thereby each of them has own advantages and disadvantages. A fractionation protocol, which was established for mammalian cell lines, was chosen (Holden & Horton 2009). It is based on sequential detergent treatment to isolate four different fractions, soluble cytosolic, organellar, nuclear and insoluble membrane enriched fraction. First, cells were treated with digitonin to permeabilize the plasma membrane and then to release the cytosolic content. Digitonin forms pores in the membrane via solubilization of cholesterol and other β -hydroxysterols (Mooney, 1988). At low digitonin concentration, cholesterol rich plasma membrane can be solubilized, whereas, intracellular membranes with low cholesterol content are not affected. Upon digitonin treatment, supernatant was collected as a cytosol-enriched fraction. Next, with mild Nonidet P-40

treatment, extract of membrane bound organelles, such as Golgi, mitochondria, and some nuclear luminal proteins were obtained as an organelle enriched fraction. The resulting pellet, predominantly nuclear fraction, was solubilized with RIPA buffer containing sodium deoxycholate to release the content of the nucleus as a soluble nuclear fraction.

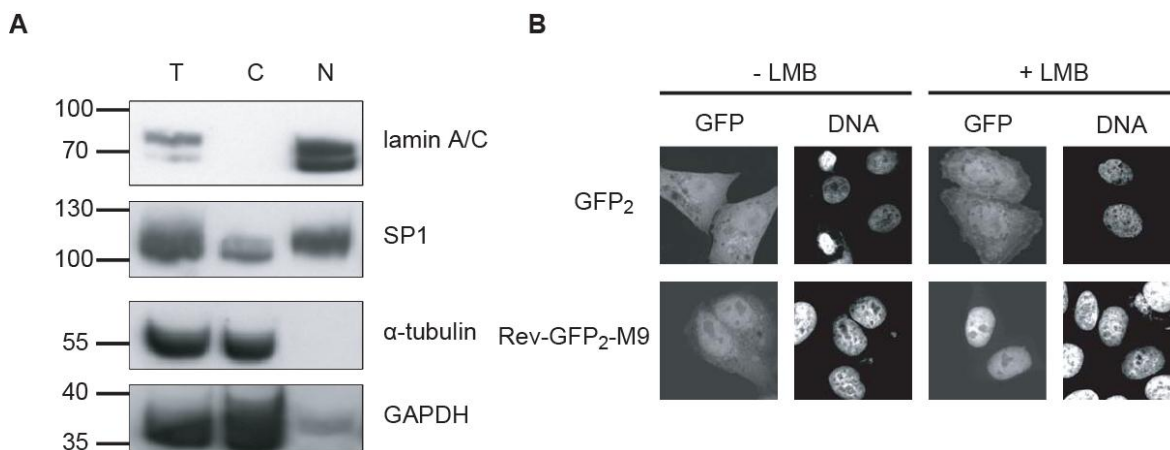


Figure 4.1 Quality control experiments.

A. Immunoblot analysis of subcellular fractions (T refers to total; C, cytosolic; N, nuclear). Indicated fractions were analyzed via SDS-PAGE, followed by immunoblotting with antibodies against lamin A/C, transcription factor SP1, GAPDH, and α -tubulin. B. Fluorescence microscopy images of HeLa cells transfected with plasmids coding for GFP₂ or Rev(48–116)-(GFP)₂-M9. After 24 h, cells were treated with or without 10 nM leptomycin B (LMB) for 3 h, fixed, and analyzed. These results were published before in (Thakar et al., 2013) and the figure was adapted with permission from the publisher.

I concentrated on MS analysis of soluble cytosolic and nuclear fractions because LMB treatment would cause localization change of proteins between those fractions. The quality of the cytosolic and the nuclear fractions was verified with several marker proteins (Figure 4.1A). The cytosolic fraction was checked with α -tubulin and GAPDH, which were predominantly present in cytosolic fraction, and absent in nuclear fraction. Lamin A/C and transcription factor SP1 were used as nuclear markers. Both of them were predominantly nuclear, and almost absent from the cytosolic fraction. Overall, high quality soluble cytosolic and nuclear fractions were obtained, which is required for quantitative mass spectrometric analysis.

To validate the inhibition of CRM1 export pathway with LMB treatment, HeLa cells were transfected with the shuttling reporter construct Rev_(48–116)-(GFP)₂-M9. It has the NES from HIV-1 Rev protein (Fischer et al. 1995) and the M9 nuclear import signal of hnRNP A1 protein that is imported via transportin (Pollard et al. 1996). This shuttling construct predominantly localized in the nucleus and with some extent in the cytosol. Inhibition of CRM1 dependent nuclear export

with brief LMB treatment resulted in exclusive localization of this construct into the nucleus (Figure 4.1B). Additionally, this shuttling construct was set as a positive control to validate that *in vivo* nuclear-to-cytosolic localization changes of individual proteins can indeed be detected with MS. These experiments were performed by and figures were provided by Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen.

4.1.2 Metabolic labeling of HeLa cells

To accurately quantify the nuclear-to-cytosolic localization changes, complete incorporation of heavy amino acids into cells is crucial. When a cell line is used for first the time for a stable isotope labeling by amino acids in cell culture (SILAC) experiments, the viability of cells and the incorporation of heavy amino acids should be monitored. For metabolic labeling, cells were initially cultured in normal media containing unlabeled L-arginine and L-lysine (“light amino acids”). Then, cells were transferred to media containing heavy counterparts of these amino acids, L-arginine-U- $^{13}\text{C}_6$ and L-lysine-U- $^{13}\text{C}_6$ - $^{15}\text{N}_2$ (“heavy amino acids”). After 5 generations, total cell extract (TCE) was isolated and proteins were separated by SDS-PAGE analysis. Proteins were in-gel digested with trypsin and resulting peptides were analyzed by LC-MS.

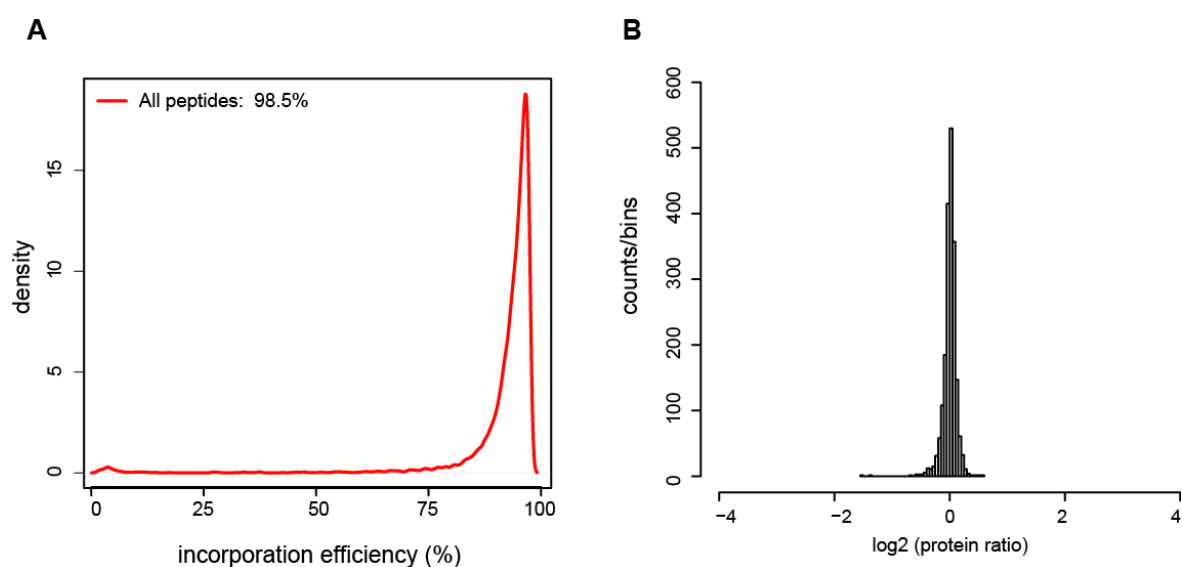


Figure 4.2 Metabolic labeling of HeLa cells.

A. Density line plot shows incorporation efficiency of heavy L-arginine and L-lysine into HeLa cells. B. Heavy and light cells were mixed 1-to-1 ratio and heavy to light protein ratios (in log₂ scale) were plotted to show distribution of protein ratios. These results were published before in (Thakar et al., 2013) and the figure was adapted with permission from the publisher.

raw data was computationally processed with the MaxQuant software (Cox et al., 2009). Incorporation efficiency was calculated based on density of all peptides containing heavy arginine and lysine relative to their light counterparts. Almost complete incorporation (~99%) of heavy amino acids was achieved after 5 cell generations (Figure 4.2A). Furthermore, to assess accuracy of the complete workflow and to validate that incorporation of heavy amino acids does not affect the proteome of HeLa P4 cells; untreated light and heavy cells were mixed in equal numbers and processed for MS analysis. Relative protein abundance changes between “light cells” and “heavy cells” were represented with histogram plot (Figure 4.2B). The x-axis represents abundance changes of proteins which is log₂ of heavy to light protein ratios. The y-axis shows protein density of protein ratios. More than 95 % of the proteins had log₂ protein ratio within standard deviation of 0.15 from 0. This tight protein ratio distribution not only shows the incorporation of heavy amino acids does not alter cellular proteome but also validates the high accuracy of the complete workflow.

4.1.3 General experimental workflow and overview of the data

HeLa cells were transfected with a plasmid containing the positive control, Rev₍₄₈₋₁₁₆₎-GFP₂-M9. Afterwards, transfected cells were split into two populations and were cultured in the media containing either “heavy” or “light” amino acids for 5 cell generations. “Light cells” were set as a control and “heavy cells” were treated with LMB. Additional to this experimental design, another experimental set was performed with swapping labeling status of LMB treatment. At this label-swap experiment “light cells” were treated with LMB and “heavy cells” were set as a control. After 3h of LMB treatment, equal number of control and LMB treated cells were mixed together. The ability of mixing of control and LMB treated cells at earliest experimental stage (before cell lysis) is one of the crucial advantages of the metabolic-based labeling technique. The control and LMB treated cells (which have different mass tags to be differentiated by MS analysis) would be processed in the same workflow. Therefore, this prevents any variation that might originate from handling errors and/or instrumental variation. Ultimately, this improves quantification accuracy; even down to 10-15 % abundance changes of proteins can be regarded as significant with SILAC based quantification (Argenzio et al., 2011). Subsequently, mixed cells were subjected to subcellular fractionation. Equal protein amount from cytosolic and nuclear fractions were run on denaturing SDS-PAGE, and then in-gel digested with trypsin. Peptides extracted from gel were subjected to LC and analyzed by MS. Afterwards, extensive data analysis and interpretation, including further bioinformatics, and microscopic characterization of potential CRM1 cargoes were performed.

To obtain general overview about the data, number of identifications from different fractions was plotted (Figure 4.3B). Almost equal number of proteins were identified from the cytosolic and the

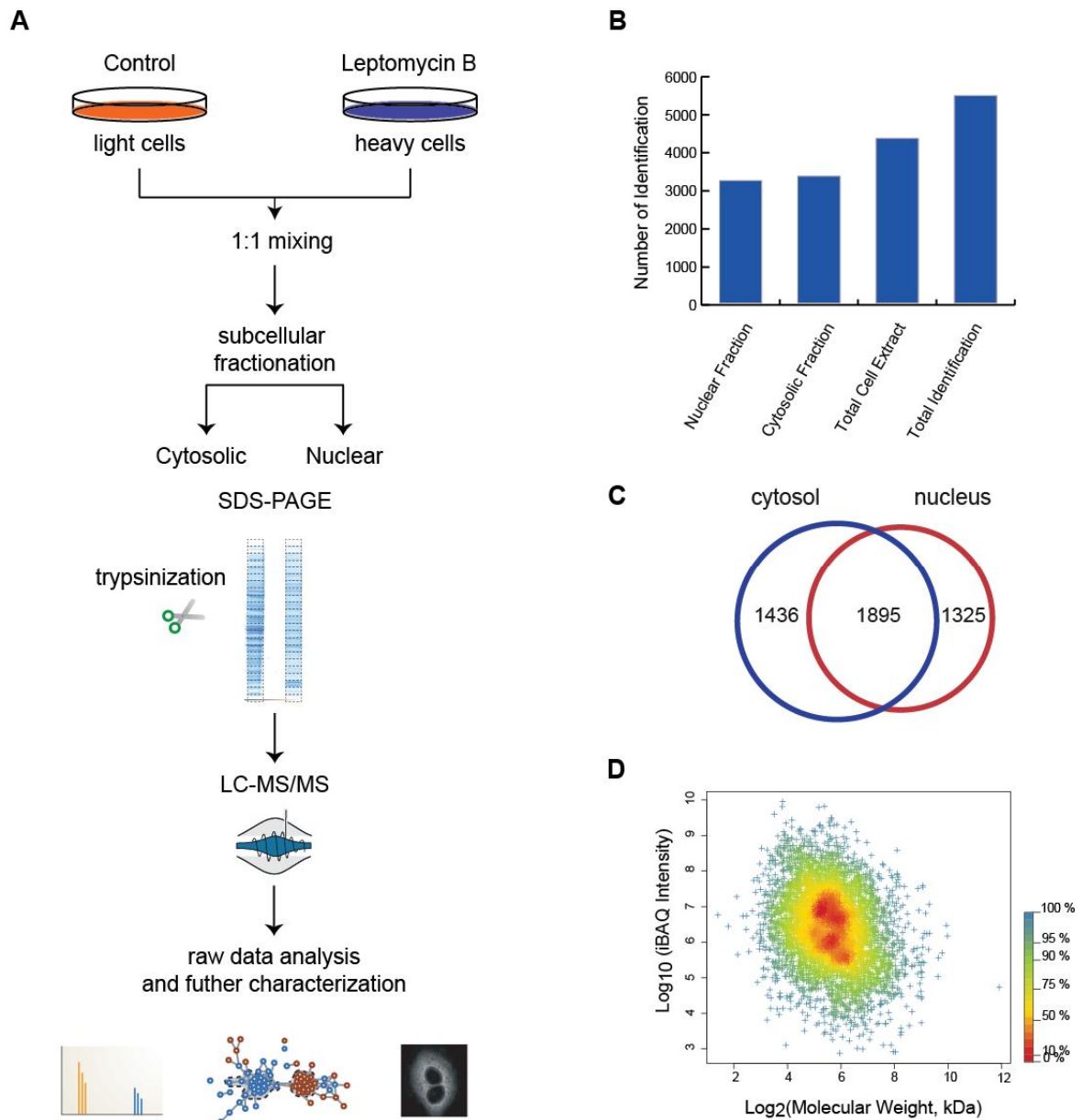


Figure 4.3 Experimental workflow and general overview.

A. Schematic representation of experimental workflow. HeLa cells were cultured with either light (Arg0 and Lys0) or heavy (Lys8 and Arg6) amino acids. Cells are mixed in equal numbers prior to subcellular fractionation. Proteins from soluble nuclear and cytosolic fractions were separated via SDS-PAGE analysis. The proteins were then in-gel digested and extracted peptides analyzed by LC-MS/MS. Generated raw data were processed with MaxQuant and further bioinformatics and microscopic characterization of data were performed. B. Protein identification across the different fractions. C. Venn diagrams of common and unique protein in cytosol and nucleus. D. Density scatterplot showing of the iBAQ intensities versus molecular weight of proteins. The color code indicates the percentage of points that are included in a region of a specific color. These results were published before in (Thakar et al., 2013) and the figure was adapted with permission from the publisher.

nuclear fractions, ~3300 and ~3200 proteins, respectively. The number of the identifications in total cell extract was ~4300 proteins and from all fractions ~5500 proteins were identified. Overall, more than half of the previously published HeLa proteome was covered (Nagaraj et al., 2011). The Venn diagram shows comparison between the cytosolic and the nuclear fractions, approximately 1400 proteins were unique to each fraction and ~1900 proteins were detected in both fractions (Figure 4.3C). Further, I wondered whether in-depth analysis of the cytosolic and the nuclear fractions cover wide range of protein expression profile. A label-free quantification approach, intensity based absolute quantification (iBAQ) (Schwanhäusser et al. 2011), was performed to estimate approximate absolute abundance of proteins. Log₁₀ of iBAQ intensities was plotted against log₂ of molecular mass of proteins (kDa) where color code represents density of the proteins (Figure 4.3D). A dynamic range of ~6 orders of magnitude in abundance was covered. This is close to the highest dynamic range detected so far in the HeLa cell proteome (Nagaraj et al., 2011). There was a tendency of low molecular weight proteins being more abundant than the high molecular weight proteins consistent with previous studies (Kulak et al., 2014; Marguerat et al., 2012). Overall, wide dynamic range in protein abundance was covered and the approach appeared to be sensitive for low abundant proteins as well.

4.2 Effect of leptomycin B on HeLa proteome

It is well established that LMB inhibits CRM1 mediated export and leads to mis-localization of proteins between the cytosol and the nucleus. In order to understand effect of the LMB in global scale, I decided to perform quantitative total proteome analysis of LMB treated HeLa cells. More specifically, total proteome analysis is necessary to assess the secondary effect of LMB. I wanted to rule out the possibility of accumulated and depleted proteins in the nuclear and/or cytosolic fractions are due to the enrichment or depletion in the total cell extract (TCE). Figure 4.4A shows the quantitative results of total cell proteome analysis after LMB treatment. This type of representation is called "Christmas tree" where the y-axis is log₁₀ of protein intensity, which is sum of all peptide intensities of a protein, and x-axis is log₂ of "heavy" to "light" protein ratio. "Heavy" to "light" protein ratio shows protein abundance changes after LMB treatment. Proteins are color coded based on their calculated p-values (see below for details). The proteins that were not affected in abundance after LMB treatment had log₂ of protein ratio around 0, which were represented as grey triangles. Proteins depleted in TCE (at left side of the figure) had log₂ of protein ratio below 0, and enriched proteins (at the right side of the figure) had log₂ of protein ratio above 0.

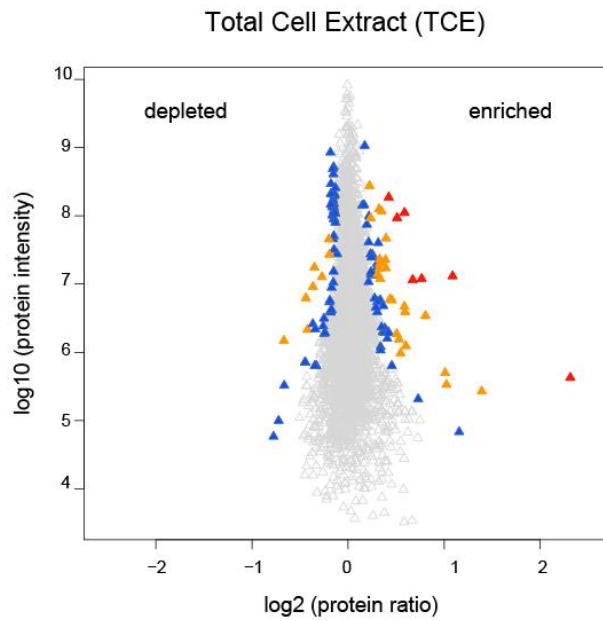
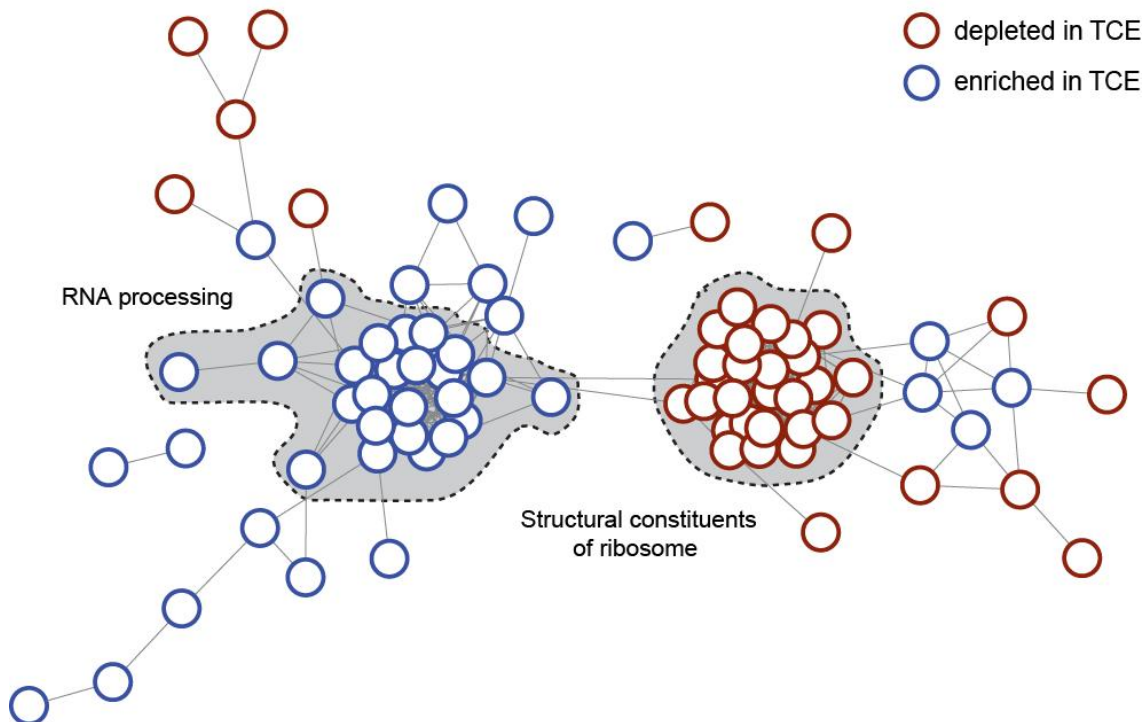
A**B**

Figure 4.4 Effect of Leptomycin B on cellular proteome.

A. Scatterplots of quantified proteins in total cell extract after LMB treatment. Proteins are colored according to significance B; gray triangles signify p values > 0.01 , blue < 0.01 , yellow < 0.0001 , and red $< 1 \times 10E-11$. B. Protein-protein association network of significantly altered proteins after LMB treatment. Red circles represent depleted proteins, blue ones enriched proteins. Grey shaded area shows protein involved in cellular process according to GO analysis. See Appendix 2 for list of the proteins. These results were published before in (Thakar et al.,2013) and the figure was adapted with permission from the publisher.

There is no consensus on a cut-off protein ratio to consider a protein “significantly changed” among the proteomics field. In fact, it is not rational to assume that a single agreed cut-off protein ratio could be applicable to all high throughput proteomics studies. Each study should be considered and judged by its own context. Therefore, I followed a universal approach (a statistical measure) to generate a list of significantly changed proteins. A statistical test (Significance B) performed by MaxQuant software (Cox and Mann, 2008) was applied to find significantly altered proteins. Principally, this statistical test identifies proteins lying “comparatively” far from the bulk of the protein distribution by considering variance of the all protein ratio distribution (Cox and Mann, 2008). This statistical test reports a p-value and a p-value of 0.01 means that 99% probability that reported protein ratio is significantly changed. Besides, it takes into consideration that quantification accuracy of an abundant protein is more accurate than low abundant proteins. Accordingly, even small changes in protein abundance down to 10-15% could be significant for high abundant proteins since they can be easily separated from the bulk (Argenzio et al., 2011). Three criteria were required to consider a protein was significantly changed in their abundance after LMB treatment. First, a protein was identified with at least one unique peptide, and was quantified with at least with 2 quantification events. Next, upper threshold p-value (calculated by MaxQuant) of 0.01 was required. Last, proteins behaving inconsistently in the label-swap experiments in terms of quantification were excluded from further analysis. Proteins matching all three criteria were considered significantly altered in their abundance after LMB treatment. This data filtering was applied for quantitative TCE and also for subcellular fractionation analysis of LMB treated HeLa cells.

As expected for a highly targeted drug, the effect of LMB on cellular proteome was specific as well. Only a small portion of the identified proteins showed enrichment or depletion (3% of all quantified proteins). 70 proteins were enriched and 58 proteins were depleted in TCE of LMB treated cells (Figure 4.4A and Appendix 2 for list of the proteins). Apart from several proteins, majority of them showed less than 75 % change in their abundance.

To assess whether there is any functional association between enriched and depleted proteins in the TCE, I performed Gene ontology (GO) analysis. Interestingly, for depleted proteins, GO analysis revealed that almost half of the proteins were part of the structural constituent of ribosomes and/or involved in translation. Among 26 depleted proteins of structural constitute of ribosome, 25 of them were belong to 60S ribosomal subunit (out of 48 identified components). On the other hand, there was only one from the 40S ribosomal subunit (out of 33 identified components). Likewise, enriched proteins showed a strong association between them. Almost

40% of them were involved in various steps of RNA processing activities. Next, a protein-protein association network from STRING database revealed more informative depiction of depleted and enriched proteins. This association network was based on relationships inferred from experimental evidence, literature, text mining and curated databases (Figure 4.4B). Each protein was represented as an individual node (circle), whereas association of proteins was connected with edges (lines). Color-coding shows either a protein was depleted (red nodes) or enriched (blue nodes) after LMB treatment. The distance or the path between nodes represents how closely these proteins associated with each other. A visual inspection of the association network reveals that majority of the enriched and depleted proteins closely clustered with each other. Enriched proteins concentrated on left side of the network and depleted proteins at right side of the network. More interestingly, two tight clusters are present in each group, which are shown as grey shaded area. Grey shaded area (on the right side of the network) containing red nodes is composed of proteins belong to GO category of the structural constitute of ribosome. Similarly, the other grey shaded (on the left side of the network) area comprises proteins belong to GO term of RNA processing.

4.3 Identification of potential CRM1 cargoes with leptomycin B treatment

4.3.1 LC-MS/MS analysis of cytosolic and nuclear fractions

Subsequent analysis of TCE; I proceeded to analyze the cytosolic and the nuclear fractions. Subcellular fractionation analysis led to identification of approximately 4600 proteins with an overlap of 1900 proteins (Figure 4.3C). It is expected that LMB treatment would cause depletion of proteins in the cytosol and accumulation of proteins in the nucleus. Therefore, either accumulation or depletion of proteins upon LMB treatment would be reflected into abundance changes of proteins in the nucleus and in the cytosol. Successfully, a group of proteins showed accumulation and depletion in the nuclear and the cytosolic fractions, respectively (Figure 4.6A and 4.6B). However, the cytosolic and the nuclear proteomes showed rather small changes in abundance. I applied stringent criteria to create a significant list of proteins that showed localization change (as explained at section 4.2). This resulted in a list of 138 proteins (see Appendix 1) that 84 proteins accumulated in the cytosolic fraction and 59 depleted in the nuclear fraction, with 5 overlap proteins (Figure 4.6C).

Ideally upon LMB treatment a protein which is accumulated in the nucleus simultaneously should have depleted in the cytosol. However the overlap between cytosol-depleted and nuclear-accumulated is rather low, only 5 proteins showed this "ideal" behavior (Figure 4.6C). These

proteins are glutamate-rich WD repeat-containing protein 1, eIF6, alpha-globin transcription factor CP2, programmed cell death protein 2-like protein (PDCD2L), and the 60S ribosomal export protein NMD3. Among those, eIF6 (Biswas et al., 2011) and NMD3 (Ho et al., 2000) were already described as CRM1 cargoes. It is worth to mention that, in this list there are well-established CRM1 cargoes that do not exhibit this "ideal" behavior. They were either depleted in the cytosol or accumulated in the nucleus. Alternatively, they were not identified in other fraction. A representative example is the positive control, Rev₍₄₈₋₁₁₆₎-GFP₂-M9. It showed up with high significance among the cytosol-depleted list. However, it could not be quantified in the nuclear fraction. Therefore, this observation underlines that for comprehensive identification of potential CRM1 cargoes, both the nuclear and the cytosolic fraction should be analyzed.

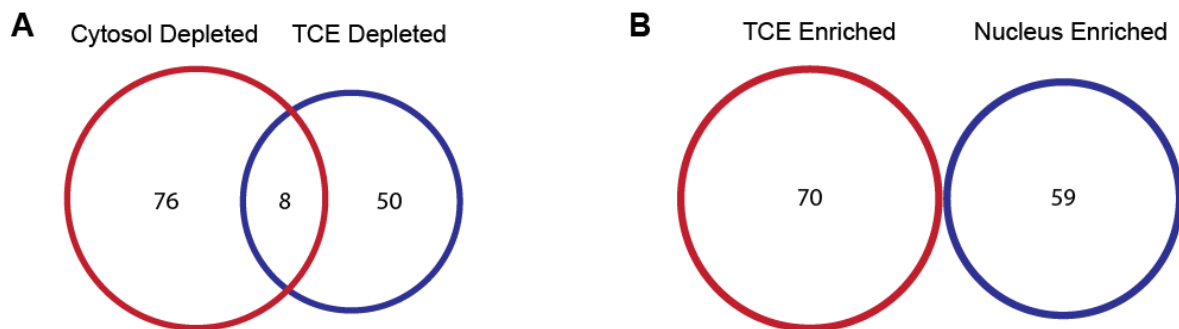


Figure 4.5 Comparison of significant lists of different experiments.

A. The Venn diagram depicts overlap between depleted protein in cytosol and total cell extract. B. The Venn diagram shows lack of overlap between list of protein in TCE enriched and nucleus enriched. These results were published before in (Thakar et al., 2013) and the figure was adapted with permission from the publisher.

As a first step to interpret data, I wondered how these 138 mis-localized proteins had behaved in LMB treated TCE analysis. There were eight proteins (AAMP, DDX3X, eIF6, GRWD1, PTPN23, RPL22L1, RRM2 and SQSTM1) out of 84 that were depleted from the cytosol as well in from the TCE (Figure 4.5A). Among those eight proteins, two of them were previously characterized as CRM1 cargoes, eIF6 and DDX3X (Biswas et al., 2011; Yedavalli et al., 2004). From these 8 proteins, one candidate was further investigated, sequestosome 1 (SQSTM1) protein. A GFP tagged version of this protein showed nuclear accumulation upon LMB treatment (Figure 4.7B). It is plausible that depletion of a protein in TCE could be result of reduction in the half-life of the protein upon mis-localization. However, protein stability under LMB treatment was not primary focus of this study and was not further investigated. The remaining 76 proteins either were not affected in the

TCE, or were not identified. Similarly, none of the accumulated proteins in the nuclear fraction showed up as enriched in the TCE (Figure 4.5B). Additionally, there was not a protein that was both depleted in nuclear fraction and in the TCE. Therefore, this indicates that enrichment or depletion of proteins in the TCE upon LMB treatment cannot account for many of the CRM1 cargoes as found in the subcellular fractions.

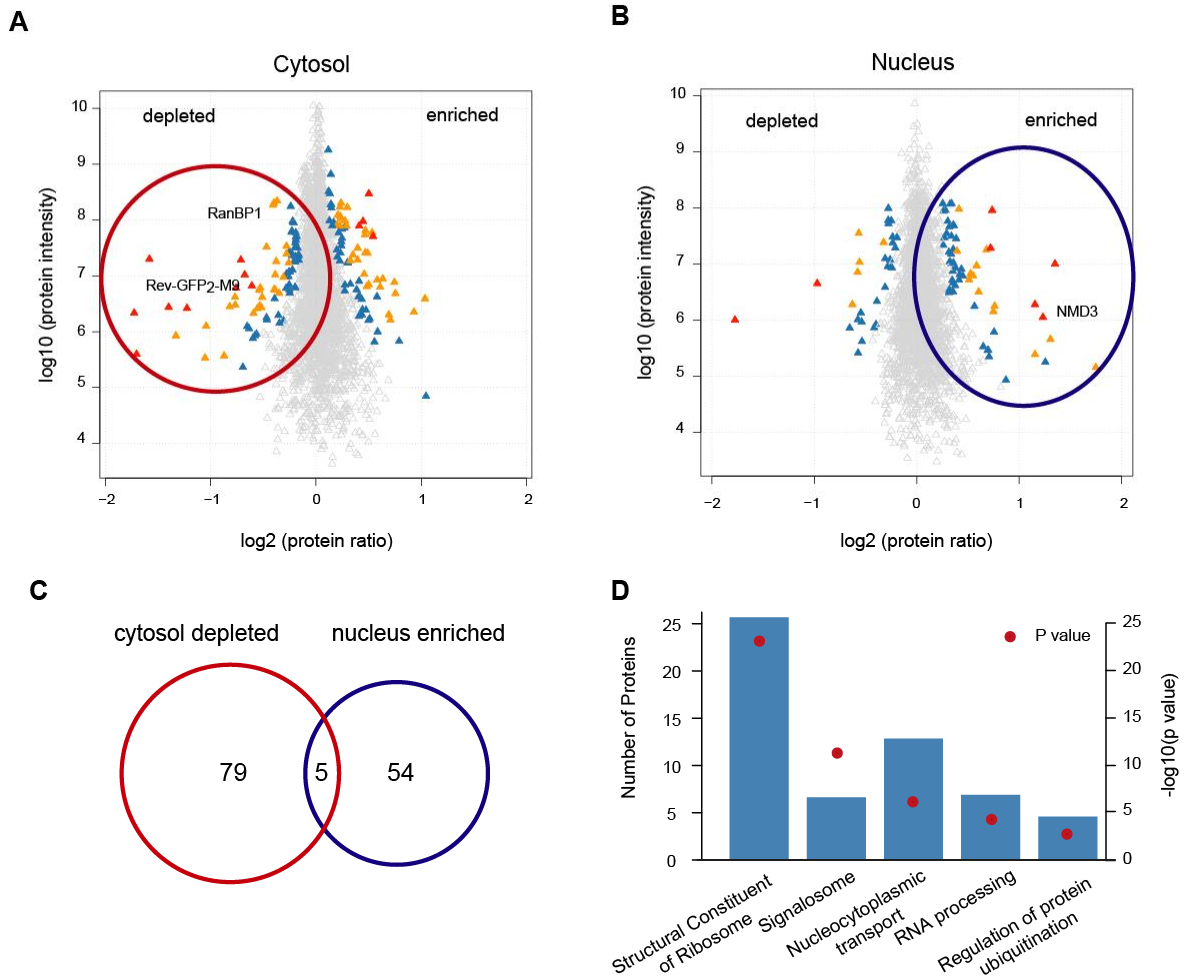


Figure 4.6 Selective inhibition of the CRM1 export pathway leads to changes in the subcellular localization of proteins.

A. and B. shows scatterplots of quantified proteins in cytosolic or nuclear fractions after LMB treatment, respectively. Coloring scheme is based on the significance of proteins; gray triangles signify p values > 0.01 , blue < 0.01 , yellow < 0.0001 , and red $< 1 \times 10E-11$. C. Venn diagrams shows overlap between cytosol depleted and nucleus enriched protein in A. and B. See Appendix 1 for details. D. Histogram plot shows significantly enriched GO terms among potential CRM1 cargoes. Bar represented number of proteins in terms, and red line shows p -values. These results were published before in (Thakar et al.,2013) and the figure was adapted with permission from the publisher.

Before proceeding to characterize potential CRM1 cargoes, I concentrated on bioinformatics dissection of the data with Gene Ontology (GO) analysis. Identified potential CRM1 cargoes were

mapped to more than 50 GO terms. Afterwards, I manually inspected all remaining terms to represent the most relevant and non-redundant term. I was able to scale down 50 GO terms into five non-redundant and informative terms (Figure 4.6D). Results were illustrated with a bar plot representing number of proteins in GO term and red dot shows corresponding p-value of the over-represented term. The highest significant GO term was the structural constitutes of ribosome with comprising approximately 20% of the potential CRM1 cargoes (for details see Appendix 1). Particularly, 25 out of 26 proteins were belonging to 60S ribosomal subunits. Previously, it was shown that CRM1 exports 60S ribosomal subunit to the cytosol with the adapter protein NMD3 (Ho et al., 2000). NMD3 was detected as a nuclear accumulated protein upon LMB treatment in this study (Figure 4.6B). Even though, the 40S ribosomal subunit was shown to be export with CRM1 (Moy and Silver, 2002), there was only one protein, RPS3A, from 40S ribosomal subunit.

The second highest category is the COP9 signalsome, which is a conserved multi-subunit protein complex involves in variety of biological processes, such as ubiquitin-proteasome pathway, DNA-damage response, and cell-cycle control (Kato and Yoneda-Kato, 2009). It is composed of 8 subunits in higher eukaryotes and seven out of eight subunits were depleted in the cytosolic fraction (see Appendix 1). This suggests that CRM1 exports the complete COP9 signalosome complex. This observation is supported by the fact that a component of COP9 signalosome, Jab1/CSN5, has already been described as being LMB sensitive (Tomoda et al. 2002). Other over-represented terms were, nucleocytoplasmic transport, RNA processing, and regulation of protein ubiquitylation. It was surprising that term nucleocytoplasmic transport was over-represented. It is composed of CRM1 itself, importin alpha-1,2,3,4,7, importin 8, importin-4 and transportin-1. This could be due to the secondary effect of the LMB. The term RNA processing is a fairly general description, as the proteins in this term represent variety of functions in RNA metabolism. Several of them are RNA binding proteins PABPC1, PABPC4, helicase DDX20 and interestingly a RNA polymerase II component, POLR2B and RNA polymerase II-associated protein 1 (RPAP1). Last the GO term was regulation protein ubiquitylation, which contains four proteins from anaphase promoting complex (APC/C). It is an E3 ubiquitin ligase involving progression of cell cycle. Given with these observations, presence of well-known CRM1 cargoes (15 cargoes, see Appendix 1) and expected indirect CRM1 binders, such as 60S ribosomal subunits, first it validates that the approach is suitable for detection of dynamic nuclear-to-cytosolic localization changes with MS. Next, it strongly shows that these mis-localized proteins are CRM1 export cargoes.

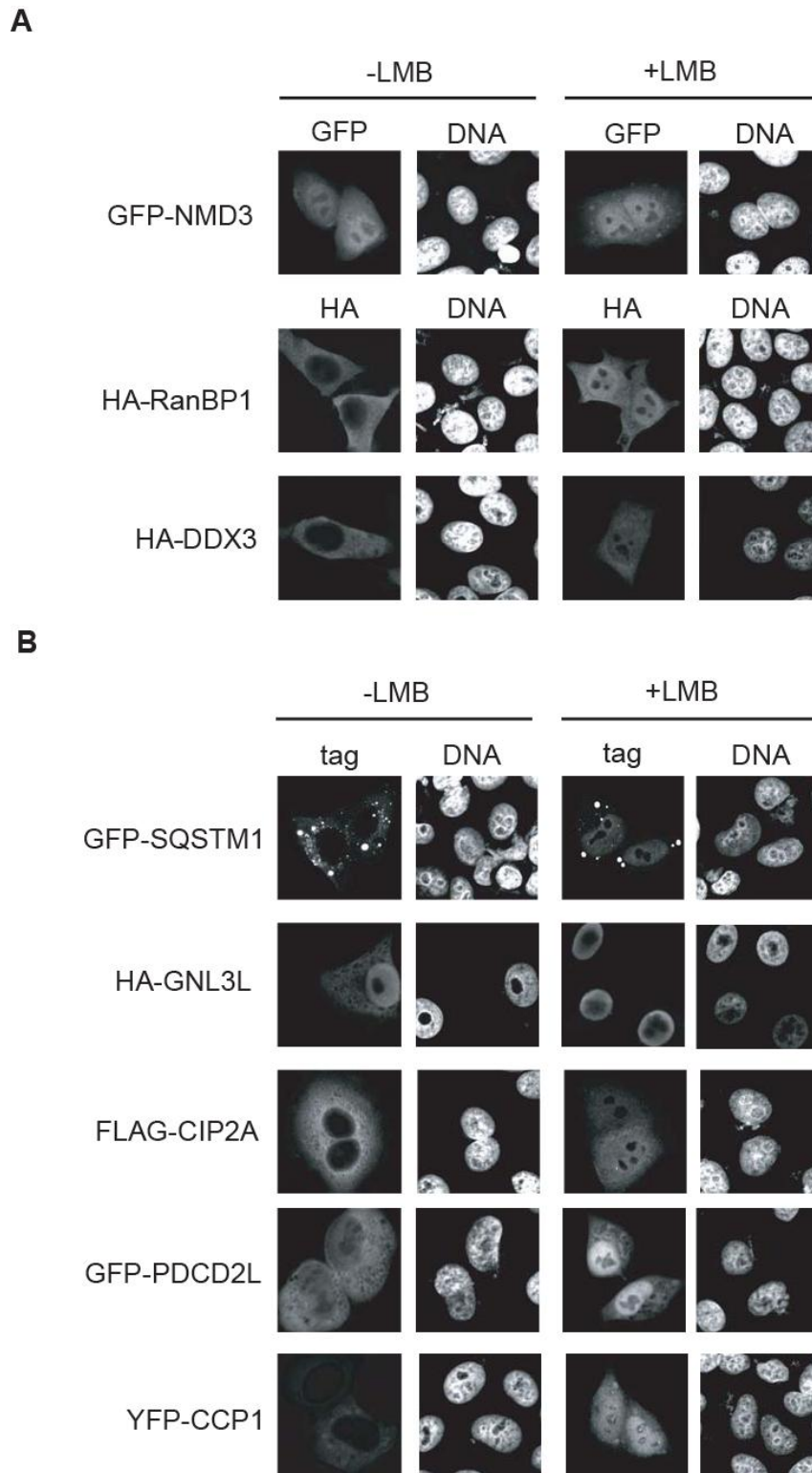


Figure 4.7 Validation of several known and novel CRM1 cargoes.

A. Fluorescence microscopy images of HeLa cells which were transfected with plasmids coding for GFP-NMD3, HA-RanBP1, HA-DDX3 (previously known CRM1 cargoes). B. GFP-SQSTM1, HA-GNL3L, FLAG-CIP2A, GFP-PDCD2L, or YFP-CCP1 (novel candidates), as indicated. For both A. and B. After 24 h of transfection, cells were treated with or without 10 nM LMB for 3 h. Detected were done via indirect immunofluorescence using anti-HA or anti-FLAG antibodies or directly via the GFP/YFP-tag. These results were published before in (Thakar et al., 2013) and the figure was adapted with permission from the publisher.

4.3.2 Validation of known and novel CRM1 cargoes

The observed potential CRM1 cargo list contains 14 known CRM1 cargoes (see Appendix 1). Among those proteins, epitope-tagged version of 3 known CRM1 cargoes NMD3, RanBP1, and DDX3 were confirmed via microscopic analysis. Indeed, HA-RanBP1 and HA-DDX3 behaved as predominant cytoplasmic proteins and then accumulated in the nucleus upon treatment with LMB (Figure 4.7A). On the other hand, GFP-NMD3 was predominantly nuclear without LMB treatment, but there was a slight shift towards nucleus upon LMB treatment (Figure 4.7A). Afterwards, novel CRM1 cargoes were investigated by microscopic analysis. In order to show the validity and reliability of the approach, proteins with varying levels of p-values (from upper, middle and lower part of the list) were selected for thorough microscopic analysis. Dr. Ketan Thakar from the Department of Biochemistry I, Faculty of Medicine, Georg-August-University of Göttingen. Tagged versions of the selected proteins (GFP-SQST1, HA-GNL3L, FLAG-CIP2A, GFP-PDCD2L, and YFP-CCP1) were expressed in HeLa cells, and localization of these constructs was investigated with and without LMB treatment. A protein with very low p value (see Appendix 1) from potential CRM1 cargo list was, Sequestosome 1 (SQSTM1/p62). SQSTM1 was initially identified as a protein that binds to polyubiquitin chains (Shin, 1998) and it was reported to be involved in development of Parkinson's disease, Alzheimer's disease, and Paget's disease, among others (Geetha et al., 2012). GFP-tagged version of this protein shows cytoplasmic localization with cytoplasmic speckles. LMB treatment caused localization towards the nucleus (Figure 4.7B). Another candidate with very low p-value was nucleotide binding protein-like 3-like protein (GNL3L). Previously it was characterized as a nucleolar protein (Meng et al., 2007). However, at Human Proteome Atlas (a consortium that generates an antibody for each proteins and documents their subcellular localization), apart from main nucleolar localization, it has been detected in the cytoplasm in some cell lines (Uhlén et al., 2005). We, as well, detected some portion of this protein in cytoplasm (Figure 4.7B). Upon LMB treatment, HA-GNL3L showed exclusive nuclear localization. Among the proteins with moderate p-value (Figure 4.6A, orange triangles), CIP2A (cancerous inhibitor of PP2A) was selected. It is direct interaction partner of c-Myc and it promotes c-Myc stabilization (Junttila et al., 2007). YFP-CIP2A showed exclusive cytoplasmic localization, and LMB strongly disturbed this cytoplasmic localization towards the nucleus. Programmed cell death protein 2-like (PDCD2L), is an example of "ideally" behaving protein (both nuclear accumulation and cytosol depletion in this screen). GFP tagged PDCD2L construct showed cytoplasmic localization, but moved into nucleus, after LMB treatment. Lastly, a protein with relatively high-level p-value; Cytosolic carboxypeptidase 1 (CCP1/NNA1) was studied. Initially it was identified in motor neurons and it was related to zinc carboxypeptides (Harris et al.,

2000). Later, it was shown that it mediates deglutamylation of target proteins, such as tubulin, and it is involved in certain types of neurodegeneration (Rogowski et al., 2010). YFP-tagged CCP1 had exclusive cytoplasmic localization; LMB treatment caused more equal distribution between the nucleus and the cytoplasm.

The thorough microscopic analysis demonstrated that, all selected candidates (having varying levels of p-values) were LMB sensitive, firmly shows reliability of the screen. Altogether, the presence of many known and expected CRM1 cargoes (direct or indirect binders) and validation of the several candidates demonstrate that the approach I presented here is well suitable to detect dynamic *in vivo* nuclear-to-cytosolic changes with MS. Furthermore, this approach dramatically expanded potential CRM1 cargo atlas by providing evidence for ~80 novel proteins.

4.4 Quantitative spatial proteomics and RanGTP dependent CRM1 interactome of *Xenopus laevis* oocytes

Quantitative MS analysis of LMB treated HeLa cell fractions led to identification of many known and novel CRM1 cargoes, and several potential cargoes with varying p-values were microscopically validated. However, this is not the most comprehensive CRM1 cargo atlas since several well-established cargoes were missing. I, therefore, decided to follow alternative approaches to generate more comprehensive picture of CRM1 cargo atlas. Moreover, to make thorough interpretation of cellular compartmentation with respect to nuclear export, there was a need for localization profiling of proteins in the nucleus and in the cytosol. For this purpose, I chose a model organism, *Xenopus laevis* oocyte, in which the nuclear and the cytosolic proteome can be obtained with unmatched purity. As an alternative approach to LMB treatment, RanGTP dependent CRM1 binders were selectively enriched via affinity based chromatographic approach. Afterwards, I correlated subcellular proteome of *X. laevis* with RanGTP dependent CRM1 binders to generate CRM1 “ exportome ” with respect to cellular compartmentation.

4.4.1 Overview of experimental workflow and data analysis

Problems encountered during subcellular fractionation of HeLa cell cultures with centrifugation based approaches challenged me to obtain pure cytosolic and nuclear proteomes. Therefore, I shifted to a different model system, *Xenopus laevis* oocytes, where I acquired cytosolic and nuclear proteomes without detectable cross-contamination. Fully-grown *X. laevis* oocytes are gigantic with diameter of ~1.3 mm and oocyte's nucleus reaches up to 100.000-fold larger volume than somatic nucleus. This allows for faithful manual dissection of the oocyte to obtain pure cytoplasmic and nuclear fraction (Figure 2.2). Figure 4.8 outlines the complete workflow of spatial

proteome mapping of *X. laevis* oocyte and affinity enrichment of RanGTP dependent CRM1 interactome.

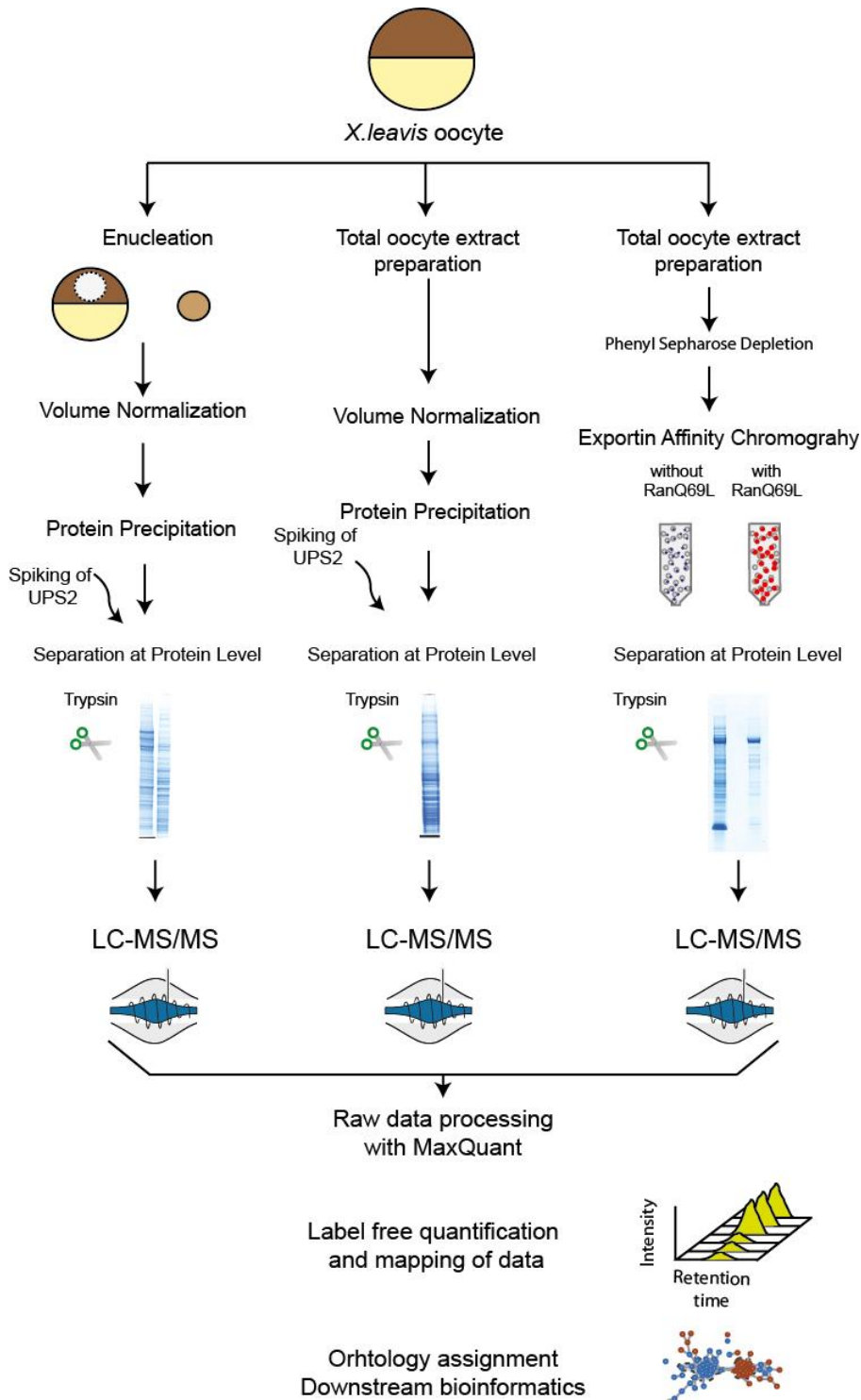


Figure 4.8 Workflow of spatial proteomics and RanGTP-dependent interactome analysis of the *Xenopus laevis* oocytes.

The complete workflow is composed of two major sections. In depth proteome analysis, affinity chromatography enrichment for CRM1 export complexes. See text for the details.

The first phase of the workflow was dedicated to in-depth proteome profiling. For this purpose, I manually isolated 60 nuclei from stage VI oocytes. I enucleated the oocytes with very fine pair of forceps with puncturing the oocyte on the animal pole (dark pigmented side) and then squeezing out the intact nucleus from the punctured site. The intact nucleus was gently but thoroughly washed to get rid of yolk and residual cytoplasmic content around the NE. The intact nucleus was transferred into absolute ethanol for further protein precipitation. Corresponding enucleated oocytes were processed to generate the cytosolic fraction. In order to directly compare the protein concentrations between the cytosolic and nuclear proteomes, I performed 1-to-10-volume normalization for the cytosolic extract (assuming volumes for yolk free-cytosol 500 nL and nucleus 50 nL). Balancing protein content of the cytosolic and nuclear proteomes not only enabled to direct protein concentration comparison but also prevents identification biases via MS. Likewise; the total oocyte extract was volume-normalized as the cytosolic extract. To obtain a rough estimate of absolute protein abundance in the fractions, Universal Protein Standards 2 (UPS2, Sigma) was spiked into all three fractions right before running them on SDS-PAGE. UPS2 is a mixture of bacterially expressed 48 human proteins spanning 6 orders of magnitude in abundance. Both UPS2 and oocyte fractions were processed in the same workflow, in gel-digested with trypsin and analyzed by LC-MS. Even though, label-based quantification approaches provide higher accuracy in quantification, they suffer from reduced identification due to the complexity generated by labeling. Therefore, a label-free quantification approach, iBAQ (Schwanhäusser et al. 2011), was utilized to calculate both relative abundance comparison between fractions and absolute protein amounts in all fractions.

The second phase of the workflow consists of affinity purification of RanGTP dependent interactome of CRM1 from *X. laevis* oocyte extract. To prepare the extract, oocytes were physically homogenized and the resulting extract was cleared with sequential low and high-speed centrifugation. During the optimization of the CRM1 affinity chromatography, I encountered enrichment of nuclear transport receptors (NTRs) and cargoes of other exportins. This could be due to the bridging effect of FG-nucleoporins that interact with immobilized CRM1 and endogenous exportins in Ran independent manner. To overcome this issue, all endogenous NTRs were depleted from the oocyte extract by phenyl sepharose matrix (Ribbeck and Görlich, 2002). For affinity-based enrichment of CRM1 export complexes, biotin tagged CRM1 was immobilized on streptavidin-agarose beads. To mimic the nuclear environment, export complexes were formed in the presence of 5 μ M RanQ69L (GTPase deficient mutant) from *X. laevis* oocyte extract.

After extensive washing, bound proteins were eluted with cold SDS sample buffer. This ensured dissociation of export complexes while the majority of biotin tagged CRM1 was preserved on the beads, thereby suppressing elution of high abundant immobilized CRM1. Elutions were run on SDS-PAGE, in-gel digested with trypsin and analyzed with LC-MS. Since metabolic labeling is not yet possible for *Xenopus*, I performed a label free approach to identify significantly enriched CRM1 interaction partners. Due to the lack of comprehensive annotation of *X. laevis* genome, identified proteins were initially mapped to human orthologs with via KEGG knowledge base. Then, I carried out following bioinformatics analysis by relying on human orthologs of *X. laevis* proteins.

4.4.2 In-depth proteome analysis of total oocyte, cytosolic and nuclear fractions

The first phase of the experimental workflow focused on characterization of the total, cytosolic and nuclear proteomes of the *Xenopus* oocyte. For this purpose, I initially, correlated the absolute abundance of UPS2 against their measured iBAQ intensities. Before performing this, peptide sequences from UPS2 were compared with theoretical peptide sequences from the *Xenopus* proteome. The common peptides were identified and individually excluded from the quantitative analysis. The measured iBAQ intensities of identified UPS2 standards were correlated with their absolute abundance with a Pearson correlation coefficient (r-value) of 0.98 and lowest detected UPS2 standard was ~0.1 fmole (Figure 4.9A). High correlation between absolute amounts and measured iBAQ intensities of UPS2 standards was achieved for the nuclear and the cytosolic fraction as well with r-values of 0.97 and 0.98, respectively (data not shown). Based on the obtained correlation equation from each fraction, absolute amounts of all identified *Xenopus* proteins were calculated. The calculated log₁₀ of absolute protein copy numbers per oocyte were decreasingly ranked and quantified UPS2 standards were depicted with yellow dots to show the quantification range (Figure 4.9B). The quantification range, the range between the highest and lowest abundant UPS2 standards, nearly covers whole total oocyte proteome (~99% of the proteome) according to available database. For the proteins outside of the quantification range, absolute amounts were calculated with data extrapolation. As a further validation, previously published individual protein concentrations (nM) (Wühr et al., 2014) either in *X. laevis* oocyte or egg were plotted against calculated protein concentration (nM) (Figure 4.9C). There was good agreement between published individual proteins concentration with UPS2 based absolute quantification, with r-value of 0.85. I was able to quantify ~6300 proteins within 6 orders of magnitude in abundance. In the total oocyte extract, protein concentration varied between 21 μM (most abundant protein, actin) to 5 pM (after extrapolation of the proteins outside of the quantification range) and median protein concentration was around at 14 nM.

The abundance values of individual proteins were used to estimate cumulative abundance of all identified proteins. The cumulative abundance reflected more drastic nature of the *X. laevis* proteome (Figure 4.9D).

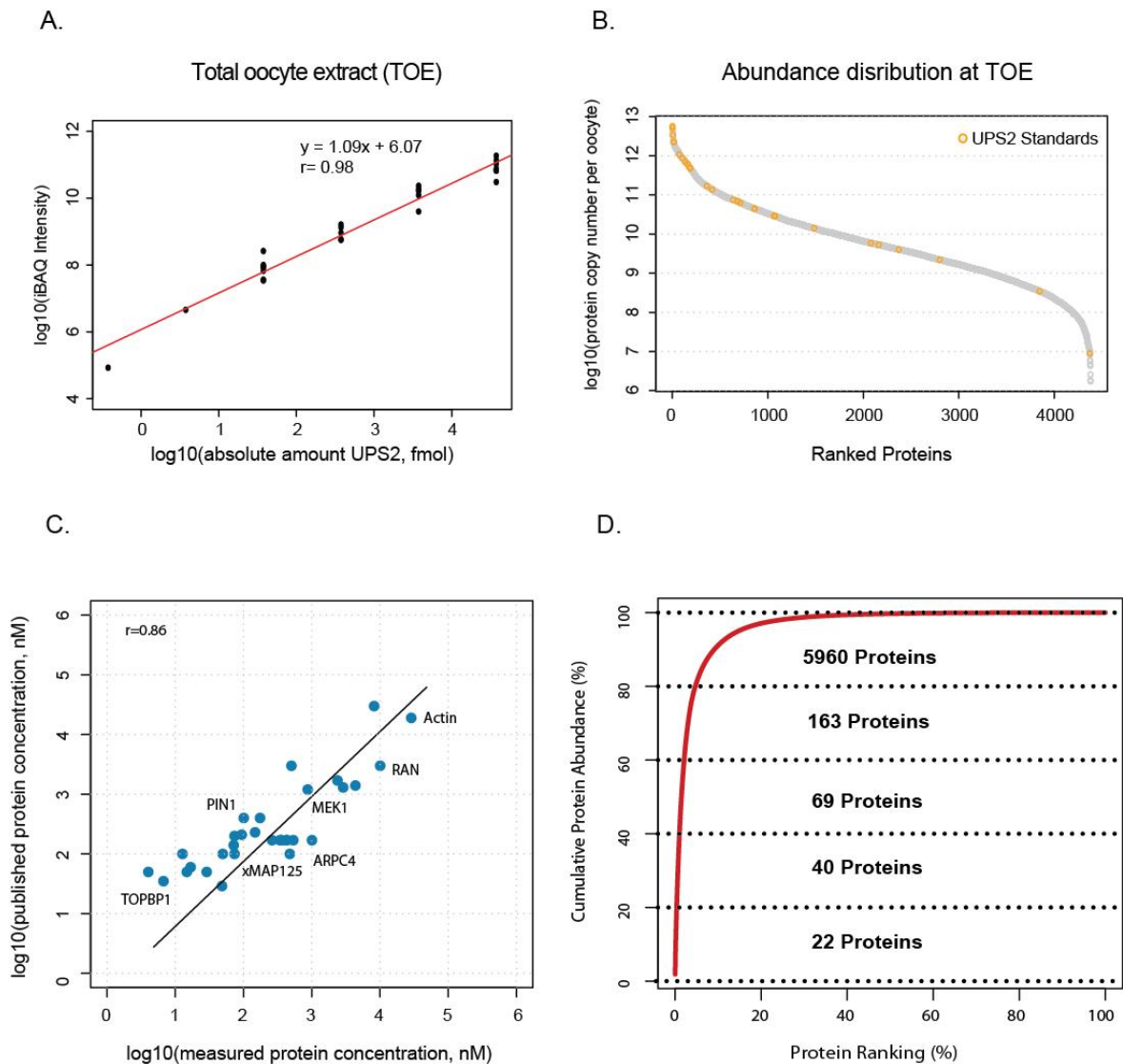


Figure 4.9 Absolute and in-depth proteome of *Xenopus laevis* oocyte.

A. The scatterplot represents regression curve of UPS2 absolute amounts against measured iBAQ intensities. B. The scatterplot of calculated absolute copy numbers per *X. laevis* oocyte proteome with spiked UPS2 proteins, where UPS2 standards were colored as yellow dots C. Scatterplot showing the correlation between previously published proteins concentration and calculated protein concentration in this study. D. Density histogram plot of cumulative abundance of all identified proteins (in all, cytosolic, nuclear and total oocyte extract) Numbers represent, how many proteins are present in each 20 % quantile.

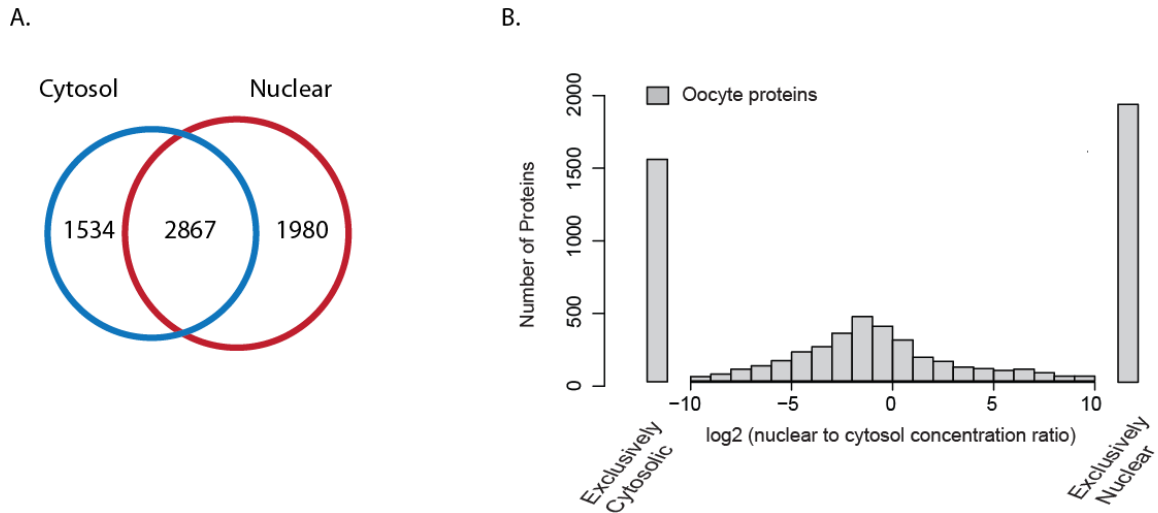


Figure 4.10 Spatial proteome of *Xenopus laevis* oocyte.

A. Venn diagram shows identification comparison between the cytosolic and the nuclear proteome of the *X. laevis* oocyte. B. Histogram plot shows \log_2 of nuclear- to-cytosolic concentration ratio distribution of quantified proteins. Compartment specific proteins were placed to outside of the ± 10 scales.

A handful of proteins constitute the majority of the protein content in the oocyte. The most abundant 22 proteins account for approximately 20 % of the protein molecules in the oocyte. On the other hand, least abundant ~6000 proteins represent protein molecules of only 20%. The most 22 abundant proteins are ribosomal proteins, glycolysis pathway members and cytoskeleton proteins, such as actin. Alone ribosomal proteins themselves constitute ~20% of the protein molecules in the oocyte. Spatial proteome analysis of *X. laevis* nuclear, cytosolic and total oocyte extract led to identification of ~4400, ~4800 and ~5500 proteins, respectively. The comparison between the cytosolic and the nuclear proteome showed that high number of proteins was compartment specific (Figure 4.10A). Approximately two-thirds of the proteins were detected in both compartments, while almost one-third of the cytosolic and the nuclear proteome were unique to the respective compartment (Figure 4.10B). The nuclear-to-cytosolic concentration ratio was calculated by iBAQ intensity values of proteins in \log_2 scale. A negative ratio represents higher cytosolic concentration, whereas positive ratio represents higher nuclear concentration. When ratios of all common proteins were plotted as frequency histogram plot, it gave a Gaussian-like distribution (Figure 4.10B). Almost all \log_2 protein ratio spans from -10 to 10. There were only 88 proteins having extreme ratios outside of ± 10 out of ~6300 proteins. Proteins showing extreme ratios and proteins not quantified in one compartment were represented outside of the scale, at the left and right side of figure (Figure 4.10B). In general, nuclear-to-cytosolic concentration distribution of proteins revealed tri-modal distribution and only a subset of

proteins (~ 500 proteins having ratios in log2 scale between -1 and 1) had equal concentration distribution across the compartments. Approximately 1600 and 2000 proteins were specific either to the cytosol or the nucleus, respectively. Most abundant proteins in the total oocyte extract mainly belong to common biological processes such as ribosomal proteins, metabolic pathway members and cytoskeleton proteins.

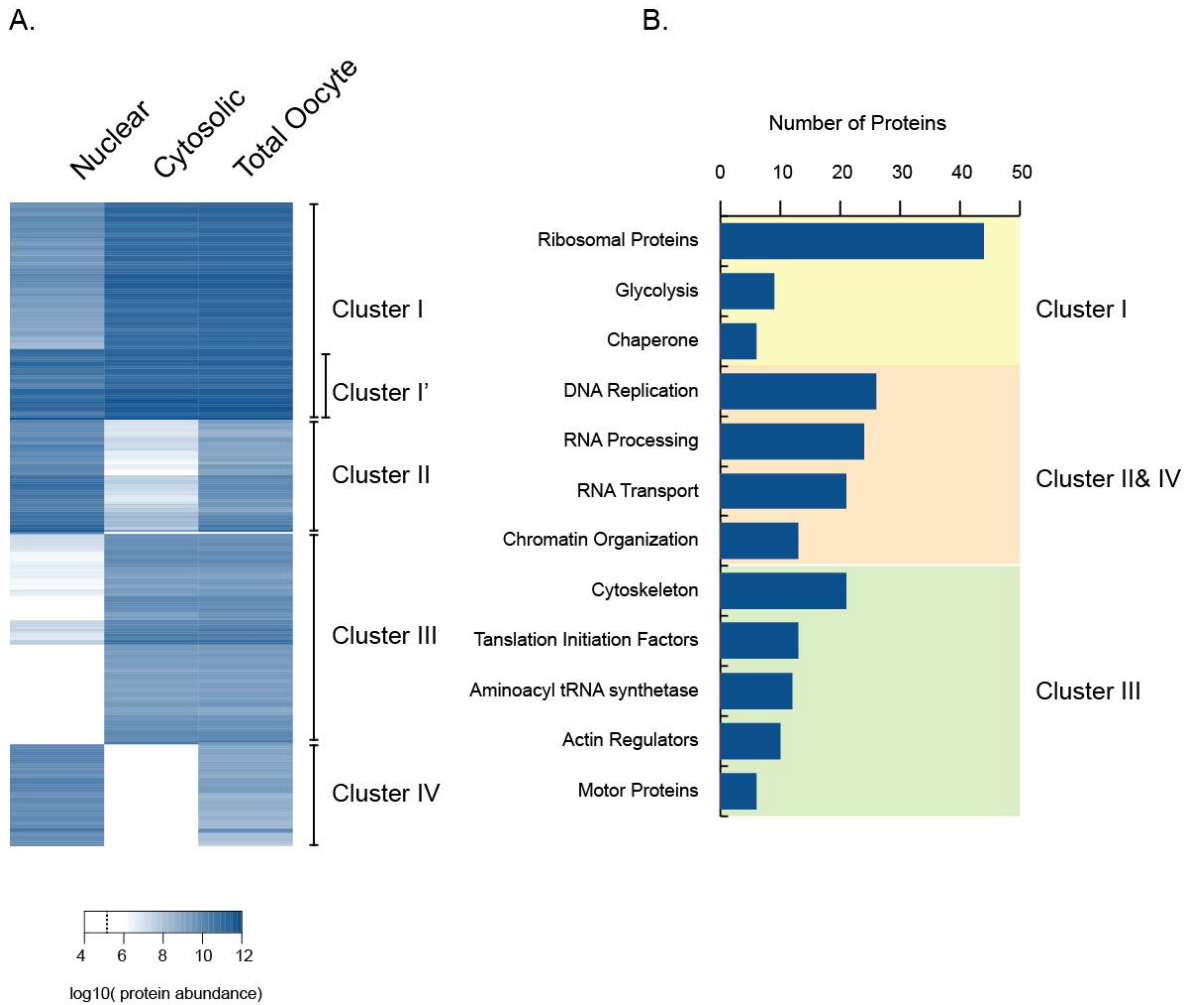


Figure 4.11 Top100 spatial proteome of *Xenopus laevis* oocyte.

A. Heat map representing hierarchical clustering of the 100 most abundant cytosolic, nuclear and total oocyte proteins. Hierarchical clustering revealed distinct specific clusters based on abundance dependant clusters B. Gene ontology (GO) analysis of specific clusters indicated at A. Histograms show number of proteins belong to specific GO terms.

During the development of the *X. laevis* oocyte, oocytes carry out massive transcription and store the produced RNA. After fertilization, fertilized egg undergoes successive 12 rounds of cell division in the absence of the transcription and utilized the stored RNAs for new proteins translation (Newport & Kirschner, 1982). Therefore, I wondered whether oocyte contains high

amount of proteins with respect to its biological specification. To approach this question systematically, I compared the 100 most abundant oocyte proteins with the 100 most abundant proteins across 47 human tissues and cell lines (Wilhelm et al., 2014b). Interestingly, two RNA binding proteins, cold-inducible RNA-binding protein B (CIRBP-B) and Y-box-binding protein 2 (YBX2) were not present among the 100 most abundant proteins of all 47 human tissues and cell lines. CIRBP is RNA binding protein that act as a negative regulator for adenylation of several mRNAs, thereby preventing their translation (Aoki et al., 2003). This is an essential mechanism to tune translation capacity of mRNAs in *Xenopus* egg and embryos (Mendez and Richter, 2001). Exportin 6 and c-Mos are prominent examples of poly(A) tail based translation regulation (Bohnsack et al., 2006; Gebauer et al., 1994). YBX2 is a major component of messenger ribonucleoprotein particles (mRNPs) involving cytosolic storage of mRNA (Murray et al., 1991).

A comparative analysis of the 100 most abundant proteins in the cytosol, nucleus and total oocyte extracts (TOE) revealed drastic differences between these proteomes. Hierarchical clustering of the 100 most abundant proteins in all fractions was depicted as a heat map and it revealed four distinct clusters (Figure 4.11A). Functional categories of the proteins that belong to these clusters were represented with bar plots (Figure 4.11B). Cluster I is composed of proteins, which are highly abundant in the TOE and cytosolic extract, and either high or relatively less abundant in the nuclear extract. The proteins of Cluster I belong to three functional groups. The most over-represented group in the Cluster I is ribosomal proteins with 50 members. As pointed before, ribosomal proteins are most abundant in the TOE, here it is revealed that that they are the most abundant protein in the cytosol as well. Additional to ribosomal proteins, there are glycolysis pathway members and several protein chaperones. A subgroup of this cluster, Cluster I' consists of proteins, which are highly abundant in all three fractions. These proteins include actin, actin interaction protein cofilin, ubiquitin, several redox homeostasis enzymes (such as Peroxiredoxin 2 and 6), several glycolysis enzymes (phosphoglycerate kinase 1, alpha enolase, triosephosphate isomerase) and some chaperones (heat shock 70kDa protein 8). It is well established that the filamentous form of actin is almost absent in all eukaryotic nuclei apart from *X. laevis* oocyte nuclei. This is due to absence of export mediator of actin, Exportin 6, in the oocytes. Actin is present in oocyte nucleus as F-actin to provide mechanical support (Bohnsack et al., 2006) and here I show that actin is almost equimolar distribution across the NE.

Cluster II and IV are nuclear specific proteins; those proteins are either completely absent or present in a very low abundance in the cytosolic fraction. Those proteins were primarily involved in DNA replication, chromatin organization which are nuclear specific molecular activities.

Additionally, RNA processing (including transcription) and RNA transport pathway members were highly abundant in the nucleus. Lastly, Cluster III comprises proteins that are exclusively or predominantly cytosolic proteins. These are translation initiation factors and aminoacyl tRNAs synthetases. Additionally, cytoskeleton proteins, particularly, molecular motors, actin regulators are exclusive to cytoplasm. In general, *X. laevis* proteome characterization shows together with many other studies, that compartmentation causes spatial distribution of distinct molecular activities. The cytosol is dominated with proteins involved in protein translation and structural integrity of oocyte, whereas the nucleus is marked with presence of DNA and RNA activities. Finally, proteins highly abundant in both compartments and TOE show primarily function that take place in both compartments.

4.4.3 CRM1 affinity chromatography and bioinformatics characterization of RanGTP dependent CRM1 interactome

With obtaining localization profile of thousands of proteins in the oocyte, I complemented this data set with affinity chromatography to selectively enrich proteins binding to CRM1 in RanGTP dependent manner. As a background control, empty beads were incubated with the oocyte extract. No specific background was retained on the beads (without CRM1 immobilization) with or without RanQ69L (Figure 4.12A). When, the *X. laevis* extracts were incubated with immobilized CRM1 to streptavidin-agarose beads, high numbers of the interaction partners were retained in the presence of 5 μ M RanQ69L (lane IV). Following elution of the proteins from beads of control (without RanQ69L) and experimental case (with RanQ69L), elutions were separately run on SDS-PAGE, proteins were cleaved into peptides with trypsin and analyzed with LC-MS in triplicates. By combining data from the control and the experimental case, I identified totally 1946 proteins with RanQ69L incubation, and 1022 proteins without RanQ69L incubation. In order to prevent the bridging effect of FG-nucleoporins, all NTRs were depleted with phenyl-sepharose matrix. Successfully, none of the NTRs were identified in both with RanQ69L and without RanQ69L incubation. I performed label-free quantification (LFQ) to identify significant RanGTP dependent CRM1 interaction partners. LFQ results were represented with volcano plot (Figure 4.12B), where x-axis is the log₂ of RanQ69L dependent enrichment ratio and y-axis is the -log₁₀ of p-value, both were calculated by a t-test. Two main criteria were set to consider a protein significant interaction partner, p-value lower than 0.01 and at least 2 MS/MS spectra counts were required for a protein per each replicate. Proteins matching these thresholds (in total 632 proteins), were colored as blue dots in the volcano plot. I wondered how RanGTP dependent CRM1 interactome localizes in the *X. laevis* oocyte. Among 632 proteins, 631 of them were mapped into the *X. laevis* proteome (see Appendix 3 for protein IDs).

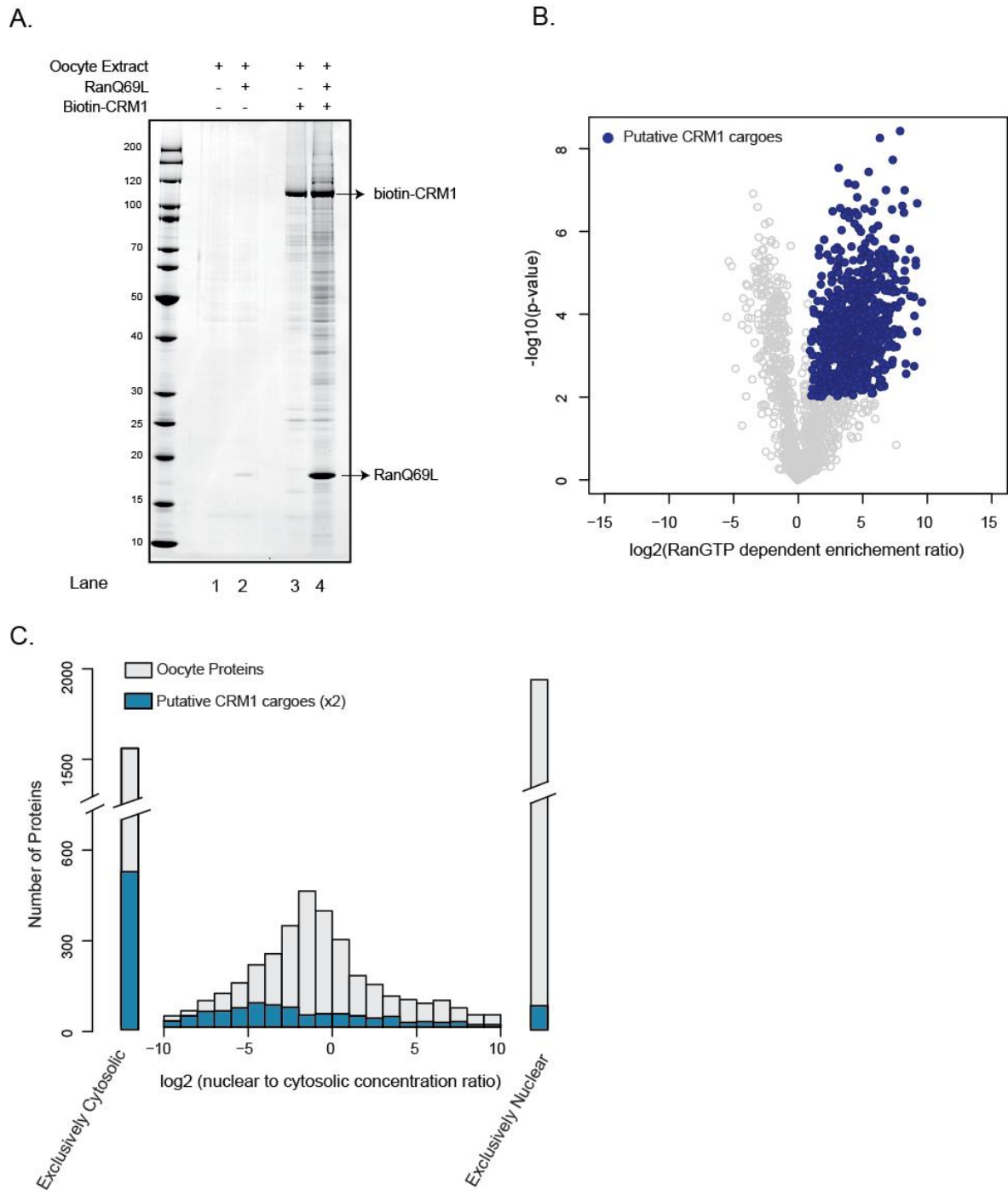


Figure 4.12 Quantitative analysis of CRM1 affinity chromatography.

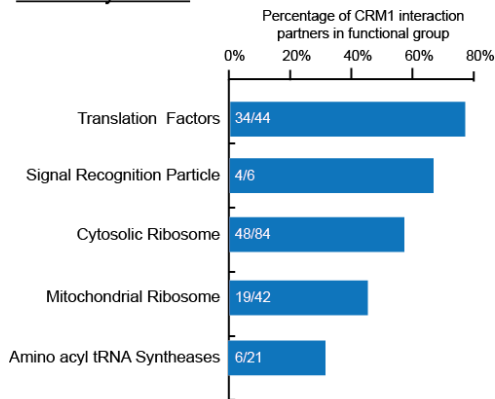
A. SDS-PAGE analysis of CRM1 affinity chromatography. *X. laevis* oocyte extract was applied to either biotin-CRM1 on immobilized streptavidin-agarose beads or “empty” streptavidin-agarose beads. Complex formation was induced with addition of 5 μM RanQ69L (lane IV). Elutions from lane 3 (without RanQ69L, as a control) and lane 4 (with RanQ69L) was analyzed with MS. B. Scatterplots represent label free quantification (LFQ) of CRM1 RanGTP dependent interactome, $-\log_{10}(\text{p-value})$ was plotted on the y-axis and \log_2 of RanGTP dependent enrichment ratio was plotted on the x-axis. Significant proteins were represented with dark blue circles, and the rest with grey circles. See Appendix 3 for protein IDs C. Mapping of the CRM1 RanGTP dependent interactome into *X. laevis* oocyte localization data (Figure 4.10 A). Histogram plot represents \log_2 (nuclear to cytosolic concentration ratios) distribution of both all quantified *X. laevis* proteins (grey bars) and CRM1 RanGTP dependent interactome (blue bars). Outside of the x-axis scale, two bars show proteins only detected either in cytosolic (left- bar; exclusively cytosolic) or nuclear (right-bar; exclusively nuclear) fraction of the oocyte.

Then, localization profiles of these RanGTP dependent CRM1 interactome were overlaid onto *X. laevis* spatial proteome (Figure 4.12C) in which grey bars represent all quantified proteins in *X. laevis*, and blue bars represent the RanGTP dependent CRM1 interactome. To make RanGTP dependent interactome visible in the figure, number of the proteins in particle bin was multiplied by two. Unlike to observed tri-modal localization distribution of *X. laevis* spatial proteome, RanGTP dependent CRM1 interactome showed uni-modal positively skewed (tailed) distribution. Moving through from positive values (nuclear localization) to negative values (cytosolic localization) the number of interaction partners increases steadily, even though, the number of the oocyte proteins decreases.

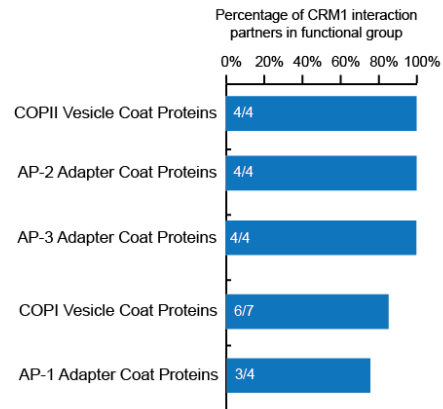
Approximately 17 % (259 out of 1534) of the exclusively cytosolic proteins were RanGTP dependent CRM1 binders. Surprisingly, there were some proteins having more prominent nuclear localization, 53 proteins have log₂ protein concentration ratios greater than 5 or had exclusive nuclear localization (only 2.4% of predominantly nuclear proteins). Remarkably, ~65% of the RanGTP dependent CRM1 interaction partners were predominantly or exclusively cytosolic, highlighting the principle function of CRM1, to export proteins out of the nucleus. I further dissected the RanGTP dependent CRM1 interactome into functional categories by using Gene ontology (GO) annotations and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Huang et al., 2009; Kanehisa et al., 2012). Afterwards, obtained functional terms were manually inspected, and redundant terms were discarded to present most representative functional groups. Figure 4.13 shows over- and under-represented functional groups among the CRM1 interactome. These functional groups were then classified into broader biological categories, which are protein metabolism, membrane vesicle trafficking, cytoskeleton, and miscellaneous. The category miscellaneous is composed of proteins and protein complexes that are not closely associated with each other. In each category, functional groups were represented on the left side of the figure to provide more detailed description. Histograms represent percentage of the RanGTP dependent CRM1 interaction partners in each functional group. The first functional category is protein synthesis and its functional groups are translation factors, cytosolic, aminoacyl tRNA synthetases and signal recognition particle (SRP). Here almost half of the cytosolic ribosome components were significantly detected with CRM1 in the presence of RanGTP. This validates and complements previous studies (Ciuffo and Brown, 2000; Ho et al., 2000; Moy and Silver, 2002). SRP is composed of six proteins and 7S RNA and it is assembled in nucleolus. Afterwards, it is exported to direct cotranslational translocation of proteins containing signal sequence into ER (Alberts et al. 2002). Previously, it was shown that this complex is LMB sensitive in yeast (Grosshans et al., 2001).

A. Over-represented protein groups among RanGTP dependent CRM1 interactome

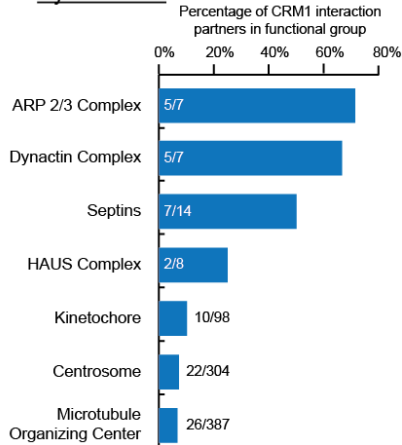
Protein synthesis



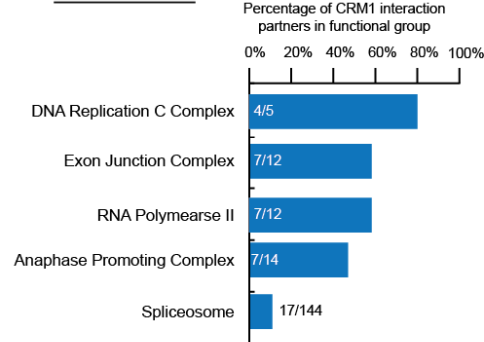
Membrane vesicle trafficking



Cytoskeleton



Miscellaneous



B. Under-represented protein groups

Metabolic Pathways

- Glycolysis
- Citrate cycle (TCA cycle)
- Pyruvate Metabolism

DNA Repair Pathways

- Base excision repair
- Nucleotide excision repair
- Mismatched repair

Ubiquitin Proteasome System

Figure 4.13 Characterization of RanGTP dependent the CRM1 interactome.

Putative CRM1 cargoes were classified into functional groups based GO analysis. Different functional groups were categorized based on cellular roles. A. Bar plots shows over-represented protein groups in translation, membrane vesicle trafficking, cytoskeleton, miscellaneous functions. Each bar represents percentage of proteins in putative RanGTP dependent CRM1 binders of particular group in. Numbers of significant proteins of a particle group in putative RanGTP dependent CRM1 interactome were shown in or out side of the bars. B. Under-represented protein groups among RanGTP dependent CRM1 interactome.

Here, I provide evidence that the vertebrate SRP particle is also exported by CRM1. The remaining functional groups are translation factors, aminoacyl tRNA synthetases. Almost all of the core translation initiation factors (24 out of 30) were detected in the presence of RanGTP. Previously, Bohnsack *et. al.* showed that tagged versions of several individual translation factors (eIF2 γ , eIF2B ϵ , eIF5, eIF4A1, eRF1) were LMB sensitive. In this study, all components of eIF2 complex, eIF2B and eIF3 complexes were significantly detected as CRM1 binders. Almost all translation initiation factors individually or in a complex interacts with CRM1 in RanGTP dependent manner. Another category of protein synthesis is aminoacyl tRNA synthetases, which load amino acids to their cognate tRNAs. Together with eEF1A, loaded tRNA brings amino acid to ribosome for polypeptide formation. Six out of 21 amino aminoacyl tRNA synthetases showed RanGTP dependency to interact with CRM1.

The second biological category encompasses the coat proteins of vesicles that mediate intracellular membrane trafficking. Based on the coat proteins, there are three main vesicle types, COPI, COPII and clathrin coated vesicles. COPI coated vesicles sort proteins from Golgi to ER, while COPII coated vesicles sort proteins from ER to Golgi (Kirchhausen, 2000). Almost all components of COPI (four out of four) and COPII (six out of seven) coat complexes were significantly detected in CRM1 interactome (Figure 4.13A, upper right panel). Apart from the coat proteins, one of the COPI coated vesicle regulatory subunit, ARF3 (ADP-ribosylation factor 3) was detected as well. It is a GTP binding protein regulating assembly of coat proteins on donor membrane (Alberts *et al.*, 2002). Additionally, ADP-ribosylation factor GTPase-activating protein 3 (a GTPase activating protein of ARFs) that promotes hydrolysis of ARF-bound GTP was one of the significantly enriched CRM1 binders (Liu *et al.*, 2001).

In addition to COPI and COPII coated vesicles, clathrin-coated vesicles mediate another type of intracellular membrane trafficking. They are coated by clathrin, that forms a triskelion structure, which is composed of three heavy and light chains (Kirchhausen, 2000). None of the clathrin chains were detected as a CRM1 binder. However, unlike to COPI and COPII vesicles, a clathrin vesicle requires an adapter protein (AP) complex, which defines the final destination of the vesicle (Kirchhausen, 2000). Even though heavy and lights chains of clathrin were not identified at all, almost all components of AP-1, AP-2 and AP-3 were significantly enriched.

The third biological category is the cytoskeleton system (Figure 4.13A lower-left panel). Proteins that belong to these category are the Arp2/3 complex, which is major actin cytoskeleton regulator, dynactin complex that is involved in bidirectional intracellular transport, and the HAUS complex that regulates mitotic spindle assembly (Alberts *et al.* 2002). Another interesting group

of proteins are septins. They are GTP binding proteins that assembled into high-order structures like filaments and rings, hence they are considered part of the cytoskeleton (Mostowy and Cossart, 2012). They are implicated in variety of cellular processes such cytokinesis, ciliogenesis and forming diffusion barriers for compartmentation of cellular domains (Mostowy and Cossart, 2012). In the KEGG septin group comprises 14 genes. Seven septin proteins were identified in *X. laevis* oocyte, and all of them interact with CRM1 in RanGTP dependent manner. In addition to these, there were many other proteins involved in cytoskeleton-based functions, such as kinetochore, centrosome and microtubule organizing factor.

The last category, named miscellaneous, is composed of several protein complexes which have diverse cellular functions. It contains group of proteins, including subunits of RNA polymerases II, DNA replication factor C, and Anaphase-promoting complex/Cyclosome (APC/C). RNA Polymerase II synthesizes pre-mRNAs and most microRNAs (Alberts et al., 2002). In the *X. laevis* oocyte, RNA polymerase components have predominant or exclusively nuclear localization (Figure 4.16). DNA replication factor C is five-subunit protein complex, which is responsible for loading of PCNA onto DNA. Four out of five subunits of this complex were enriched in CRM1 affinity chromatography. APC/C is E3 ubiquitin ligase that is essential for progression through cell cycle by targeting degradation of cell cycle proteins. APC/C has equimolar distribution between the nucleus and the cytosol (Figure 4.18) in the *X. laevis* oocyte and five members of this complex were interacted with CRM1 in RanGTP dependent manner. Four members of the APC/C with common three proteins, Cdc23/ANAPC8, ANAPC1, and ANAPC5, were also depleted in the nuclear fraction after LMB treatment in HeLa cells (see Appendix 1).

Apart from these over-represented functional groups, I next investigated cellular processes, which were under-represented in CRM1 interactome. Three main cellular processes were under-represented in CRM1 interactome, these are metabolic pathways, and DNA associated pathways and ubiquitin-proteasome system. These cellular processes did not have significant number of CRM1 binders.

4.4.4 Comparison of two orthogonal methods to identify CRM1 export cargoes

Two orthogonal methods employed in two different organism lead to identification of ~780 potential CRM1 cargoes. With LMB treatment 138 human, and with CRM1 affinity chromatography 632 *X. laevis* candidates were identified. These 632 *X. laevis* proteins were initially mapped to human orthologs and after removing isoforms and allelic variants; two potential CRM1 cargo pools were compared. 52 proteins were identified with these orthogonal methods (Figure 4.14). Among these 52 proteins, 15 of them were 60S ribosomal proteins.

Additionally, this overlap contains several previously characterized CRM1 cargoes; some of them are eIF6, DIAPH3, RanBP1. Among the five proteins, proteins (SQSTMS1, CIP2A, and CCP1) that were validated before (Figure 4.7), were detected as a significant binders in CRM1 affinity chromatography as well. Another common set of proteins is several members of APC/C. LMB treatment led to identification of four, affinity chromatography led to five members of APC/C with three overlap.

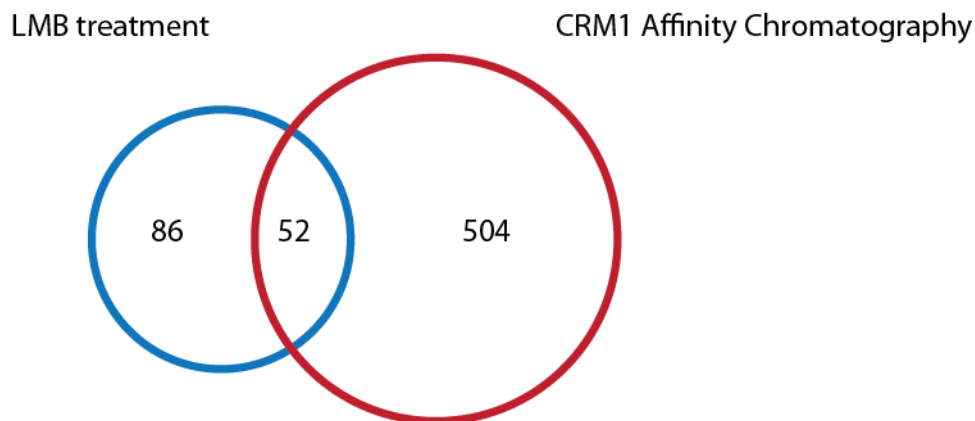


Figure 4.14 Comparison of potential CRM1 cargoes between LMB treatment in HeLa cells and CRM1 affinity chromatography from the *X. laevis* oocyte extract.

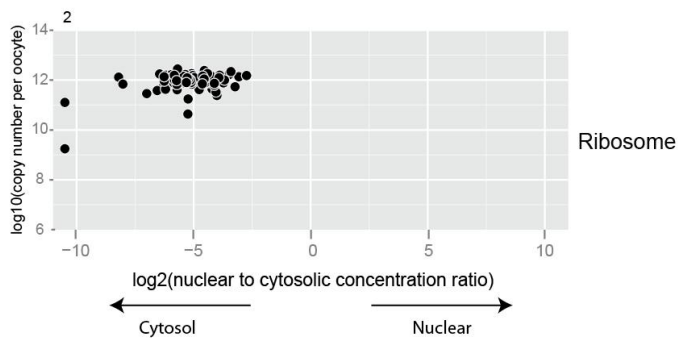
Interestingly, two proteins, which are involved in RNA Polymerase II based transcription, were common in both data sets. One of them is DNA-directed RNA polymerase, II subunit RPB2 (POLR2B) that is required for Pol II catalytic activity. Other one is RNA polymerase II associated proteins I (RPAP1).

Apart from the common proteins, high number of different cargoes was identified. Almost 5 times more candidates were detected with CRM1 affinity chromatography. Several NTRs were detected significant with LMB treatment, whereas none of the NTRs were identified at all with affinity chromatographic enrichment, showing successful NTR depletion with phenyl sepharose matrix. An interesting difference between HeLa and *Xenopus* system was the COP9 signalosome. All components of COP9 signalosome were depleted in the cytosolic fraction with LMB treatment in HeLa cells. However, none of the COP9 signalosome components were detected as CRM1 binders from the *X. laevis* extract.

4.5 Molecular pathway dissection of *Xenopus laevis* oocyte proteome with respect to RanGTP interactome of CRM1

Having obtained the localization profile of thousands proteins in the oocyte and the RanGTP dependent interactome of CRM1, I proceeded with inspection of these two datasets with respect to molecular pathways in *X. laevis* oocytes. In order to visualize the localization distribution of proteins, I chose a scatter plot representation in which each dot represents a quantified protein. I will refer to this kind of representation as a localization profile, or *loc-profile* for simplicity. As an example, *loc-profile* of ribosomal proteins was plotted at Figure 4.15A. The X-axis is log₂ of nuclear to cytosolic concentration ratio and y-axis shows log₁₀ of calculated protein copy numbers per oocyte. Negative values represent cytosolic localization and positive values represent nuclear localization.

A.



B.

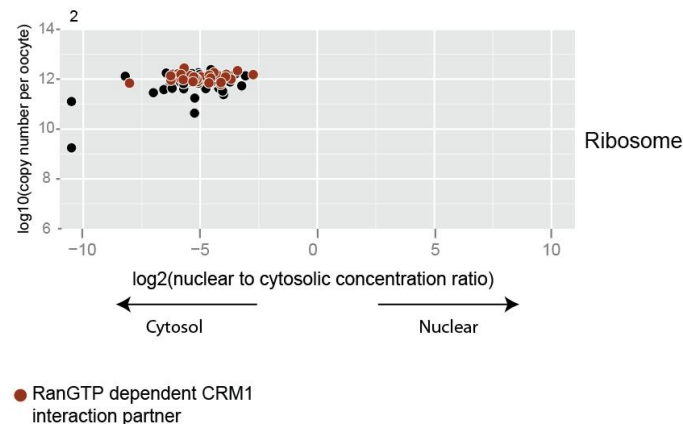


Figure 4.15 Representation of localization profiling of molecular pathways.

A. Scatterplot representation of nuclear-to-cytosolic concentration distribution of ribosomal proteins. The X-axis is the log₂ of the nuclear to cytosolic concentration ratio. Y-axis represents the log₁₀ of calculated protein copy numbers per oocyte. B. Same as A in which significant RanGTP dependent CRM1 interaction partners are additionally colored in red.

For instance, log₂ ratio of -5 would mean that the protein is approximately 32 times more concentrated in the cytosol with respect to nucleus. To incorporate proteins that were only

identified in one compartment into the plot, they were placed out of the ± 10 scale. Numbers right above the figures show how many proteins were placed outside of the ± 10 scales. A small fraction of quantified proteins presented extreme \log_2 protein ratios < -10 and > 10 (88 proteins out of ~ 6300 proteins). These proteins were placed outside the scales as well. RanGTP dependent CRM1 binders incorporated in *loc-profiles* as a color code. In case a protein was significantly identified as a CRM1 binder, it was represented as a red dot in the figure (Figure 4.15B). *Loc-profiles* of three essential molecular pathways were further dissected with respect to RanGTP interactome. These pathways are RNA metabolism, protein synthesis and ubiquitin proteasome system.

4.5.1 RNA Metabolism

The central dogma of the molecular biology explains the flow of the genetic information in biological systems (Crick, 1970). Briefly, genetic information present in DNA is transferred to mobile RNA molecules, which serve as templates for protein synthesis. I will partially follow the same path as the central dogma and I will explain spatial distribution of molecular mechanisms responsible for RNA metabolism, protein synthesis and finally ubiquitin-proteasome system. Meanwhile, I will correlate spatial distribution of these pathways with the RanGTP-dependent interactome of CRM1.

RNA molecules emerge in the nucleus synthesized by RNA polymerases. Eukaryotes possess three types of RNA polymerases for synthesis of variety of RNA molecules. In addition to the RNA polymerases, there are other factors needed for faithful RNA production. These factors form pre-initiation complex, which melt DNA to release the RNA polymerase from promoter to elongation mode, namely basal transcription factors. Additionally, some accessory proteins are required for proper communication of transcription activators with RNA polymerases and basal transcription factors, such as mediator complex. During *X. laevis* oogenesis, the oocytes carry out extensive transcription to produce vast amounts RNAs, which will be used at early embryogenesis. Stage IV oocytes become practically silent in transcription until midblastula transition (Newport & Kirschner 1982). Even though, the amount of the RNA polymerases does not change significantly until the gastrulation, their activity is almost undetectable in *X. laevis* oocytes and eggs (Roeder 1974). However, elegant microinjection experiments showed that oocyte and egg were still compatible to transcribe exogenous DNA for a certain time (Newport & Kirschner 1982). These studies proposed that suppression of the transcription machinery is an active process involving on DNA and/or chromatin structure rather than regulating the abundance of RNA polymerases. Nevertheless, it is possible that cell might employ multiple and complementary strategies to

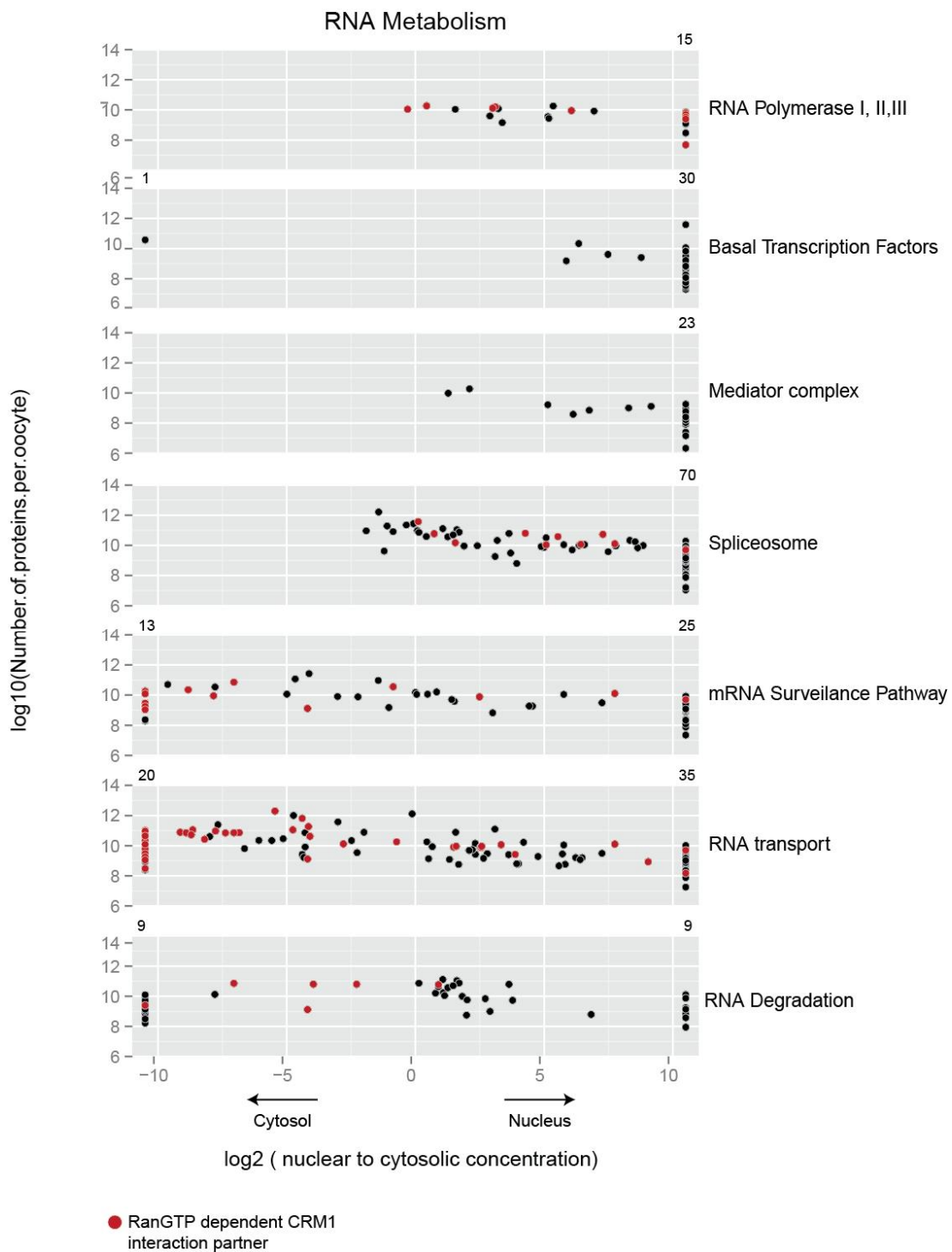


Figure 4.16 Localization profile of RNA metabolism.

Scatterplots show the localization profiles of protein groups or molecular pathways, related to RNA metabolism. X-axis is the log₂ of nuclear-to-cytosolic concentration ratio, and y-axis is log₁₀ of absolute copy number per oocyte. Each dot represents a quantified protein in the oocyte and those are in red additionally represents RanGTP dependent CRM1 binders.

regulate temporal organization of cellular activities. A prominent example is the prevention of DNA re-replication after S phase. Two important players of initiation of DNA synthesis are Cdc6 and MCM complex, they involve in formation pre-initiation complex. After initiation of DNA synthesis, the active concentration of Cdc6 is reduced by degradation and the MCM complex is spatially regulated by nuclear export to ensure that its concentration in the nucleus is limited (Nguyen et al., 2001). Therefore, to scrutinize whether a similar spatial regulation is present in *X. laevis* oocytes for RNA transcription, I concentrated on *loc-profiles* of not only RNA polymerases, but also whole RNA metabolism (Figure 4.16). The first three panels show *loc-profiles* of RNA polymerases, basal transcription factors and mediator complex. Aside from TFIIIA, most of the protein components of RNA polymerases, basal transcription factors and components of mediator complex have either prominent or exclusive nuclear localization. As supportive evidence this results, none of the components of basal transcription factors and mediator complex were shown to be CRM1 binders. However, I cannot completely rule out the possibility that exclusively cytosolic (or nuclear) proteins were out of the detection limit in other compartment due to the very scarce abundance.

All the protein components of RNA polymerase and mediator complex were detected. However, they dramatically varied in abundance with several orders of magnitude. On the contrary, several basal transcription factors were not identified, particularly, three subunit of TFIID complex and several TATA box binding proteins for RNA polymerase I. TFIIIA showed exclusive cytoplasmic localization. This is in line with previous studies where nuclear amount of TFIIIA is barely detectable (Shastry et al., 1984).

While one of the major functions of TFIIIA is to regulate 5S RNA transcription, it has an additional role in *X. laevis* oogenesis (Camier et al., 1995). It predominantly associates with 5S rRNA and is involved in storage of 5S RNA as a 7S RNP in the cytosol (Shastry, 1996). Furthermore, several components of RNA polymerase II, POLR2B, POLR2C, POLR2J2, showed almost equimolar distribution and they were RanGTP dependent CRM1 binders. It is supported by the fact that POLR2B showed localization changes after LMB treatment (Appendix I).

Subsequent to RNA synthesis, pre-mRNA splicing takes place in the nucleus. Spatial regulation must ensure that no unspliced mRNA is available for translation in the cytoplasm. Therefore, splicing must take place before mRNA is exported to the cytosol. At Figure 4.16 the fourth panel validates that majority of splicing machinery had predominant or exclusive nuclear localization (68 proteins exclusively detected in the nucleus). Some accessory proteins, which might have

additional functions, have equimolar distribution. Such as several Lsm proteins, apart from their primary role in maturation of U6 snRNPs, they are involved in mRNA degradation (Houseley and Tollervey, 2009; Will and Lührmann, 2001). Only ten proteins showed RanGTP dependent interaction with CRM1 out of more than 100 protein members of spliceosome. Some of these CRM1 binders are not *bona fide* splicing proteins and they have auxiliary roles in splicing such as, nuclear cap-binding protein subunit 1 (CBP80) which is part of the RNA cap binding complex (CBC) involving in U snRNA export via CRM1 (Ohno et al., 2000) or MAGOH which is part of the exon junction complex (EJC) that deposited on exon-exon boundaries (Bono et al., 2006). In general, core-splicing machinery revealed predominant or exclusive nuclear localization and mainly splicing proteins are devoid of RanGTP dependent interaction with CRM1.

The Last three panels of the Figure 4.16 illustrate *loc-profile* of proteins involved in mRNA surveillance, RNA transport and RNA degradation. These pathways are not mostly sequential and are not completely isolated from each other. They are mechanistically coupled and share common proteins. Collective work of studies from past 40 years has showed that these molecular processes span both the nucleus and the cytosol (Bentley, 2014). Overall, nuclear-to-cytosolic distribution of proteins involved in mRNA surveillance, RNA transport and RAN degradation suggests that these pathways require collective work of proteins localized to nucleus, cytosol and shuttling proteins between two compartments. Several proteins involved in this processes were CRM1 dependent interaction partner and they were mainly predominant cytosolic proteins with several exceptions. The prominent examples are regulator of nonsense transcripts 1 and 2 (UPF1 and UPF2). They are recruited to mRNA upon detection of premature termination codon (PTC), following phosphorylation and interaction with EJC appears to be a signal for Non-sense mRNA decay (NMD) (Bentley, 2014). Both these proteins were exclusively cytosolic and this observation is supported that both of them were RanGTP dependent CRM1 binders.

4.5.2 Protein Synthesis

In prokaryotes, translation takes place while the nascent transcript is still being synthesized. However, this might cause serious problems in eukaryotes due to the presence of introns. Eukaryotic pre-mRNAs must be spliced before translation. Otherwise, it would result in translation of truncated proteins, or in the worst-case these truncated versions could have dominant negative effect. Eukaryotes avoid this problem by physically separating transcription and translation by the NE. However, several studies claimed the presence of nuclear translation around 10-15 % of the cellular translation as a proof reading mechanism for mRNA (David et al., 2012; Iborra et al., 2001). Spatial distribution of nine GFP-tagged translation factors in human cell

cultures showed that the nucleus contains very low amount of translation factors and active nuclear export of them dis-favours the concept of such nuclear translation (Bohnsack et al., 2002). Therefore, to approach this issue in broader perspective, I concentrated on *loc-profiles* of protein

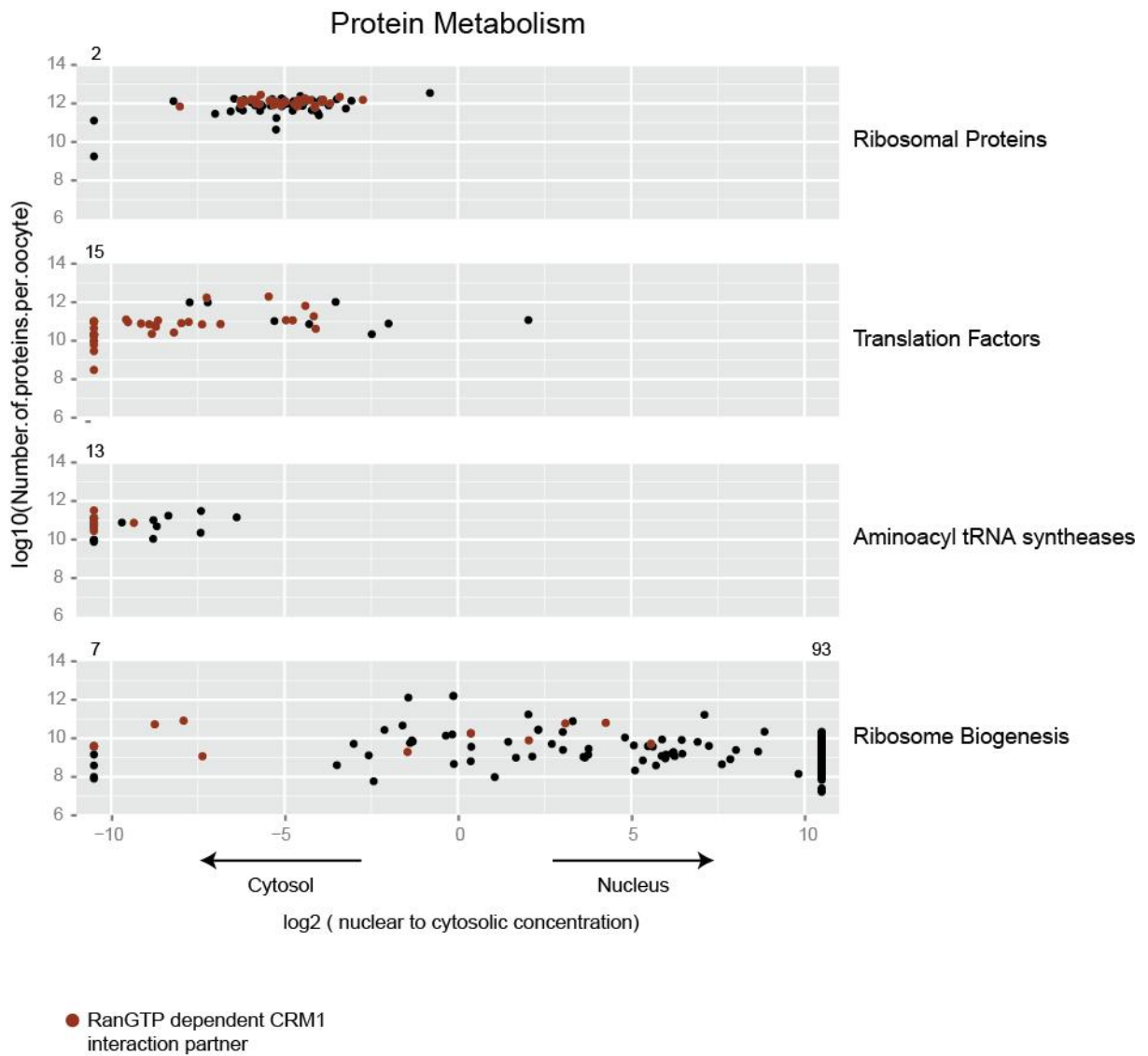


Figure 4.17 Localization profiling of protein synthesis.

Scatterplots show localization profile of proteins groups or molecular pathways, which are related to protein synthesis. X-axis is the log2 of nuclear-to-cytosolic concentration ratio, and y-axis is log10 of absolute copy number per oocyte. Each dot represents a quantified protein in the oocyte those in red additionally represent RanGTP dependent CRM1 binder.

synthesis system, namely ribosomal proteins, translation factors, aminoacyl tRNA synthetase, and ribosome biogenesis. Successfully, I identified and quantified almost all ribosomal proteins, 46 proteins of the 60S ribosomal and 32 proteins of 40S ribosomal subunits (Figure 4.17, first panel). 60S and 40S ribosomal subunits mature as ribonucleoprotein complexes in the nucleolus. However, at a steady state, almost all of the ribosomal proteins showed predominant cytosolic localization, apart from 2 exclusively cytosol localized proteins. They roughly have log₂ protein ratio of -5, in other words they are ~ 32 times more concentrated in the cytosol with respect to the nucleus. As expected, the majority of them have similar copy numbers to each other. Among 78 identified ribosomal proteins, 30 of the them (13 proteins from 40S and 17 from 60S subunit) were RanGTP dependent CRM binder. It was already known that 60S and 40S ribosomal subunits are exported via CRM1, where NMD3 is export adapter of 60S subunit (Ho et al., 2000).

The second panel at Figure 4.17 represents *loc-profiles* of translation factors. Previously, it was shown that GFP fusion of nine translation factors had predominant (>100 fold higher concentration in cytosol) cytosolic localization in human cells (Bohnsack et al., 2002). In this study, almost all translation factors were identified, 33 initiation, six elongation and two termination (these numbers include isoforms and allelic variants) factors. In the *X. laevis* oocyte, translation factors has similar localization profile as human orthologs (Bohnsack et al., 2002). Majority of translation factors were not detected at all in the nucleus or they were very low abundant in the nucleus (Figure 4.17, second panel). Unlike to other translation factors, eukaryotic initiation factor 6 (eIF6) was slightly more concentrated in the nucleus. This could be explained by the fact that eIF6 participates in ribosome biogenesis (Sanvito et al., 1999).

Interestingly, many of the core translation initiation factors (24 out of 30), were identified as RanGTP dependent interaction partners of CRM1. All components of eIF2, eIF2B and eIF3 complexes were RanGTP dependent CRM1 binders, this suggests that CRM1 interact not only as a individual proteins, but also as a complex. Overall, this study shows that translation factors are either absent or present with very low abundance in the nucleus and almost all of them bind to CRM1 in RanGTP dependent manner. Other sets of molecules that are indirectly needed for translation are aminoacyl-tRNAs. Along with elongation factor, eEF1A, they bring amino acids to the ribosome for incorporation to polypeptides. The enzymes that load amino acids to cognate tRNAs are aminoacyl tRNA synthetases. *Loc-profile* of these enzymes shows similar trend as translation factors. They had either exclusive or predominant cytosolic localization (Figure 4.17, third panel).

Ribosomal biogenesis starts at the nucleolus with synthesis of ribosomal RNAs (rRNA), 18S, 5.8S and 25S. Then, series of modifications takes place on the rRNA and with assembling of ribosomal proteins on rRNAs leads to emergence of pre-40S and pre-60S ribosomal subunits. Figure 4.17 panel four reveals that majority of the ribosomal biogenesis factors are nuclear, some of them show equimolar distribution and several are exclusive cytosolic proteins. These results were consistent with previous studies that ribosome biogenesis is governed at the nucleus. Only several proteins have been identified as RanGTP dependent CRM1 interaction partners (11 out of 167 proteins). This time, not all interaction partners show predominant cytosolic localization. Some of them have predominant cytosolic and some have equimolar distribution between the cytosol and the nucleus.

4.5.3 Ubiquitin Proteasome Pathway

Apart from compartment specific molecular pathways as mentioned before, in this chapter I will concentrate on an example of molecular pathway spanning two compartments based on *X. laevis* spatial proteomics analysis. The ubiquitin-proteasome system (UPS) is a major pathway of non-lysosomal protein degradation. In addition to protein degradation, ubiquitin system has key roles in a variety of cellular mechanism such as cell cycle, cell division, development, and signaling. In molecular aspects, attachment of an ubiquitin molecule on a substrate protein requires concerted action of three sets of enzyme, activation enzyme E1, conjugation enzyme E2 and ligation enzyme E3. Initially, ubiquitin is activated and covalently attached to E1 enzyme in ATP dependent manner. In humans, there is only one protein that is capable of activating ubiquitin, UBA1. Afterwards, conjugating enzymes E2 catalyzes transfer of ubiquitin from E1 to active site of E2 enzymes. Human genome contains ~30 of E2 enzymes. Finally, E3 ubiquitin ligase mediates recognition of a substrate and then attachment of ubiquitin (in some cases with help of E2 enzyme) via isopeptide bond on lysine residues of the substrate (Hershko and Ciechanover, 1998; Kerscher, 2006). Figure 4.18 first panel shows localization profile of ubiquitin, E1 enzyme and E2 enzymes detected in the *X. laevis* oocyte. Ubiquitin itself and single E1 activating enzyme have equimolar distribution (orange and blue dot, respectively). Even though, there were two compartments specific E2s, most of them showed almost equimolar distribution. In general, the process of ubiquitin conjugation into E2 enzymes was not spatially restricted either to the nucleus or to the cytosol. This observation was supported by the finding that only a single E2 enzyme was RanGTP dependent CRM1 binder. On the other hand, E3 ligases presented more distinct *localization profile* and several E3 ligases were spatially restricted either in the cytosol (18 E3 ligases) or in the nucleus (15 E3 ligases) (Figure 4.18 second panel). Among 18 cytosol specific E3 ligases, four of them were RanGTP dependent CRM1 binder.

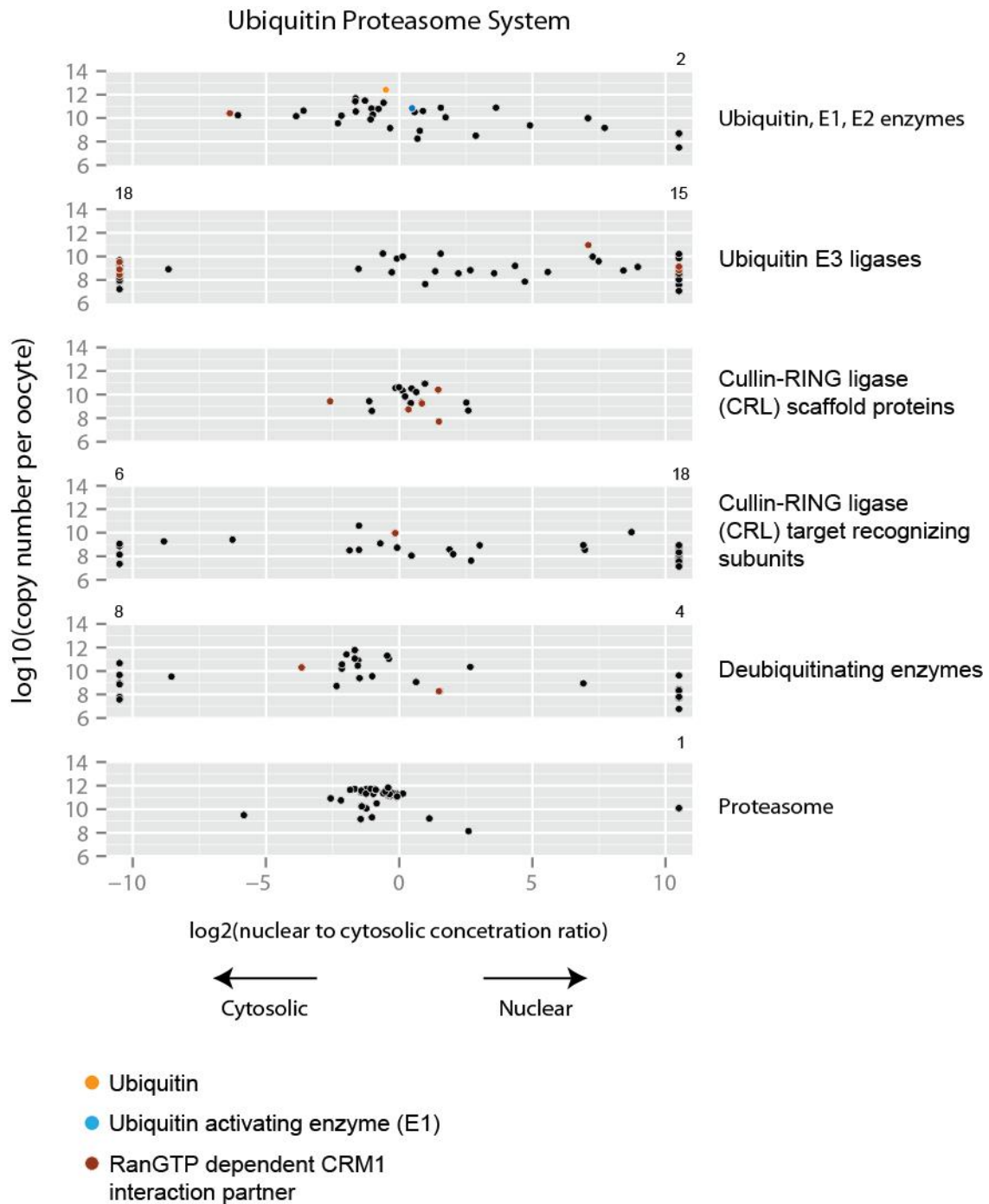


Figure 4.18 Localization profiling of ubiquitin-proteasome system.

Scatterplots show localization profile of proteins groups or molecular pathways, which are related to ubiquitin proteasome system. X-axis is the \log_2 of nuclear-to-cytosolic concentration ratio, and y-axis is \log_{10} of absolute copy number per oocyte. Each dot represents a quantified protein in the oocyte those in red additionally represent RanGTP-dependent CRM1 binder.

Three E3 ligases having predominant or exclusively nuclear localization were CRM1 binders. This suggests that they could be regulated CRM1 cargoes. A group of E3 ligases is considered as modular E3 ligases; archetypes of this group are cullin-RING ligase (CRL) family. They are composed of core scaffold (cullin proteins) with interchangeable substrate recognizing subunits. A typical and best-characterized one is the S phase kinase-associated protein 1 (SKP1)-cullin 1 (CUL1)-F-box protein (SCF) complexes in which F-box protein determines the substrate specificity (Skaar et al., 2013).

The loc-profiling of these scaffold proteins shows that they were not spatially restricted and were present both in the cytosol and the nucleus (Cullin related E3 ligase, APC/C, is also included in this category) (Figure 4.18, panel four). Several proteins in this group were CRM1 binders, which were components of anaphase promoting complex/cyclosome (APC/C).

Similar to E3 ligases, CRL target recognizing subunits had distinct localization profile. Some of the target recognizing subunits had spatially restricted either in the cytosol or in the nucleus, and some of them presented almost equimolar distribution. Additionally, CRL target recognizing subunits were devoid of CRM1 interaction.

Deubiquitinating enzymes (DUBs) are the proteases that cleave ubiquitin and ubiquitin like proteins. They have several roles in the UPS pathway: i) maturation to ubiquitin monomer, ii) recycling of ubiquitin, iii) reversing the ubiquitin and ubiquitin-like modification (Reyes-Turcu et al. 2009). Based on their function in UPS, it is expected that the whole family should not be spatially restricted. Even though some of the DUBs were compartment specific, almost half of the DUBs showed equimolar distribution. There were only two DUBs showing interaction with CRM1 in RanGTP dependent manner (Figure 4.18, panel five). Based on these observations, subsequently, I checked the loc-profile of the proteasome. Since ubiquitination takes place in both compartments, protein components of the proteasome should be present in both compartments as well. 20S and 19S subunits of the proteasome did not show spatial restricted localization, they showed almost equimolar distribution between the cytosol and the nucleus. Only one member of the proteasome had exclusive nuclear localization, proteasome activator subunit 3 (PSME3). It is subunit of the 11S REG-gamma proteasome regulator that modulates the endopeptidases activity of the proteasome (Li et al. 2000).

5 Discussion

5.1 Effects of LMB on the HeLa proteome

Quantitative analysis of LMB treated HeLa cells revealed that LMB is indeed a very specific agent that alters the abundance of only a small portion of the HeLa proteome (3% of the identified proteome) (Figure 4.4A and Appendix 2 for list of the proteins). Additionally, majority of the proteins had relatively small changes (less than 75% depletion or enrichment) in their abundance. Interestingly, depleted and enriched proteins in the TCE were closely associated with each other. Approximately half of the depleted proteins were structural constituents of the ribosome, particularly members of 60S ribosomal subunit (Figure 4.4B). Subcellular fractionation analysis also showed that many 60S ribosomal subunit proteins accumulated in the nucleus upon LMB treatment. Taken together, these observations indicate that inhibiting the export of the 60S ribosomal subunit promote their turnover (e.g. by proteasome), thereby resulting in depletion of 60S ribosomal subunit proteins in the TCE. On the other hand, 40% of the TCE enriched proteins were involved in RNA processing/splicing. This could be due to the indirect effect of the LMB treatment. Mature U snRNPs are imported into the nucleus by the adapter Snurportin 1 (SPN1) and the recycling of SPN1 into cytosol is mediated by CRM1 (Paraskeva et al., 1999). Hence, inhibition of nuclear export of SPN1 might create shortage of mature U snRNPs in the nucleus. Subsequently, this might induce over expression of proteins involved in RNA processing/splicing to compensate decreased levels of U snRNPs in the nucleus.

5.2 MS-based quantitative analysis of LMB treated HeLa cell fractions

The MS-based approach established for detection of proteins that undergo nuclear-to-cytosolic localization changes led to identification of 138 proteins. 59 of these proteins accumulated in the nucleus and 84 of them were depleted in the cytosol with 5 overlapping proteins (Figure 4.6C). A shuttling construct, Rev₍₄₈₋₁₁₆₎-GFP₂-M9, was set up as a positive for localization change. This construct was predominantly localized to nucleus under physiological conditions, and it became exclusively nuclear after LMB treatment, which was documented by microscopic analysis (Figure 4.1B). The MS analysis, successfully, detected the positive control as a depleted protein in the cytosol. Detection of the positive control as a significant hit firstly validates the approach and secondly it stresses that the workflow is sensitive enough to detect even for predominantly nuclear proteins. However, the positive control could not be quantified in the nuclear fraction. It was identified with a single peptide with only one quantification event. Since two quantification events were required for each protein to be quantified, it was not reported as a significant in the

nuclear fraction. It is possible that its peptides were masked due to the presence of other more abundant co-eluting peptides.

Many known CRM1 cargoes (direct or indirect binders) were depleted and/or enriched either in the cytosol or the nucleus, respectively. Almost one-third of the potential LMB sensitive CRM1 cargoes were the proteins of 60S ribosomal subunit or COP9 signalosome (see Appendix 1). In contrast to the presence of many 60S ribosomal proteins, there was only one component of 40S subunit, RPS3A. Previously, siRNA knock-down of CRM1 was shown to result in the inhibition of 40S ribosomal subunit export (Wild et al., 2010). Possibly, longer LMB treatment is required for MS based detection of 40S ribosomal proteins. Among potential CRM1 cargoes, 14 proteins were already documented as LMB sensitive (see Appendix I with references). From those 14 cargoes, three of them (NMD3, RanBP1, DDX3) were confirmed by microscopic analysis (Figure 4.7A). Additionally, nuclear accumulation of five novel cargoes (CCP1, CIP2A, SQSTM1, GNL3L and PDCD2L) was verified by microscopic analysis (Figure 4.7B). Taken together, validation of the candidates and presence of many known CRM1 cargoes confirm the credibility of the list and overall quality of the approach.

Under physiological conditions, upon inhibition of CRM1 export pathway, an ideal CRM1 cargo should be depleted from the cytosol and accumulated in the nucleus. However, the overlap between proteins depleted from the cytosol and enriched in the nucleus, respectively, was relatively low, only five candidates. There could be several reasons to account for this observation. First, cross-contamination between the cytosol and the nucleus due to the subcellular fractionation might decrease the strictness of the localization of proteins in these compartments. Second, this could be explained by relative quantification by MS. A significant detection of a protein in both compartments depends on a sufficient relative amount changes both in the cytosol and in the nucleus. In case of a scenario where a cargo is predominantly cytosolic (assuming 100X absolute abundance in the cytosol) with scarce amount (1X amount) in nucleus, upon LMB treatment a fraction of the protein would accumulate in the nucleus. Hypothetically, if 5X amount accumulated in the nucleus, it would be significant only in the nuclear fraction. Relative nuclear abundance change in the nucleus would be 500%; however, cytoplasmic relative change would only be 5%. Therefore, cytosolic depletion would not be detected as significant as the nuclear accumulation. Hence, for a cargo to be significantly identified in both compartments, sufficient relative abundance change in both fractions is required. Furthermore, this underlines that it is necessary to analyze both the nuclear and the cytosolic fractions for more comprehensive cargo identification.

A group of the proteins presenting localization changes could be attributed to secondary effects of the LMB. Several NTRs showed mis-localization upon LMB treatment. CRM1 itself and several importins – importin alphas, transportin, importin 4 and 8 appeared as a significant hit. Under physiological conditions NTRs leave the nucleus in the form of a complex with RanGTP. At the cytoplasmic site concerted action of RanGAP and RanBP1 results in removal of NTR bound RanGTP and subsequent hydrolysis of GTP. RanBP1 is already known CRM1 cargoes and appeared as a significant hit in this screen as well (Zolotukhin and Felber, 1997). Upon inhibition of CRM1 export, RanBP1 starts to accumulate inside the nucleus, which depletes the RanGTP gradient across the NE, thereby leading to retention of the NTRs in the nucleus.

These results do not yet represent the most comprehensive list of CRM1 cargo atlas. Several well-established CRM1 cargoes are missing, such as Snurportin-1 and several translation initiation factors. There could be several reasons to account for this. First, nuclear accumulation of CRM1 cargoes essentially depends on the nuclear influx rate of a cargo. Even though, a protein is exported by CRM1, cytosolic retention -either being in big complex or being tethered to stable structures (e.g. cytoskeleton)- could prevent its nuclear accumulation in a short time. The prominent examples for this situation are several translation initiation factors. Majority of translation initiation factors are part of big complexes and cannot accumulate in the nucleus shortly after the LMB treatment. It was previously reported that, for several translation factors (eIF2B epsilon, eIF5 , eRF1) LMB sensitive nuclear accumulation was only achieved after tagging them with nuclear localization signal (NLS) (Bohnsack et al., 2002). Hence, the approach I presented here is more suitable for identification of CRM1 cargoes with high nuclear influx rates. Second, a drawback of the approach is lack of a selective enrichment step for the potential cargoes. Therefore, low abundant cargoes could be missed in the background proteome. However, advances in MS instrumentation are rapidly evolving and mass spectrometers are becoming more sensitive. Therefore, this drawback could be overcome by using more sensitive mass spectrometers. Third, some potential might not be expressed at all in HeLa cells. Fourth, export of the cargo might depend on certain stimuli, such as a posttranslational modification. For example, export of cyclin D1 is mediated only after its phosphorylation (Benzeno et al. 2006). Fifth, for a shuttling cargo, nuclear import might suppress the nuclear export so that complete inhibition of export might not induce adequate accumulation of the cargo in the nucleus. Finally, due to the subcellular fractionation, there could be cross-contamination between the cytosol and the nucleus.

Due to the challenges in *in silico* identification of CRM1 export cargoes; novel and complementary approaches are needed to tackle with this important biological problem. Despite several drawbacks, MS-based approach established in this thesis was validated to detect *in vivo* localization changes between the nucleus and the cytosol. It was applied to investigate CRM1 cargoes with LMB treatment and novel CRM1 cargoes were identified, as well as many known cargoes. The thorough microscopic validation strengthened the reliability of the approach. Furthermore, this approach can easily be applied to other cell lines under different growth conditions with alternative fractionation protocols to generate a more comprehensive CRM1 cargo atlas. Other nucleocytoplasmic pathways can also be studied in a similar manner to expand the cargo repertoire of other exportins or importins. Additionally, employing of this kind of an approach to investigate nucleoporins can also yield interesting results since as some nucleoporins were previously shown to affect the localization of certain proteins (Wälde et al. 2012).

5.3 Quantitative proteome analysis of *X. laevis* oocyte

Centrifugation based approaches for isolation of the nuclear and the cytosolic proteomes were challenging to obtain pure fractions in cultured cells. Hence, I shifted to a different model organism, *X. laevis*, from which I can obtain the cytosolic and the nuclear proteomes with unmatched purity. Absolute amounts of ~4400, ~4800 and ~5500 proteins in the cytosolic, nuclear, and total oocyte extracts, respectively, were estimated with spiking of protein standards with known amounts (UPS2, Sigma). In total, proteomics of *X. laevis* oocyte provided MS-based quantitative information for ~6300 proteins. This high number of identification level by was achieved by employing molecular weight based separation at the protein level and followed by trypsin digestion of proteins.

A drawback of *X. laevis* model system is the lack of well-annotated genomic data. This dramatically hampers proteomics studies, which requires well-annotated genomic sequence for faithful protein identification. In order to increase the identification level, generated MS data were searched not only against *X. laevis* database and but also against an evolutionarily close species *X. tropicalis*. During the preparation of this thesis, a study compiled an mRNA derived protein reference database for *X. laevis* and based on this database, authors reported identification of ~11000 proteins from *X. laevis* eggs (Wühr et al., 2014). This indicates that my spatial proteome analysis of *X. laevis* oocyte is not at the saturation level for protein identification. This could be due to the two reasons. First, their mRNA derived reference database is more comprehensive protein sequence database than the one I used. Authors reported that this database increased identification by up to 30% compared to conventional *X.*

X. laevis protein databases. Initial testing of this mRNA-based protein database has resulted in 15% more protein identification for my data set (data not shown). Next, it appears that I could not reach the complete depth for the protein identification. Complementary techniques can be employed to increase the level of protein identification. Instead of a fractionation at protein-level, a peptide level fractionation technique such as peptide isoelectric focusing, size exclusion or ion exchange chromatography can be employed. Alternatively, protein digestion with different endopeptidase can be performed to generate differential peptide spectrum for MS detection. An unavoidable drawback of *X. laevis* model system is the presence of high number allelic protein variants as a result of pseudotetraploid genome. This leads to sequencing of more peptides species from same protein, therefore decreasing the proteins identification level.

Quantitative analysis of *X. laevis* proteome revealed interesting observations. Only a handful of proteins dominate majority of the protein molecules in the oocyte. The 22 most abundant proteins constitute 25 % of the total protein molecules in the oocyte. A striking example is, actin, which is the most abundant protein it comprising ~2% of the total protein molecules. Additionally, ribosomal proteins cover almost 20% of the all protein molecules in the oocyte. These observations underline the fact that *X. laevis* oocyte spends most of its resources to produce many copies of a few proteins, similar to yeast and humans cells (Marguerat et al. 2012; Nagaraj et al. 2011).

The most abundant proteins in the total oocyte extract belong to ribosomal proteins, chaperones, metabolic pathway members, cytoskeleton proteins, and several RNA binding proteins. In general, ribosomes, metabolic pathway members, and cytoskeleton proteins are highly abundant across all organisms and cell types (Marguerat et al. 2012 Nagaraj et al. 2011). Still, presence of huge amounts of ribosomal proteins, metabolic pathway members may reflect oocyte's unique physiological needs. During fertilization, apart from equal number of chromosome contribution from sperm and egg, egg provides majority of the biological material, which will be needed in early embryonic development. Therefore, oocyte stockpiles huge amounts of biological material (protein, RNA, lipids) for later usage. Even though, oocyte is quite active in protein synthesis during development, it has been reported that only a portion of the ribosome is active and they are stored for later use (Taylor et al., 1985). Metabolic pathway members, particularly, glycolysis members might be needed to contribute energy requirement of the giant oocyte during oogenesis and later during embryogenesis. Highly abundant cytoskeleton proteins (several of them are actin, cofilin, gelsolin) are needed for mechanical stability of the giant oocyte.

After fertilization, egg will undergo 12 rounds of cell division without G1 phase. During these cell divisions, fertilized egg is practically silent in transcription and uses stored mRNAs for new protein synthesis (Roeder, 1974; Newport & Kirschner, 1982). The comparison of 100 most abundant *X. laevis* proteins with 100 most abundant proteins of 47 human tissues and cell lines led to detection of two RNA binding proteins that are not significantly expressed at human tissues and cell lines (Wilhelm et al. 2014). These are cold-inducible RNA-binding protein B (CIRBP) and Y-box-binding protein 2 (YBX2). Both of them are essential for embryonic development. On one hand, YBX2 binds to mRNA and marks mRNAs for cytoplasmic storage (Tafari & Wolffe, 1990). Highly abundant expression of YBX2 makes perfect biological sense for an organism to store enormous amounts of RNA. On the other hand, CIRBP may act negative regulator for adenylation of mRNAs, thereby preventing their translation (Aoki et al. 2003). This is an essential mechanism to tune translation capacity of stored mRNAs in *X. laevis* oocyte and embryos in the absence of transcription (Mendez & Richter 2001). Exportin 6 and c-Mos are prominent examples of poly(A)-tail-based translation regulation (Bohnsack et al. 2006; Gebauer et al. 1994). As a result, 100 most abundant *X. laevis* proteins are involved in common biological processes, such as protein translation, and energy production. However, presence of two highly abundant RNA binding proteins highlights that biological specificity of the oocytes can be reflected by presence of signature proteins.

Spatial proteome analysis revealed not only qualitative but also quantitative differences between the cytosolic and the nuclear proteomes. Quantitative localization distribution of the cytosolic and nuclear proteomes revealed tri-modal distribution. Two-third of the identified proteins were found in both compartments, unexpectedly only a small fraction (~500 proteins) of identified proteins had equimolar distribution in both compartments (Figure 4.10B). On the other hand, almost one-third of the identified proteins were unique to each compartment, where the nucleus contained both higher number of compartment specific and total protein species (Figure 4.10B).

The comparative analysis of 100 most abundant proteins in the cytosolic, the nuclear and the total oocyte fraction revealed dramatic proteome difference among those. Hierarchical clustering revealed four distinct proteins groups based on their abundance. Abundance based clustering indicates which molecular pathways are dominated in which compartment. The Cluster I composed of proteins, which are highly abundant in the cytosol, and to some extent (apart from Cluster I') in the nucleus (Figure 4.11A). Majority of these proteins were ribosomal proteins, glycolysis pathway members and some chaperones. Since ribosomal biogenesis starts in the nucleolus, it is expected to detect ribosomal proteins in the nucleus. Interestingly, there was a

group of proteins, which were highly abundant in all fractions (Figure 4.11A, Cluster 1'). These proteins indicate cellular processes, which are predominantly takes place both in the cytosol and in the nucleus. One of them was actin, due to the lack of Exportin 6 expression till egg stage; actin can diffuse into the nucleus, and it forms filaments to provide mechanical stability to the giant nucleus (Bohnsack et al., 2006). GTP-binding protein Ran is another high abundant which were present in almost equal concentration in the cytosol and in the nucleus. Considering essential role of the Ran in nucleocytoplasmic transport equimolar distribution of is anticipated. Another protein is ubiquitin; polyubiquitination is signal for protein degradation that I previously shown that ubiquitin proteasome system is not spatially restricted (Figure 4.18). Interestingly, there were several glycolysis pathway members and ATP producing enzymes, such as creatine kinase, nucleoside diphosphate kinase A1, phosphoglycerate kinase. High nuclear abundance of several ATP producing enzymes raises the possibility of local ATP production in the nucleus. Due to the enormous size of the nucleus, passive diffusion of ATP would be inefficient to supply all energy requirement of the oocyte. Apart from these ATP producing enzymes, several glycolysis pathway members were highly abundant in the nucleus, such alpha-enolase and transketolase.

Nuclear proteome was dominant with compartment specific proteins involved in DNA replication, chromatin organization, and RNA processing. On the other hand, the cytosol contains predominantly proteins involved in protein synthesis, like translation initiation factors, and aminoacyl tRNA synthetases. Apart from these, cytoskeleton and cytoskeleton-associated proteins, such as actin regulators and motor proteins were highly abundant. As a result, spatial proteome of *X. laevis* analysis underlines that compartmentation results in spatial distribution of distinct molecular activities.

5.4 RanGTP dependent CRM1 interactome with respect to spatial proteome of *Xenopus laevis*

5.4.1 CRM1 affinity chromatography

RanGTP dependent CRM1 affinity chromatography was employed to selectively enrich CRM1 binders from *X. laevis* oocyte extract. Subsequent to quantitative MS analysis, 632 *X. laevis* proteins were detected as significant CRM1 binders among identified ~2000 proteins. Previously, this methodology was successfully applied for identification of other exportin cargoes, such as Exportin 6, Exportin 5, Exportin t, and Exportin 7 (Mingot et al. 2004; Bohnsack et al. 2002; Bohnsack et al. 2004). However, in these studies cargo identification was restricted to visual detection of enriched bands. Recent technical, computational and experimental improvements in

MS-based workflow enabled researches not only faithful identification but also accurate quantification of thousands of proteins within hours of MS analysis time. Therefore, I did not restrict the identification of cargoes based on visual detection of enriched bands. Complete analysis of SDS-PAGE lane was resulted in identification of hundreds of potential cargoes. Additionally, a key improvement in this workflow is the considerable reduction of the unspecific background by using low affinity matrix (Figure 4.12A). Having low background on the matrix provided identification of more and even low abundant proteins. During optimization of the workflow, enrichment of other exportin cargoes was encountered. This could be due to bridging effect via FG-nucleoporins. FG-nucleoporins can interact with NTRs with RanGTP independent manner. With addition of RanGTP, endogenous exportins can form export complexes. They can indirectly interact with immobilized CRM1 via FG-nucleoporins, and this can results in enrichment of other exportin cargos on the matrix. Therefore, endogenous NTRs were depleted via phenyl sepharose matrix to prevent unspecific enrichment of other exportin cargoes. This not only eliminates false positives but also increases the available cargoes to interact with CRM1. Successfully, none of the NTRs were identified with and without RanQ69L incubation. Therefore, I am confident about not to selectively enrich other exportins cargoes.

5.4.2 Localization Profile of RanGTP dependent CRM1 interactome

X. laevis proteome showed tri-modal localization distribution between the cytosol and the nucleus. However, RanGTP dependent CRM1 interactome resulted in uni-modal positively skewed (tailed) distribution (Figure 4.12). This behavior perfectly illustrates the principle function of the CRM1, to export proteins out of the nucleus. A striking observation is that, ~17 % of the exclusively cytosolic proteins were CRM1 binders. This observation indicates that the permeability barrier of the NPC is imperfect to completely exclude the access of cytosolic proteins into the nucleus. The nuclear influx of proteins is matter of time and in sufficient time even molecules larger than 40kDa may diffuse into the nucleus. Since *X. laevis* oocyte can stay in G2-like state for months, even high molecular weight complexes (e.g. Dynactin 390 kDa) can have enough time to leak into the nucleus. Furthermore, following the nuclear envelope break down, cytoplasmic and nuclear content of higher eukaryotes are intermixed. Reformation of nucleocytoplasmic compartmentation not only depends on nuclear import of proteins, but also nuclear export of cytoplasmic proteins. Therefore, nuclear export is needed to preserve identity of the nucleus. The enormous number of potential CRM1 cargoes (~780) demonstrates that CRM1 is the principle exportin to contribute nucleocytoplasmic compartmentation.

Surprisingly, spatial proteome analysis and CRM1 affinity chromatography showed that, several CRM1 binders were predominantly or exclusively nuclear proteins, such as several components of DNA replication factor C, and RNA polymerase II (Figure 4.13A). The detection of these complexes as a CRM1 binder could be explained following reasons. First, they could be cell-regulated CRM1 cargoes. Under physiological conditions, they could be inaccessible to CRM1. They could be well missing an adaptor or a PTM that provides recognition by CRM1. This adaptor could be maintained in the cytosol or PTM could be induced via a cytoplasmic factor. Second, RanGTP dependent interaction might not necessarily be related to nuclear export as occurs during interphase. It could be due to the mitotic function of CRM1. Previously, it was documented that CRM1 partially localizes to kinetochore and centrosomes (Arnaoutov et al., 2005; Wang et al., 2005). In a RanGTP dependent manner, CRM1 recruits RanGAP1 and RanBP2 to kinetochores (Arnaoutov et al., 2005) and nucleophosmin to centrosome (Wang et al., 2005). It was proposed that, as analogy to importin β function during mitosis, CRM1 could be responsible for recruiting several proteins to kinetochore and centrosome for mitotic specific events. Alternatively, RanGTP dependent interaction could be occurring due to different reason such as activity regulation or quenching.

5.4.3 Protein Synthesis

Function of the CRM1 in protein synthesis can be considered in two aspects. First, it supplies the cytosol with essential macromolecules, which are needed for protein synthesis, such as nuclear export of 40S and 60S ribosomal subunits and signal recognition particle (SRP). 40S and 60S ribosomal subunits undergo series of maturation steps in the nucleolus and exported into the cytosol with CRM1 (Ho et al., 2000; Moy and Silver, 2002). Majority of the cytosolic ribosome components (48 out of 84) were detected as a CRM1 binder. SRP is a ribonucleoprotein complex of 6 protein and 7S RNA. It is assembled in the nucleolus and exported into cytosol to direct cotranslational translocation of proteins into ER containing signal sequence (Akopian et al., 2013). In yeast, it was shown that SRP was LMB sensitive (Grosshans et al., 2001). Here, four out of six SRP complex proteins were found as a significant binder, providing an evidence that vertebrate SRP complex could be exported with CRM1.

Second, CRM1 serves to prevent protein synthesis in the nucleus. In prokaryotes transcription and translation is coupled. This could pose serious problems in eukaryotes due to presence of intron contains genes. Hence, eukaryotes overcome this issue by physically separating these two processes with the NE. However, there were several studies claimed the presence of a nuclear translation up to 10-15 % of the cellular translation as a proof reading mechanism for mRNAs

(Iborra et al., 2001). First, for such a proof reading mechanism to happen for every mRNA, the nucleus must contain huge amount and complete set of proteins involving in the translation. In this study, I provided two lines of evidence that such a process of nuclear translation could be inefficient process. First, spatial proteome analysis of *X. laevis* oocyte showed that majority of the translation factors; aminoacyl tRNA synthetases are either exclusively or predominantly cytosolic factors (Figure 4.17). I cannot completely rule out the possibility that exclusively cytosolic (or nuclear) proteins were out of the detection limit. However, I can confidently conclude that exclusive cytosolic translation factors should be ~100 times less concentrated. Aminoacyl tRNA synthetases also had high cytoplasmic concentration implying that aminoacylation of tRNAs is an activity governed in the cytosol. Secondly, majority of the (32 out of 44) translation factors and aminoacyl tRNA synthetases, 6 out of 21 were detected as a CRM1 binder. This indicates that all these factors were actively excluded from the nucleus. Since protein synthesis is concerted action of so many factors at the same time, even presence of several factors at limiting concentrations would make nuclear translation such an inefficient process. Besides, it is possible that CRM1 would form a trimeric complex with these proteins in the presence of RanGTP in the nucleus. This could be a form of inactivation mechanism to reduce active form of these proteins in the nucleus. As a result, detection of such high number of proteins involved in protein synthesis at very low concentrations in the nucleus and as a being potential CRM1 cargoes disfavors the concept of nuclear translation.

5.4.4 Vesicle coat proteins and cytoskeleton

A function of the CRM1 is to counteract the leakage of cytosolic proteins to preserve the identity of the nucleus. This is well-illustrated with presence of so many proteins involved in cytoskeleton-based functions as a CRM1 binder, such as ARP 2/3 complex, dynactin complex, HAUS complex, septins, and many other individual proteins involving kinetochore and microtubule organizing center. It is surprising that even though, *X. laevis* oocyte contains actin in the nucleus, main actin nucleator, Arp2/3 complex is present very scarce amount in the oocyte (data not shown) and it is a CRM1 binder (Figure 4.13A). This suggests that nuclear actin filament nucleation in *X. laevis* oocyte is Arp-independent. A recent study originating from our research group identified nuclear actin interactome in *X. laevis* oocyte and detected inverted formin-2 (INF2) as an actin binder (Samwer et al., 2013). INF2 was previously reported to be a potent nuclear actin nucleator in *X. laevis* oocyte (Chhabra and Higgs, 2006).

Dynactin is multi-subunit protein complex needed for activity of the cytoplasmic dynein for various functions, such as organellar transport (Schroer, 2004). Additionally it is also participate in

some other processes such as chromosome alignment and spindle organization during mitosis (Echeverri et al., 1996). HAUS complex is eight-subunit protein complex and it also participates in mitotic spindle assembly. It co-localizes to interphase centrosomes and to mitotic spindle microtubules (Lawo et al., 2009). Another group of proteins related to cytoskeleton are septins. They are GTP binding proteins that can assemble into high-order structures such as filaments and rings (Mostowy and Cossart, 2012). Nuclear exclusion (at least temporarily) of these proteins and complexes makes perfect biological sense. Uncontrolled and ill-timed nuclear accumulation of these proteins and proteins complexes might have severe consequences; such as they might untimely initiate mitotic events, alternatively presence of them might interfere with nuclear-based functions.

Another interesting group of proteins, which were identified as CRM1 binders were vesicle coat complexes. These were COPI, COPII vesicle coat complexes and adapter protein complexes (AP-1, AP-2 and AP-3) of clathrin vesicles (Figure 4.13B). Formation of a vesicle on donor membrane starts with recruiting of GTP binding protein (ADP-ribosylation factors, abbreviated ARF) on the donor membrane. Then, vesicle coat complexes drive formation of vesicles. After pinching off the vesicle, with hydrolysis of GTP on ARFs, these coat complexes are disassembled and re-cycled for formation of new vesicles (Kirchhausen, 2000). Active nuclear exclusion of these vesicle coat complexes could be one way of suppression mechanism to prevent vesicle budding from inner nuclear membrane.

5.4.5 RNA metabolism and ubiquitin proteasome system

Two cellular processes that I investigated in detail were RNA metabolism and ubiquitin-proteasome system (UPS). RNA metabolism is a good example of a cellular process starting in the nucleus and then extending into the cytosol. Apart from TFIIIA, majority of the proteins involved in RNA synthesis (RNA polymerase components, basal transcription factors and mediator complex) had prominent or exclusive cytosolic localization (Figure 4.16). TFIIIA has dual functions in *X. laevis* oogenesis; it regulates 5S RNA transcription and associates with 5S rRNA and is involved in storage of 5S RNA as a 7S RNP in the cytosol. Exclusive cytoplasmic localization of TFIIIA was consistent with previous studies where nuclear amount of it was barely detectable (Shastry et al., 1984).

Interestingly, there were several components from the RNA polymerase II. POLR2B, POLR2C, POLR2J2 showed almost equimolar distribution and they were found as significant RanGTP dependent CRM1 binders in *X. laevis*. It is possible that these RNA polymerase components are stored in the cytosol for later usage as TFIIIA in *X. laevis*. At Human Proteome Atlas (a consortium

that generates an antibody for each protein and documented localization of proteins via immunofluorescence), small fraction of these POLR2B, POLR2C had cytosolic localization (Uhlén et al., 2005). More interestingly, one of them, POLR2B, showed cytosolic depletion upon LMB treatment in HeLa cells (Appendix 1). Given these observations, it possible that these proteins might either stored in the cytosol for later use or they might have additional role in the cytosol and in some extent CRM1 might keep these proteins cytosolic.

Similarly, huge number of the proteins involving in splicing had predominant or exclusive nuclear localization (68 proteins exclusively in the nucleus). In general, proteins participating in splicing are devoid of CRM1 interaction. Only 10 proteins out of more than 100 proteins were found to CRM1. However, some of these proteins are not primary splicing proteins; instead they have auxiliary roles in splicing. For instance CBP80, it is a component of the RNA cap binding complex (CBC), or MAGOH, it is part of the exon junction complex (EJC) (Bono et al., 2006). Overall, spatial *X. laevis* proteome suggested that proteins involved in splicing are spatially confined to nucleus and these proteins would be limiting in case of any cytoplasmic *bona fide* splicing activities.

Following RNA-based processes, mRNA surveillance, RNA export, and RNA degradation involves activity of proteins localized to both compartments (Figure 4.16). Several proteins involved in this processes were CRM1-dependent interaction partners and they were mainly predominant cytosolic proteins with several exceptions. This suggests that some steps of these pathways could be spatially restricted, and CRM1 might be responsible to maintain exclusive cytosolic localization. A prominent example were regulator of nonsense transcripts 1 and 2 (UPF1 and UPF2) which are main regulators of Non-sense mRNA decay (NMD) (Bentley, 2014).

On the other hand, ubiquitin–proteasome system (UPS) is an example of cellular process taking place in both compartments. The main player of the UPS, ubiquitin, and the single E1 enzyme (UBA1) had almost equimolar distribution across the both compartments. In general, majority of the E2 ligases were detected in both compartments (Figure 4.18). Similarly, almost equimolar distribution was seen for Cullin-RING ligase (CRL) scaffold proteins as well. Unlike to these proteins, proteins that define target specificity for ubiquitination (E3 ligases and CRL target recognizing subunits) had more distinct localization profile. Majority of the E3 ligases and CRL target recognizing subunits were compartment specific. Overall, UPS system is a cellular process that governs both in the cytosol and in the nucleus. However, since E3 ligases dictate target specificity; spatial restricted localization of E3 ligase, CRL target recognizing subunits and might have effect on their substrate recognition profile.

5.5 Comparison between LMB treatment in HeLa cells and CRM1 affinity chromatography from *X. laevis* extract

The work presented in this thesis focused on expanding of CRM1 known cargo repertoire with employing two orthogonal approaches. First, a novel MS-based approach was established to detect nuclear-to-cytosolic localization changes of proteins upon LMB treatment. Next, RanGTP dependent CRM1 interaction partners were identified via affinity chromatography from *X. laevis* oocyte extract. The former was based on identification of proteins with high nuclear influx rate upon CRM1 export inhibition. The latter involves selective enrichment of proteins interacting with CRM1 in RanGTP dependent manner. Due to different principles which these approaches rely on, different numbers of potential cargoes were identified ~140 candidates with LMB treatment from HeLa cells and ~640 candidates with affinity chromatography from *X. laevis* extract. Approximately 5 times more potential cargo was identified with affinity chromatography and overlap was 52 proteins (Figure 4.14). Despite of relatively less cargo identifications, LMB treatment represents *in vivo* experimental system. Under physiological conditions, true CRM1 cargoes would show cytoplasmic depletion and simultaneously nuclear accumulation upon LMB treatment. However, due to the presence of several limitations; i) experimental limitations in obtaining pure cytosolic and nuclear fractions, ii) MS-based instrumental limitations to identify all existing proteins, iii) low nuclear influx of export cargoes due to the cytosolic retention might have resulted in less potential cargo identification.

On the other hand, affinity chromatography captures CRM1 binders under defined *in vitro* conditions. There is two drawback of this experimental system; i) *in vitro* conditions might be non-optimum (e.g. salt concentration) for several CRM1 binders, ii) it selects high affinity binders; hence several low affinity binders could be missed. Despite being *in vitro* experimental system, due to the presence of selective enrichment of CRM1 binders and substantial reduced background (due to several reasons, low affinity matrix, CRM1 not being a limiting to capture low abundant cargoes) affinity chromatography outperformed in terms of potential cargo identification.

The observed difference in protein identification between these two approaches could be due to the experimental principles on which they rely on. One explanation for difference can be attributed to secondary effect of the LMB treatment, for instance mis-localization of several nuclear transport receptors. On the other hand, CRM1 affinity chromatography enriched several proteins might not be true CRM1 export substrates, such as exclusively nuclear localized proteins. Additionally, I cannot completely rule out the possibility that some of the enriched proteins could

be interaction partners of true CRM1 cargoes. Furthermore, organismal difference might be another reason for the observed difference. A representative example could be COP9 signalosome. This complex was depleted in the cytosolic fraction upon LMB treatment in HeLa cells, but it was not enriched with affinity chromatography from *X. laevis* extract. However, COP9 signalosome were enriched from HeLa cells as a CRM1 binder in a study originating from our research lab (Kiri). Therefore, this strongly suggests that several candidates could be due to the organismal difference.

5.6 Conclusion

Considering the essential role of the nuclear export in eukaryotic cells, expanding known cargo repertoire of most promiscuous exportin, CRM1, is significantly crucial. The work presented here concentrated on application of two orthogonal approaches to address this biological important question. First, a novel MS-based quantitative approach has been established to detect dynamic *in vivo* localization changes between the nucleus and the cytosol in HeLa cells upon LMB treatment. Successfully, this resulted in identification of many known and novel CRM1 cargoes (totally ~140), which some of them validated by microscopic analysis. Second, affinity chromatographic based approach was performed to enrich RanGTP dependent CRM1 binders from *X. laevis* oocyte extract. This led to identification of large number of CRM1 binders (~640), owing to selective enrichment. These two approaches dramatically expanded the known CRM1 cargo repertoire and the results provided valuable insights into biological implications of CRM1 dependent nuclear export. Majority of the potential CRM1 cargoes are part of the cytoplasmic activities that should be (at least temporarily) prevented in the nucleus, such as, translation, certain cytoskeleton-based activities and intracellular vesicle trafficking. Furthermore, this comprehensive catalog of candidate CRM1 cargoes offers a valuable resource for *in silico* prediction of nuclear export signals. For this purpose, categorization of candidate CRM1 cargoes into soluble protein complexes is needed to narrow down candidates of direct CRM1 binders. Furthermore, the relative abundance of ~6300 proteins were quantitatively measured in the cytosol and the nucleus, assessing the degree of compartmentation of eukaryotic cell, and the spatial distribution of three distinct molecular activities. Besides, these results revealed that distinct proteomes occupy the cytosol and the nucleus of *X. laevis* both in qualitative and quantitative manner. Additionally, spatial distribution of candidate CRM1 cargoes in the *X. laevis* oocyte showed that ~17 % exclusively cytosolic proteins are actively exported by CRM1. This suggests that the permeability barrier the NPC is alone imperfect to prevent leakage of cytosolic

proteins into nucleus and CRM1 contributes to eukaryotic compartmentation by counteracting this leakage.

6 References

- Adachi, Y., and 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. *J. Cell Biol.* *108*, 1195–1207.
- Akopian, D., Shen, K., Zhang, X., and Shan, S. (2013). Signal recognition particle: an essential protein-targeting machine. *Annu. Rev. Biochem.* *82*, 693–721.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell*.
- Aoki, K., Matsumoto, K., and Tsujimoto, M. (2003). *Xenopus* cold-inducible RNA-binding protein 2 interacts with ElrA, the *Xenopus* homolog of HuR, and inhibits deadenylation of specific mRNAs. *J. Biol. Chem.* *278*, 48491–48497.
- Argenzio, E., Bange, T., Oldrini, B., Bianchi, F., Peesari, R., Mari, S., Di Fiore, P.P., Mann, M., and Polo, S. (2011). Proteomic snapshot of the EGF-induced ubiquitin network. *Mol. Syst. Biol.* *7*, 462.
- Arnautov, A., Azuma, Y., Ribbeck, K., Joseph, J., Boyarchuk, Y., Karpova, T., McNally, J., and Dasso, M. (2005). Crm1 is a mitotic effector of Ran-GTP in somatic cells. *Nat. Cell Biol.* *7*, 626–632.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1994). *Current protocols in molecular biology* (New York: Greene Publishing Associates and Wiley-Interscience).
- Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* *15*, 163–175.
- Benzeno, S., Lu, F., Guo, M., Barbash, O., Zhang, F., Herman, J.G., Klein, P.S., Rustgi, A., and Diehl, J.A. (2006a). Identification of mutations that disrupt phosphorylation-dependent nuclear export of cyclin D1. *Oncogene* *25*, 6291–6303.
- Benzeno, S., Lu, F., Guo, M., Barbash, O., Zhang, F., Herman, J.G., Klein, P.S., Rustgi, A., and Diehl, J.A. (2006b). Identification of mutations that disrupt phosphorylation-dependent nuclear export of cyclin D1. *Oncogene* *25*, 6291–6303.
- Bischoff, F.R., and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* *354*, 80–82.
- Bischoff, F.R., Klebe, C., Kretschmer, J., Wittinghofer, A., and Ponstingl, H. (1994). RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proc. Natl. Acad. Sci. U. S. A.* *91*, 2587–2591.
- Biswas, A., Mukherjee, S., Das, S., Shields, D., Chow, C.W., and Maitra, U. (2011). Opposing action of casein kinase 1 and calcineurin in nucleo-cytoplasmic shuttling of mammalian translation initiation factor eIF6. *J. Biol. Chem.* *286*, 3129–3138.

- Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A.J.R. (2009). Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* *4*, 484–494.
- Bohnsack, M.T., Regener, K., Schwappach, B., Saffrich, R., Allee, R., and Lu, D.- (2002). Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to control protein translation to the cytoplasm. *Mol. Cell* *21*, 6205–6215.
- Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* *10*, 185–191.
- Bohnsack, M.T., Stüven, T., Kuhn, C., Cordes, V.C., and Görlich, D. (2006). A selective block of nuclear actin export stabilizes the giant nuclei of *Xenopus* oocytes. *Nat. Cell Biol.* *8*, 257–263.
- Bono, F., Ebert, J., Lorentzen, E., and Conti, E. (2006). The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* *126*, 713–725.
- Brown, D.D., and Dawid, I.B. (1968). Specific gene amplification in oocytes. Oocyte nuclei contain extrachromosomal replicas of the genes for ribosomal RNA. *Science* *160*, 272–280.
- Brown, D.D., and Gurdon, J.B. (1964). ABSENCE OF RIBOSOMAL RNA SYNTHESIS IN THE ANUCLEOLATE MUTANT OF *XENOPUS LAEVIS*. *Proc. Natl. Acad. Sci. U. S. A.* *51*, 139–146.
- Brownawell, A.M., and Macara, I.G. (2002). Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J. Cell Biol.* *156*, 53–64.
- Camier, S., Dechampesme, A.M., and Sentenac, A. (1995). The only essential function of TFIIIA in yeast is the transcription of 5S rRNA genes. *Proc. Natl. Acad. Sci. U. S. A.* *92*, 9338–9342.
- Charneau, P., Mirambeau, G., Roux, P., Paulous, S., Buc, H., and Clavel, F. (1994). HIV-1 reverse transcription. A termination step at the center of the genome. *J. Mol. Biol.* *241*, 651–662.
- Chaurand, P., Norris, J.L., Cornett, D.S., Mobley, J.A., and Caprioli, R.M. (2006). New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *J. Proteome Res.* *5*, 2889–2900.
- Chhabra, E.S., and Higgs, H.N. (2006). INF2 is a WASP homology 2 motif-containing formin that severs actin filaments and accelerates both polymerization and depolymerization. *J. Biol. Chem.* *281*, 26754–26767.
- Ciufo, L.F., and Brown, J.D. (2000). Nuclear export of yeast signal recognition particle lacking Srp54p by the Xpo1p/Crm1p NES-dependent pathway. *Curr. Biol.* *10*, 1256–1264.
- Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* *25*, 1422–1423.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.

- Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V, and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* 4, 698–705.
- Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561–563.
- David, A., Dolan, B.P., Hickman, H.D., Knowlton, J.J., Clavarino, G., Pierre, P., Bennink, J.R., and Yewdell, J.W. (2012). Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J. Cell Biol.* 197, 45–57.
- Dong, J.-M., Lau, L.-S., Ng, Y.-W., Lim, L., and Manser, E. (2009a). Paxillin nuclear-cytoplasmic localization is regulated by phosphorylation of the LD4 motif: evidence that nuclear paxillin promotes cell proliferation. *Biochem. J.* 418, 173–184.
- Dong, X., Biswas, A., Süel, K.E., Jackson, L.K., Martinez, R., Gu, H., and Chook, Y.M. (2009b). Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* 458, 1136–1141.
- Drivas, G.T., Shih, A., Coutavas, E., Rush, M.G., and D'Eustachio, P. (1990). Characterization of four novel ras-like genes expressed in a human teratocarcinoma cell line. *Mol. Cell. Biol.* 10, 1793–1798.
- Dumont, J.N. (1972). Oogenesis in *Xenopus laevis* (Daudin). I. Stages of oocyte development in laboratory maintained animals. *J. Morphol.* 136, 153–179.
- Echeverri, C.J., Paschal, B.M., Vaughan, K.T., and Vallee, R.B. (1996). Molecular characterization of the 50-kD subunit of dynactin reveals function for the complex in chromosome alignment and spindle organization during mitosis. *J. Cell Biol.* 132, 617–633.
- Edman, P. (1949). A method for the determination of amino acid sequence in peptides. *Arch. Biochem.* 22, 475.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* (80-). 246, 64.
- Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Lührmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82, 475–483.
- Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997). CRM1 Is an Export Receptor for Leucine-Rich. 90, 1051–1060.
- Gadal, O., Strauss, D., Kessl, J., Trumpower, B., Tollervey, D., and Hurt, E. (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. *Mol. Cell. Biol.* 21, 3405–3415.
- Gall, J.G. (1967). Octagonal nuclear pores. *J. Cell Biol.* 32, 391–399.

- Gebauer, F., Xu, W., Cooper, G.M., and Richter, J.D. (1994). Translational control by cytoplasmic polyadenylation of c-mos mRNA is necessary for oocyte maturation in the mouse. *EMBO J.* *13*, 5712–5720.
- Geetha, T., Vishwaprakash, N., Sycheva, M., and Babu, J.R. (2012). Sequestosome 1/p62: across diseases. *Biomarkers* *17*, 99–103.
- Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W., and Gygi, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 6940–6945.
- Gontan, C., Güttler, T., Engelen, E., Demmers, J., Fornerod, M., Grosveld, F.G., Tibboel, D., Görlich, D., Poot, R. a, and Rottier, R.J. (2009). Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *J. Cell Biol.* *185*, 27–34.
- Görlich, D., Prehn, S., Laskey, R. a, and Hartmann, E. (1994). Isolation of a protein that is essential for the first step of nuclear protein import. *Cell* *79*, 767–778.
- Görlich, D., Panté, N., Kutay, U., Aebi, U., and Bischoff, F.R. (1996). Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J.* *15*, 5584–5594.
- Görlich, D., Seewald, M.J., and Ribbeck, K. (2003). Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J.* *22*, 1088–1100.
- Grosshans, H., Deinert, K., Hurt, E., and Simos, G. (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. *J. Cell Biol.* *153*, 745–762.
- Guan, T., Kehlenbach, R.H., Schirmer, E.C., Kehlenbach, A., Fan, F., Clurman, B.E., Arnheim, N., and Gerace, L. (2000). Nup50, a nucleoplasmically oriented nucleoporin with a role in nuclear protein export. *Mol. Cell. Biol.* *20*, 5619–5630.
- Gurdon, J., Elsdale, T., and Fischberg, M. (1958). Sexually mature individuals of *Xenopus leavis* from the transplantation of single somatic nuclei. *Nature* *182*, 64–65.
- Güttler, T., and Görlich, D. (2011). Ran-dependent nuclear export mediators: a structural perspective. *EMBO J.* *30*, 3457–3474.
- Güttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., and Görlich, D. (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat. Struct. Mol. Biol.* *17*, 1367–1376.
- Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Doglio, A., Bertrand, E., Macara, I.G., and Dargemont, C. (2003). Exportin-5 mediates nuclear export of minihelix-containing RNAs. *J. Biol. Chem.* *278*, 5505–5508.
- Harhaj, E.W., and Sun, S.C. (1999). Regulation of RelA subcellular localization by a putative nuclear export signal and p50. *Mol. Cell. Biol.* *19*, 7088–7095.

- Harland, R.M., and Grainger, R.M. (2011). *Xenopus* research: metamorphosed by genetics and genomics. *Trends Genet.* *27*, 507–515.
- Harris, A., Morgan, J.I., Pecot, M., Soumare, A., Osborne, A., and Soares, H.D. (2000). Regenerating motor neurons express *Nna1*, a novel ATP/GTP-binding protein related to zinc carboxypeptidases. *Mol. Cell. Neurosci.* *16*, 578–596.
- Hellmuth, K., Lau, D.M., Bischoff, F.R., Künzler, M., Hurt, E., and Simos, G. (1998). Yeast *Los1p* has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Mol. Cell. Biol.* *18*, 6374–6386.
- Hellsten, U., Harland, R.M., Gilchrist, M.J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N.H., Shu, S., Taher, L., et al. (2010). The genome of the Western clawed frog *Xenopus tropicalis*. *Science* *328*, 633–636.
- Hershko, a, and Ciechanover, a (1998). The ubiquitin system. *Annu. Rev. Biochem.* *67*, 425–479.
- Ho, J.H., Kallstrom, G., and Johnson, A.W. (2000). *Nmd3p* is a *Crm1p*-dependent adapter protein for nuclear export of the large ribosomal subunit. *J. Cell Biol.* *151*, 1057–1066.
- Houseley, J., and Tollervey, D. (2009). The many pathways of RNA degradation. *Cell* *136*, 763–776.
- Huang, D.W., Sherman, B.T., and Lempicki, R. a (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* *4*, 44–57.
- Huang, T.T., Kudo, N., Yoshida, M., and Miyamoto, S. (2000). A nuclear export signal in the N-terminal regulatory domain of $\text{I}\kappa\text{B}\alpha$ controls cytoplasmic localization of inactive NF- κB / $\text{I}\kappa\text{B}\alpha$ complexes. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 1014–1019.
- Hutten, S., and Kehlenbach, R.H. (2006). *Nup214* is required for CRM1-dependent nuclear protein export in vivo. *Mol. Cell. Biol.* *26*, 6772–6785.
- Iborra, F.J., Jackson, D.A., and Cook, P.R. (2001). Coupled transcription and translation within nuclei of mammalian cells. *Science* *293*, 1139–1142.
- Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* *4*, 1265–1272.
- Izaurralde, E., Lewis, J., and Gamberi, C. (1995). A cap-binding protein complex mediating U snRNA export. *Nature* *376*, 709–712.
- Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I.W., and Görlich, D. (1997). The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J.* *16*, 6535–6547.
- Junttila, M.R., Puustinen, P., Niemelä, M., Ahola, R., Arnold, H., Böttzauw, T., Ala-aho, R., Nielsen, C., Ivaska, J., Taya, Y., et al. (2007). CIP2A Inhibits PP2A in Human Malignancies. *Cell* *130*, 51–62.

- Kalab, P., Weis, K., and Heald, R. (2002). Visualization of a Ran-GTP gradient in interphase and mitotic *Xenopus* egg extracts. *Science* *295*, 2452–2456.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* *40*, D109–14.
- Karas, M., and Hillenkamp, F. (1988). Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10.000 Daltons. *Anal. Chem.* *2301*, 2299–2301.
- Kato, J., and Yoneda-Kato, N. (2009). Mammalian COP9 signalosome. *Genes Cells* *14*, 1209–1225.
- Kehlenbach, R.H., Dickmanns, A., and Gerace, L. (1998). Nucleocytoplasmic shuttling factors including Ran and CRM1 mediate nuclear export of NFAT In vitro. *J. Cell Biol.* *141*, 863–874.
- Kerscher, O. (2006). Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu. Rev. Cell Dev. ...* 159–180.
- Kirchhausen, T. (2000). Three ways to make a vesicle. *Nat. Rev. Mol. Cell Biol.* *1*, 187–198.
- Kirli, K. Exploration of cargo spectrum and NES patterns recognized by the exportin CRM1.
- Klebe, C., Bischoff, F.R., Ponstingl, H., and Wittinghofer, A. (1995). Interaction of the Nuclear GTP-Binding Protein Ran with Its Regulatory Proteins RCC1 and RanGAP1. *Biochemistry* *34*, 639–647.
- Krüger, M., Moser, M., Ussar, S., Thievensen, I., Luber, C.A., Forner, F., Schmidt, S., Zanivan, S., Fässler, R., and Mann, M. (2008). SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* *134*, 353–364.
- Kudo, N., Wolff, B., Sekimoto, T., Schreiner, E.P., Yoneda, Y., Yanagida, M., Horinouchi, S., and Yoshida, M. (1998). Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1. *Exp. Cell Res.* *242*, 540–547.
- Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* *11*, 319–324.
- Kurisaki, A., Kurisaki, K., Kowanez, M., Sugino, H., Yoneda, Y., Heldin, C.-H., and Moustakas, A. (2006). The mechanism of nuclear export of Smad3 involves exportin 4 and Ran. *Mol. Cell. Biol.* *26*, 1318–1332.
- Kutay, U., and Güttinger, S. (2005). Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol.* *15*, 121–124.
- Kutay, U., Bischoff, F., Kostka, S., Kraft, R., and Gorlich, D. (1997). Export of importin a from the nucleus is mediated by a specific nuclear transport factor. *Cell* *90*, 1061–1071.
- Kutay, U., Lipowsky, G., Izaurralde, E., Bischoff, F.R., Schwarzmaier, P., Hartmann, E., and Görlich, D. (1998). Identification of a tRNA-specific nuclear export receptor. *Mol. Cell* *1*, 359–369.
- Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* *227*, 680–685.

- Larance, M., Bailly, A.P., Pourkarimi, E., Hay, R.T., Buchanan, G., Coulthurst, S., Xirodimas, D.P., Gartner, A., and Lamond, A.I. (2011). Stable-isotope labeling with amino acids in nematodes. *Nat. Methods* 8, 849–851.
- Lawo, S., Bashkurov, M., Mullin, M., Ferreria, M.G., Kittler, R., Habermann, B., Tagliaferro, A., Poser, I., Hutchins, J.R.A., Hegemann, B., et al. (2009). HAUS, the 8-subunit human Augmin complex, regulates centrosome and spindle integrity. *Curr. Biol.* 19, 816–826.
- Lipowsky, G., Bischoff, F.R., Schwarzmaier, P., Kraft, R., Kostka, S., Hartmann, E., Kutay, U., and Görlich, D. (2000). Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *EMBO J.* 19, 4362–4371.
- Liu, J.X. (2006). *Xenopus Protocols: Cell Biology and Signal Transduction* (Springer Science & Business Media).
- Liu, X., Zhang, C., Xing, G., Chen, Q., and He, F. (2001). Functional characterization of novel human ARFGAP3. *FEBS Lett.* 490, 79–83.
- Loiodice, I., Alves, A., Rabut, G., Van Overbeek, M., Ellenberg, J., Sibarita, J.-B., and Doye, V. (2004). The entire Nup107-160 complex, including three new members, is targeted as one entity to kinetochores in mitosis. *Mol. Biol. Cell* 15, 3333–3344.
- Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E.M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* 25, 117–124.
- Malim, M.H., Hauber, J., Le, S.Y., Maizel, J. V, and Cullen, B.R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* 338, 254–257.
- Malim, M.H., McCarn, D.F., Tiley, L.S., and Cullen, B.R. (1991). Mutational definition of the human immunodeficiency virus type 1 Rev activation domain. *J. Virol.* 65, 4248–4254.
- Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bähler, J. (2012). Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell* 151, 671–683.
- Maurer, P., Redd, M., Solsbacher, J., Bischoff, F.R., Greiner, M., Podtelejnikov, a V, Mann, M., Stade, K., Weis, K., and Schlenstedt, G. (2001). The nuclear export receptor Xpo1p forms distinct complexes with NES transport substrates and the yeast Ran binding protein 1 (Yrb1p). *Mol. Biol. Cell* 12, 539–549.
- McAlister, G.C., Huttlin, E.L., Haas, W., Ting, L., Jedrychowski, M.P., Rogers, J.C., Kuhn, K., Pike, I., Grothe, R.A., Blethrow, J.D., et al. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* 84, 7469–7478.
- Melchior, F., Paschal, B., Evans, J., and Gerace, L. (1993). Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. *J. Cell Biol.* 123, 1649–1659.

- Melchior, F., Sweet, D.J., and Gerace, L. (1995). Analysis of Ran/TC4 function in nuclear protein import. *Methods Enzymol.* *257*, 279–291.
- Mendez, R., and Richter, J.D. (2001). Translational control by CPEB: a means to the end. *Nat. Rev. Mol. Cell Biol.* *2*, 521–529.
- Meng, L., Zhu, Q., and Tsai, R.Y.L. (2007). Nucleolar trafficking of nucleostemin family proteins: common versus protein-specific mechanisms. *Mol. Cell. Biol.* *27*, 8670–8682.
- Michalski, A., Cox, J., and Mann, M. (2011). More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J. Proteome Res.* *10*, 1785–1793.
- Miki, T., Okawa, K., Sekimoto, T., Yoneda, Y., Watanabe, S., Ishizaki, T., and Narumiya, S. (2009). mDia2 shuttles between the nucleus and the cytoplasm through the importin- α / β - and CRM1-mediated nuclear transport mechanism. *J. Biol. Chem.* *284*, 5753–5762.
- Mingot, J.M., Kostka, S., Kraft, R., Hartmann, E., and Görlich, D. (2001). Importin 13: a novel mediator of nuclear import and export. *EMBO J.* *20*, 3685–3694.
- Mingot, J.-M., Bohnsack, M.T., Jäkle, U., and Görlich, D. (2004). Exportin 7 defines a novel general nuclear export pathway. *EMBO J.* *23*, 3227–3236.
- Mirgorodskaya, O.A., Kozmin, Y.P., Titov, M.I., Körner, R., Sönksen, C.P., and Roepstorff, P. (2000). Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (^{18}O) -labeled internal standards. *Rapid Commun. Mass Spectrom.* *14*, 1226–1232.
- Mohr, D., Frey, S., Fischer, T., Güttler, T., and Görlich, D. (2009). Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J.* *28*, 2541–2553.
- Monecke, T., Güttler, T., and Neumann, P. (2009). Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* (80-.). *69*, 1087–1091.
- Mooney, R.A. (1988). Use of digitonin-permeabilized adipocytes for cAMP studies. *Methods Enzymol.* *159*, 193–202.
- Moore, M., and Blobel, G. (1992). The two steps of nuclear import, targeting to the nuclear envelope and translocation through the nuclear pore, require different cytosolic factors. *Cell* *69*, 939–950.
- Mostowy, S., and Cossart, P. (2012). Septins: the fourth component of the cytoskeleton. *Nat. Rev. Mol. Cell Biol.* *13*, 183–194.
- Moy, T.I., and Silver, P.A. (2002). Requirements for the nuclear export of the small ribosomal subunit. *J. Cell Sci.* *115*, 2985–2995.
- Murdoch, K., Loop, S., Rudt, F., and Pieler, T. (2002). Nuclear export of 5S rRNA-containing ribonucleoprotein complexes requires CRM1 and the RanGTPase cycle. *Eur. J. Cell Biol.* *81*, 549–556.

- Murray, A.W., and Kirschner, M.W. (1989). Cyclin synthesis drives the early embryonic cell cycle. *Nature* 339, 275–280.
- Murray, M.T., Krohne, G., and Franke, W.W. (1991). Different forms of soluble cytoplasmic mRNA binding proteins and particles in *Xenopus laevis* oocytes and embryos. *J. Cell Biol.* 112, 1–11.
- Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pääbo, S., and Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* 7, 548.
- Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9, 255–262.
- Neville, M., and Rosbash, M. (1999). The NES-Crm1p export pathway is not a major mRNA export route in *Saccharomyces cerevisiae*. *EMBO J.* 18, 3746–3756.
- Nguyen, V.Q., Co, C., and Li, J.J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* 411, 1068–1073.
- Nikolov, M., Schmidt, C., and Urlaub, H. (2012). Quantitative mass spectrometry-based proteomics: an overview. *Methods Mol. Biol.* 893, 85–100.
- Nix, D.A., Fradelizi, J., Bockholt, S., Menichi, B., Louvard, D., Friederich, E., and Beckerle, M.C. (2001). Targeting of zyxin to sites of actin membrane interaction and to the nucleus. *J. Biol. Chem.* 276, 34759–34767.
- Ohno, M., Segref, a, Bachi, a, Wilm, M., and Mattaj, I.W. (2000). PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* 101, 187–198.
- Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376–386.
- Panté, N., and Kann, M. (2002). Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol. Biol. Cell* 13, 425–434.
- Paraskeva, E., Izaurralde, E., Bischoff, F.R., Huber, J., Kutay, U., Hartmann, E., Lührmann, R., and Görlich, D. (1999). CRM1-mediated recycling of snurportin 1 to the cytoplasm. *J. Cell Biol.* 145, 255–264.
- Perander, M., Bjorkoy, G., and Johansen, T. (2001). Nuclear import and export signals enable rapid nucleocytoplasmic shuttling of the atypical protein kinase C lambda. *J. Biol. Chem.* 276, 13015–13024.
- Petit, M.M., Fradelizi, J., Golsteyn, R.M., Ayoubi, T.A., Menichi, B., Louvard, D., Van de Ven, W.J., and Friederich, E. (2000). LPP, an actin cytoskeleton protein related to zyxin, harbors a nuclear export signal and transcriptional activation capacity. *Mol. Biol. Cell* 11, 117–129.
- Philpott, A., and Yew, P. (2005). The *Xenopus* cell cycle. *Cell Cycle Control* 296, 95–112.

- Philpott, A., and Yew, P.R. (2008). The *Xenopus* cell cycle: an overview. *Mol. Biotechnol.* *39*, 9–19.
- Piccotti, P., Rinner, O., Stallmach, R., Dautel, F., Farrah, T., Domon, B., Wenschuh, H., and Aebersold, R. (2010). High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat. Methods* *7*, 43–46.
- Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J. Cell Biol.* *110*, 883–894.
- Ribbeck, K., and Görlich, D. (2001). Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* *20*, 1320–1330.
- Ribbeck, K., and Görlich, D. (2002). The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* *21*, 2664–2671.
- Richards, S.A., Carey, K.L., and Macara, I.G. (1997). Requirement of guanosine triphosphate-bound ran for signal-mediated nuclear protein export. *Science* *276*, 1842–1844.
- Rogowski, K., van Dijk, J., Magiera, M.M., Bosc, C., Deloulme, J.C., Bosson, A., Peris, L., Gold, N.D., Lacroix, B., Grau, M.B., et al. (2010). A family of protein-deglutamylating enzymes associated with neurodegeneration. *Cell* *143*, 564–578.
- Romano, M., Rosanova, P., Anteo, C., and Limatola, E. (2004). Vertebrate yolk proteins: a review. *Mol. Reprod. Dev.* *69*, 109–116.
- Rout, M.P., and Blobel, G. (1993). Isolation of the yeast nuclear pore complex. *J. Cell Biol.* *123*, 771–783.
- Sambrook, J., and Russell, D.W. (2001). *Molecular Cloning: A Laboratory Manual (Third Edition)*.
- Samuels, A.L., Klinken, S.P., and Ingley, E. (2009). Liar, a novel Lyn-binding nuclear/cytoplasmic shuttling protein that influences erythropoietin-induced differentiation. *Blood* *113*, 3845–3856.
- Samwer, M., Dehne, H.-J., Spira, F., Kollmar, M., Gerlich, D.W., Urlaub, H., and Görlich, D. (2013). The nuclear F-actin interactome of *Xenopus* oocytes reveals an actin-bundling kinesin that is essential for meiotic cytokinesis. *EMBO J.* 1–17.
- Sanvito, F., Piatti, S., Villa, A., Bossi, M., Lucchini, G., Marchisio, P.C., and Biffo, S. (1999). The beta4 integrin interactor p27(BBP/eIF6) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly. *J. Cell Biol.* *144*, 823–837.
- Sasaki, T., Kojima, H., Kishimoto, R., Ikeda, A., Kunimoto, H., and Nakajima, K. (2006). Spatiotemporal regulation of c-Fos by ERK5 and the E3 ubiquitin ligase UBR1, and its biological role. *Mol. Cell* *24*, 63–75.
- Schmidt, C. (2010). Absolute and relative quantification of proteins in large protein-RNA assemblies by mass spectrometry submitted by.
- Schroer, T.A. (2004). DYNACTIN.

- Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337–342.
- Seiser, R.M., Sundberg, A.E., Wollam, B.J., Zobel-Thropp, P., Baldwin, K., Spector, M.D., and Lycan, D.E. (2006). Ltv1 is required for efficient nuclear export of the ribosomal small subunit in *Saccharomyces cerevisiae*. *Genetics* *174*, 679–691.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* *13*, 2498–2504.
- Shastry, B.S. (1996). Transcription factor IIIA (TFIIIA) in the second decade. *J. Cell Sci.* *109 (Pt 3)*, 535–539.
- Shastry, B.S., Honda, B.M., and Roeder, R.G. (1984). Altered levels of a 5 S gene-specific transcription factor (TFIIIA) during oogenesis and embryonic development of *Xenopus laevis*. *J. Biol. Chem.* *259*, 11373–11382.
- Sheets, M.D., Wu, M., and Wickens, M. (1995). Polyadenylation of c-mos mRNA as a control point in *Xenopus* meiotic maturation. *Nature* *374*, 511–516.
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* *1*, 2856–2860.
- Shin, J. (1998). P62 and the sequestosome, a novel mechanism for protein metabolism. *Arch Pharm Res* *21*, 629–633.
- Skaar, J.R., Pagan, J.K., and Pagano, M. (2013). Mechanisms and function of substrate recruitment by F-box proteins. *Nat. Rev. Mol. Cell Biol.* *14*, 369–381.
- Smith, A.E., Slepchenko, B.M., Schaff, J.C., Loew, L.M., and Macara, I.G. (2002). Systems analysis of Ran transport. *Science* *295*, 488–491.
- Soufi, B., and Macek, B. (2014). Stable isotope labeling by amino acids applied to bacterial cell culture. *Methods Mol. Biol.* *1188*, 9–22.
- Stade, K., Ford, C.S., Guthrie, C., and Weis, K. (1997). Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* *90*, 1041–1050.
- Stüven, T., Hartmann, E., and Görlich, D. (2003). Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. *EMBO J.* *22*, 5928–5940.
- Sury, M.D., Chen, J.-X., and Selbach, M. (2010). The SILAC fly allows for accurate protein quantification in vivo. *Mol. Cell. Proteomics* *9*, 2173–2183.
- Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguéz, P., Doerks, T., Stark, M., Müller, J., Bork, P., et al. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* *39*, D561–8.
- Tanaka, K., Waki, H., Ido, Y., Akita, S., and Yoshida, Y. (1988). Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-flight Mass Spectrometry. *2*, 151–153.

- Taylor, M.A., Johnson, A.D., and Smith, L.D. (1985). Growing *Xenopus* oocytes have spare translational capacity. *Proc. Natl. Acad. Sci. U. S. A.* *82*, 6586–6589.
- Team, R.D.C. (2010). R : A language and environment for statistical computing. Vienna, Austria : R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>.
- Thomas, F., and Kutay, U. (2003). Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *J. Cell Sci.* *116*, 2409–2419.
- Tomoda, K., Kubota, Y., Arata, Y., Mori, S., Maeda, M., Tanaka, T., Yoshida, M., Yoneda-Kato, N., and Kato, J. (2002a). The cytoplasmic shuttling and subsequent degradation of p27Kip1 mediated by Jab1/CSN5 and the COP9 signalosome complex. *J. Biol. Chem.* *277*, 2302–2310.
- Tomoda, K., Kubota, Y., Arata, Y., Mori, S., Maeda, M., Tanaka, T., Yoshida, M., Yoneda-Kato, N., and Kato, J. (2002b). The cytoplasmic shuttling and subsequent degradation of p27Kip1 mediated by Jab1/CSN5 and the COP9 signalosome complex. *J. Biol. Chem.* *277*, 2302–2310.
- Tunquist, B.J., and Maller, J.L. (2003). Under arrest: cytostatic factor (CSF)-mediated metaphase arrest in vertebrate eggs. *Genes Dev.* *17*, 683–710.
- Uetrecht, C., Barbu, I.M., Shoemaker, G.K., van Duijn, E., and Heck, A.J.R. (2011). Interrogating viral capsid assembly with ion mobility-mass spectrometry. *Nat. Chem.* *3*, 126–132.
- Uhlén, M., Björling, E., Agaton, C., Szigartyo, C.A.-K., Amini, B., Andersen, E., Andersson, A.-C., Angelidou, P., Asplund, A., Asplund, C., et al. (2005). A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* *4*, 1920–1932.
- Wang, W., Budhu, A., Forgues, M., and Wang, X.W. (2005). Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. *Nat. Cell Biol.* *7*, 823–830.
- Watson, M.L. (1954). Pores in the mammalian nuclear membrane. *Biochim. Biophys. Acta* *15*, 475–479.
- Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell* *82*, 463–473.
- Wild, T., Horvath, P., Wyler, E., Widmann, B., Badertscher, L., Zemp, I., Kozak, K., Csucs, G., Lund, E., and Kutay, U. (2010). A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol.* *8*, e1000522.
- Wilhelm, B.G., Mandad, S., Truckenbrodt, S., Kröhnert, K., Schäfer, C., Rammner, B., Koo, S.J., Claßen, G.A., Krauss, M., Haucke, V., et al. (2014a). Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science* *344*, 1023–1028.
- Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A.M., Lieberenz, M., Savitski, M.M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., et al. (2014b). Mass-spectrometry-based draft of the human proteome. *Nature* *509*, 582–587.
- Will, C.L., and Lührmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.* *13*, 290–301.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., and Wisniewski, J.R. (2009). Universal sample preparation method for proteome analysis Supp. Nat. Methods 6, 359–362.

Wühr, M., Freeman, R.M., Presler, M., Horb, M.E., Peshkin, L., Gygi, S.P., and Kirschner, M.W. (2014). Deep Proteomics of the *Xenopus laevis* Egg using an mRNA-Derived Reference Database. Curr. Biol. 24, 1467–1475.

Xu, D., Grishin, N. V, and Chook, Y.M. (2012). NESdb: a database of NES-containing CRM1 cargoes. Mol. Biol. Cell 23, 3673–3676.

Yang, Q., Rout, M.P., and Akey, C.W. (1998). Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. Mol. Cell 1, 223–234.

Yedavalli, V.S.R.K., Neuveut, C., Chi, Y.-H., Kleiman, L., and Jeang, K.-T. (2004). Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell 119, 381–392.

Zemp, I., Wild, T., O'Donohue, M.-F., Wandrey, F., Widmann, B., Gleizes, P.-E., and Kutay, U. (2009). Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. J. Cell Biol. 185, 1167–1180.

Zolotukhin, A.S., and Felber, B.K. (1997). Mutations in the nuclear export signal of human ran-binding protein RanBP1 block the Rev-mediated posttranscriptional regulation of human immunodeficiency virus type 1. J. Biol. Chem. 272, 11356–11360.

Appendix 1 List of proteins that were significantly depleted from the cytosol and/or accumulated in the nucleus in HeLa cells upon LMB treatment

IPI ID	Protein Name	Gene Name	Behaviour	P-value	Validation
IPI00179473	Sequestosome-1	SQSTM1	Dep.	2.26E-103	
IPI00942390	Protein diaphanous homolog 3	DIAPH3	Dep.	6.20E-38	(Miki et al., 2009)
IPI00005132	Guanine nucleotide-binding protein-like 3-like protein	GNL3L	Acc.	7.68E-27	
IPI00328987	Bystin	BYSL	Dep.	8.44E-26	
IPI00845373	Nuclear factor NF-kappa-B p100 subunit	NFKB2	Dep.	3.20E-23	(Huang et al., 2000).
-----	Rev(49-116)-GFP2-M9	Positive Control	Dep.	1.75E-20	
IPI00037599	Alpha-globin transcription factor CP2	TFCP2	Dep.&Acc.	2.30E-20	
IPI00292894	Pre-rRNA-processing protein TSR1 homolog	TSR1	Dep.	1.01E-16	
IPI00797406	60S ribosomal export protein NMD3	NMD3	Dep.&Acc.	9.97E-15	(Ho et al., 2000).
IPI00153032	Protein LTV1 homolog	LTV1	Dep.	4.82E-14	(Seiser et al., 2006) (Yeast Ortholog)
IPI00745793	G2/mitotic-specific cyclin-B1	CCNB1	Dep.	5.79E-14	Toyoshima <i>et al.</i>
IPI00385042	Nucleolar GTP-binding protein 1	GTPBP4	Acc.	3.10E-13	
IPI00867735	UPF0488 protein C8orf33	C8orf33	Acc.	4.09E-13	
IPI00010105	Eukaryotic translation initiation factor 6	EIF6	Dep.&Acc.	1.54E-11	(Biswas et al., 2011).
IPI00386448	Transcription factor p65	RELA	Dep.	2.46E-11	(Harhaj and Sun, 1999).
IPI00927731	Angio-associated migratory cell protein	AAMP	Dep.	4.65E-11	
IPI00409679	SHC-transforming protein 3	SHC3	Dep.	2.49E-10	
IPI00101186	RRP12-like protein	RRP12	Dep.	2.46E-09	
IPI00306406	Serine/threonine-protein kinase RIO2	RIO2	Dep.	2.62E-09	(Zemp et al., 2009)
IPI00402657	RNA polymerase II-associated protein 1	RPAP1	Dep.	2.69E-09	
IPI00217952	Glucosamine-fructose-6-phosphate aminotransferase [isomerizing] 1	GFPT1	Acc.	4.76E-08	
IPI00946732	Ribonucleoside-diphosphate reductase subunit M2	RRM2	Dep.	5.00E-08	
IPI00414127	Ran-specific GTPase-activating protein	RANBP1	Dep.	1.32E-07	(Zolotukhin and Felber, 1997).
IPI00215637	ATP-dependent RNA helicase DDX3X	DDX3X	Dep.	2.13E-07	(Yedavalli et al., 2004)
IPI00219575	Bleomycin hydrolase	BLMH	Acc.	4.48E-07	
IPI00305092	Partner of Y14 and mago	WIBG	Dep.	8.26E-07	
IPI00743157	Histone-lysine N-methyltransferase NSD3	WHSC1L1	Acc.	1.26E-06	
IPI00019962	UPF0534 protein C4orf43	C4orf43	Acc.	1.46E-06	
IPI00008437	Probable ribosome biogenesis protein RLP24	RPL24L	Acc.	1.54E-06	
IPI00027831	Glutamate-rich WD repeat-containing protein 1	GRWD1	Dep.&Acc.	1.67E-06	
IPI00024524	RNA-binding protein PNO1	PNO1	Dep.	2.07E-06	
IPI00014319	Influenza virus NS1A-binding protein	IVNS1ABP	Dep.	2.40E-06	
IPI00015808	Nucleolar GTP-binding protein 2	GNL2	Acc.	3.19E-06	

IPI00335437	Ankyrin repeat and zinc finger domain-containing protein 1	ANKZF1	Dep.	7.07E-06	(Samuels et al., 2009). (Mouse ortholog)
IPI00154283	Protein CIP2A	CIP2A	Dep.	8.11E-06	
IPI00016639	Protein kinase C iota type	PRKCI	Dep.	1.72E-05	(Perander et al., 2001)
IPI00031647	Programmed cell death protein 2-like	PDCD2L	Dep.&Acc.	2.81E-05	
IPI00555917	Paxillin variant	PXN	Dep.	3.52E-05	(Dong et al., 2009a)
IPI00018240	Protein SDA1 homolog	SDAD1	Acc.	4.31E-05	
IPI00171127	Ubiquitin-associated protein 2	UBAP2	Dep.	5.33E-05	
IPI00513803	Mitogen-activated protein kinase kinase kinase 2	MAP3K2	Dep.	5.41E-05	
IPI00939419	Highly similar to Notchless homolog 1	cDNA FLJ58655	Acc.	8.56E-05	
IPI00060627	Coiled-coil domain-containing protein 124	CCDC124	Dep.	9.80E-05	
IPI00966114	Survival motor neuron protein	SMN1	Dep.	0.00011913	
IPI00102096	Stromal membrane-associated protein 1	SMAP1	Dep.	0.00012846	
IPI00889000	Neurochondrin	NCDN	Dep.	0.0001719	
IPI00031651	Uncharacterized protein C7orf50	C7orf50	Acc.	0.00018072	
IPI00182180	OTU domain-containing protein 6B	OTUD6B	Dep.	0.00018493	
IPI00005780	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit	OGT	Dep.	0.00020934	
IPI00382821	Cysteine-rich and transmembrane domain-containing protein 1	C5orf32	Acc.	0.00026095	
IPI00298961	Exportin-1	XPO1	Acc.	0.0002883	
IPI00219774	cAMP-dependent protein kinase type II-alpha regulatory subunit	PRKAR2A	Acc.	0.00031557	
IPI00798041	Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	PPP2R2A	Dep.	0.00036676	
IPI00163187	Fascin	FSCN1	Acc.	0.00043747	
IPI00027717	Component of gems 4	GEMIN4	Dep.	0.0004536	
IPI00034006	Tyrosine-protein phosphatase non-receptor type 23	PTPN23	Dep.	0.00056495	
IPI00023704	Lipoma-preferred partner	LPP	Dep.	0.00058463	(Petit et al., 2000).
IPI00456262	Phosphofurin acidic cluster sorting protein 1	PACS1	Dep.	0.00064223	
IPI00301561	Thyroid receptor-interacting protein 6	TRIP6	Dep.	0.00069608	
IPI00410485	Serine/threonine-protein kinase TAO3	TAOK3	Dep.	0.00079929	
IPI00027808	DNA-directed RNA polymerase II subunit RPB2	POLR2B	Dep.	0.00082804	
IPI00642904	PABPC4 protein	PABPC4	Acc.	0.00085444	
IPI00744211	Eukaryotic translation initiation factor 4E type 2	EIF4E2	Dep.	0.00091374	
IPI00419880	40S ribosomal protein S3a	RPS3A	Acc.	0.0010838	
IPI00009958	COP9 signalosome complex subunit 5**	COPS5	Dep.	0.0011685	(Tomoda et al. 2002).
IPI00291525	Dimethyladenosine transferase 1	TFB1M	Acc.	0.0011756	
IPI00033907	Anaphase-promoting complex subunit 1	ANAPC1	Dep.	0.001699	
IPI00306127	THUMP domain-containing protein 3	THUMPD3	Dep.	0.0018264	

IPI0005904	Probable ATP-dependent RNA helicase DDX20	DDX20	Dep.	0.0020039	
IPI0009010	TRM112-like protein	TRMT112	Dep.	0.0020478	
IPI0003318	Zinc finger HIT domain-containing protein 2	ZNHIT2	Dep.	0.0021597	
IPI00103252	RWD domain-containing protein 4A	RWDD4A	Dep.	0.0021659	
IPI00514340	Methyltransferase-like protein 13	METTL13	Dep.	0.0021892	
IPI00030247	Cyclin-T1	CCNT1	Acc.	0.0022504	
IPI00290198	Interleukin-18	IL18	Dep.	0.0022604	
IPI00029601	Src substrate cortactin	CTTN	Dep.	0.0023874	
IPI00032355	Pumilio homolog 1	PUM1	Acc.	0.0027117	
IPI00647635	PERQ amino acid-rich with GYF domain-containing protein 2	GIGYF2	Dep.	0.0027142	
IPI00926625	Zyxin	ZYX	Dep.	0.0028753	(Nix et al., 2001)
IPI00002549	Anaphase-promoting complex subunit 2	ANAPC2	Dep.	0.0030246	
IPI00217862	U3 small nucleolar RNA-interacting protein 2	RRP9	Acc.	0.0030824	
IPI01014546	Protein arginine N-methyltransferase 1	PRMT1	Dep.	0.0032087	
IPI00008247	Anaphase-promoting complex subunit 5	ANAPC5	Dep.	0.003478	
IPI00844193	Zinc finger protein 593	ZNF593	Acc.	0.0034921	
IPI00410722	Uncharacterized protein C11orf48	C11orf48	Acc.	0.0037138	
IPI00008524	Polyadenylate-binding protein 1	PABPC1	Acc.	0.0043014	
IPI00238725	Coiled-coil domain-containing protein 57	CCDC57	Dep.	0.0044434	
IPI00294435	Pre-mRNA-splicing factor SLU7	SLU7	Acc.	0.0046267	
IPI00218775	FK506-binding protein 5	FKBP5	Dep.	0.0052098	
IPI00064162	Deubiquitinating protein VCIP135	VCIP1	Dep.	0.005374	
IPI00418530	Cytosolic carboxypeptidase 1	AGTPBP1	Dep.	0.0053982	
IPI00745613	Exosome complex exonuclease RRP41	EXOSC4	Acc.	0.0057676	
IPI00005822	Cell division cycle protein 23 homolog	CDC23	Dep.	0.0061449	
IPI00334190	Stomatin-like protein 2	STOML2	Acc.	0.0085422	
IPI00746351	Exosome complex exonuclease RRP44	DIS3	Acc.	0.0091731	
IPI00922751	highly similar to Solute carrier family 2, facilitated glucosetransporter member 14	SLC2A14	Acc.	0.0095978	
IPI ID	Protein Name	Gene Name	Behaviour	P-value	Validation
IPI00550021	60S ribosomal protein L3	RPL3	Acc.	9.53E-13	
IPI00220344	60S ribosomal protein L36a	RPL36A	Acc.	2.22E-06	
IPI00215719	60S ribosomal protein L18	RPL18	Acc.	1.12E-05	
IPI00555744	60S ribosomal protein L14	RPL14	Acc.	2.62E-05	
IPI00003918	60S ribosomal protein L4	RPL4	Acc.	0.00012162	
IPI00977661	60S ribosomal protein L17	RPL17	Acc.	0.00028423	
IPI00012772	60S ribosomal protein L8	RPL8	Acc.	0.0003441	
IPI00026202	60S ribosomal protein L18a	RPL18A	Acc.	0.00035221	
IPI00376798	60S ribosomal protein L11	RPL11	Acc.	0.00064352	
IPI00456758	60S ribosomal protein L27a	RPL27A	Acc.	0.00066624	
IPI00790342	60S ribosomal protein L6	RPL6	Acc.	0.00071045	

IPI00030179	60S ribosomal protein L7	RPL7	Acc.	0.00072922	
IPI00219153	60S ribosomal protein L22	RPL22	Acc.	0.0091367	
IPI00412579	60S ribosomal protein L10a	RPL10A	Acc.	3.89E-05	
IPI00027270	60S ribosomal protein L26	RPL26	Acc.	0.0013662	
IPI00031691	60S ribosomal protein L9	RPL9	Acc.	0.0015339	
IPI00414860	60S ribosomal protein L37a	RPL37A	Acc.	0.002027	
IPI00219156	60S ribosomal protein L30	RPL30	Acc.	0.0021554	
IPI00304612	60S ribosomal protein L13a	RPL13A	Acc.	0.00367	
IPI00927658	60S ribosomal protein L32	RPL32	Acc.	0.0040289	
IPI00029731	60S ribosomal protein L35a	RPL35A	Acc.	0.0045674	
IPI00947070	60S ribosomal protein L22-like 1	RPL22L1	Dep.	0.0065763	
IPI00026302	60S ribosomal protein L31	RPL31	Acc.	0.007363	
IPI00247583	60S ribosomal protein L21	RPL21	Acc.	0.0077835	
IPI00219160	60S ribosomal protein L34	RPL34	Acc.	0.0078394	
COP9 Signalosome					
IPI ID	Protein Name	Gene Name	Behaviour	P-value	Validation
IPI00018813	COP9 signalosome complex subunit 2	COPS2	Dep.	0.0067943	
IPI00018813	COP9 signalosome complex subunit 2	COPS2	Dep.	0.0067943	
IPI00018813	COP9 signalosome complex subunit 2	COPS2	Dep.	0.0067943	
IPI00301419	COP9 signalosome complex subunit 7a	COPS7A	Dep.	0.0044831	
IPI00414289	COP9 signalosome complex subunit 1	GPS1	Dep.	0.0024419	
IPI00171844	COP9 signalosome complex subunit 4	COPS4	Dep.	0.0022753	
IPI00163230	COP9 signalosome complex subunit 6	COPS6	Dep.	0.0017123	
IPI00009958	COP9 signalosome complex subunit 5	COPS5	Dep.	0.0011685	(Tomoda et al., 2002b)
IPI00009480	COP9 signalosome complex subunit 8	COPS8	Dep.	0.00044628	
Nucleocytoplasmic transport related					
IPI ID	Protein Name	Gene Name	Behaviour	P-value	Validation
IPI00477040	Nucleoporin NUP188 homolog	NUP188	Dep.	2.00E-11	
IPI00748807	Nuclear pore complex protein Nup160	NUP160	Dep.	4.74E-07	
IPI00012578	Importin subunit alpha-4	KPNA4	Dep.	8.27E-07	
IPI00002214	Importin subunit alpha-2	KPNA2	Dep.	1.45E-06	
IPI01015268	Importin subunit alpha-7	KPNA6	Dep.	2.87E-05	
IPI00299033	Importin subunit alpha-3	KPNA3	Dep.	2.95E-05	
IPI00007401	Importin-8	IPO8	Dep.	0.00014127	
IPI00298961	Exportin-1	XPO1	Acc.	0.0002883	

IPI00303292	Importin subunit alpha-1	KPNA1	Dep.	0.00042629	
IPI00024364	Transportin-1	TNPO1	Dep.	0.00093646	
IPI00398009	Importin-4	IPO4	Dep.	0.0019125	
IPI00397904	Nuclear pore complex protein Nup93	NUP93	Dep.	0.0073096	
IPI00477040	Nucleoporin NUP188 homolog	NUP188	Dep.	2.00E-11	
IPI00748807	Nuclear pore complex protein Nup160	NUP160	Dep.	4.74E-07	

Appendix 2 List of proteins that were significantly depleted/enriched in the total cell extract in HeLa cells upon LMB treatment

IPI ID	Protein Name	Gene Name	Behaviour	P-value
IPI00015953	Nucleolar RNA helicase 2	DDX21	Enriched	1.45E-43
IPI00007797	Fatty acid-binding protein, epidermal	FABP5	Enriched	4.91E-25
IPI00479217	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	Enriched	2.73E-23
IPI00101186	RRP12-like protein	RRP12	Enriched	2.78E-22
IPI00396378	Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1	Enriched	5.08E-18
IPI00789551	Putative uncharacterized protein MATR3	MATR3	Enriched	1.55E-17
IPI00215965	Heterogeneous nuclear ribonucleoprotein A1	HNRNPA1	Enriched	5.62E-13
IPI00003377	Splicing factor, arginine/serine-rich 7	SFRS7	Enriched	7.82E-10
IPI00641829	Spliceosome RNA helicase BAT1	BAT1	Enriched	2.02E-09
IPI00012382	U1 small nuclear ribonucleoprotein A	SNRPA	Enriched	1.21E-08
IPI00011274	Heterogeneous nuclear ribonucleoprotein D-like	HNRPDL	Enriched	1.91E-08
IPI00742682	Nucleoprotein TPR	TPR	Enriched	3.15E-08
IPI00219038	Histone H3.3	H3F3A	Enriched	6.53E-07
IPI00477313	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRNPC	Enriched	2.75E-06
IPI00301936	cDNA FLJ60076, highly similar to ELAV-like protein 1	ELAVL1	Enriched	3.93E-06
IPI00008708	Ribosomal L1 domain-containing protein 1	RSL1D1	Enriched	3.99E-06
IPI00784224	Zinc finger RNA-binding protein	ZFR	Enriched	4.31E-06
IPI00419373	Heterogeneous nuclear ribonucleoprotein A3	HNRNPA3	Enriched	5.23E-06
IPI00302850	Small nuclear ribonucleoprotein Sm D1	SNRPD1	Enriched	1.09E-05
IPI00011913	Heterogeneous nuclear ribonucleoprotein A0	HNRNPA0	Enriched	2.11E-05
IPI00017963	Small nuclear ribonucleoprotein Sm D2	SNRPD2	Enriched	2.58E-05
IPI00844578	ATP-dependent RNA helicase A	DHX9	Enriched	3.76E-05
IPI00028911	Dystroglycan	DAG1	Enriched	4.07E-05

IPI00418471	Vimentin	VIM	Enriched	4.08E-05
IPI00154590	MKI67 FHA domain-interacting nucleolar phosphoprotein	MKI67IP	Enriched	4.78E-05
IPI00016334	Cell surface glycoprotein MUC18	MCAM	Enriched	5.96E-05
IPI00029267	U2 small nuclear ribonucleoprotein B''	SNRPB2	Enriched	8.38E-05
IPI00021405	Lamin-A/C	LMNA	Enriched	9.04E-05
IPI00413611	DNA topoisomerase 1	TOP1	Enriched	9.07E-05
IPI01012315	PTGFRN protein	PTGFRN	Enriched	9.13E-05
IPI00022048	Prostaglandin F2 receptor negative regulator	PTGFRN	Enriched	0.000115
IPI00297477	U2 small nuclear ribonucleoprotein A'	SNRPA1	Enriched	0.000129
IPI00171903	Heterogeneous nuclear ribonucleoprotein M	HNRNPM	Enriched	0.000129
IPI00028931	Desmoglein-2	DSG2	Enriched	0.000152
IPI00010204	Splicing factor, arginine/serine-rich 3	SFRS3	Enriched	0.00017
IPI00000041	Rho-related GTP-binding protein RhoB	RHOB	Enriched	0.000311
IPI00013877	Heterogeneous nuclear ribonucleoprotein H3	HNRNPH3	Enriched	0.000385
IPI00220609	Nucleoporin SEH1-like	SEH1L	Enriched	0.00046
IPI00007928	Pre-mRNA-processing-splicing factor 8	PRPF8	Enriched	0.000493
IPI00420014	U5 small nuclear ribonucleoprotein 200 kDa helicase	SNRNP200	Enriched	0.000529
IPI00014474	A-kinase anchor protein 8	AKAP8	Enriched	0.000839
IPI00977773	cDNA FLJ59472, highly similar to Tripeptidyl-peptidase 1 (EC3.4.14.9)	TPP1	Enriched	0.000994
IPI00654555	NOP2 protein	NOP2	Enriched	0.001083
IPI00940237	ATP-dependent RNA helicase DDX39	DDX39	Enriched	0.001085
IPI00018195	Mitogen-activated protein kinase 3	MAPK3	Enriched	0.001414
IPI00032827	Pre-mRNA branch site protein p14	SF3B14	Enriched	0.001498
IPI00604620	Nucleolin	NCL	Enriched	0.001865
IPI00334713	Heterogeneous nuclear ribonucleoprotein A/B	HNRNPAB	Enriched	0.002022
IPI00004290	Digestive organ expansion factor homolog	DEF	Enriched	0.002182
IPI00002519	Serine hydroxymethyltransferase, cytosolic	SHMT1	Enriched	0.002365
IPI00879750	Small nuclear ribonucleoprotein Sm D3	SNRPD3	Enriched	0.002422
IPI00479191	Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	Enriched	0.002516
IPI00291200	Nuclear pore complex protein Nup133	NUP133	Enriched	0.002672
IPI00876941	Caveolin	CAV1	Enriched	0.002823
IPI00910816	cDNA FLJ59238, highly similar to SNW domain-containing protein 1	SNW1	Enriched	0.002842
IPI00027834	Heterogeneous nuclear ribonucleoprotein L	HNRNPL	Enriched	0.002851
IPI00290110	Programmed cell death protein 4	PDCD4	Enriched	0.002981
IPI00031697	Transmembrane protein 109	TMEM109	Enriched	0.004213
IPI00640525	Carrier family 6 , member 8 variant	PPGB	Enriched	0.004231
IPI00796835	Intron-binding protein aquarius	AQR	Enriched	0.004326
IPI00465294	Cell division cycle 5-like protein	CDC5L	Enriched	0.005704
IPI00003309	DNA-directed RNA polymerases I, II, and III subunit RPABC3	POLR2H	Enriched	0.005835
IPI00006379	Nucleolar protein 58	NOP58	Enriched	0.006179
IPI00163391	Putative methyltransferase METT10D	METT10D	Enriched	0.006312

IPI00028888	Heterogeneous nuclear ribonucleoprotein D0	HNRNPD	Enriched	0.006703
IPI00059292	Protein mago nashi homolog 2	MAGOHB	Enriched	0.007931
IPI00024524	RNA-binding protein PNO1	PNO1	Enriched	0.008081
IPI00300127	N-acetyltransferase 10	NAT10	Enriched	0.008101
IPI00006620	Monocarboxylate transporter 7	SLC16A6	Enriched	0.009264
IPI00290204	U1 small nuclear ribonucleoprotein 70 kDa	SNRNP70	Enriched	0.09274
IPI00034006	Tyrosine-protein phosphatase non-receptor type 23	PTPN23	Depleted	5.6983E-09
IPI00010105	Eukaryotic translation initiation factor 6	EIF6	Depleted	8.6405E-09
IPI00011118	Ribonucleoside-diphosphate reductase subunit M2	RRM2	Depleted	3.1402E-08
IPI00927731	Angio-associated migratory cell protein	AAMP	Depleted	1.3862E-06
IPI00790503	Myosin-10	MYH10	Depleted	0.000010252
IPI00414860	60S ribosomal protein L37a	RPL37A	Depleted	0.000022197
IPI00385042	Nucleolar GTP-binding protein 1	GTPBP4	Depleted	0.000031465
IPI00179473	Sequestosome-1	SQSTM1	Depleted	0.000034601
IPI00641181	MARCKS-related protein	MARCKSL1	Depleted	0.00023619
IPI00003918	60S ribosomal protein L4	RPL4	Depleted	0.00047464
IPI00328118	Sperm-associated antigen 5	SPAG5	Depleted	0.00053002
IPI00014400	DnaJ homolog subfamily B member 12	DNAJB12	Depleted	0.00053821
IPI00166010	CCR4-NOT transcription complex subunit 1	CNOT1	Depleted	0.00060002
IPI00555744	60S ribosomal protein L14	RPL14	Depleted	0.00060314
IPI00016802	NAD-dependent deacetylase sirtuin-1	SIRT1	Depleted	0.00064492
IPI00027831	Glutamate-rich WD repeat-containing protein 1	GRWD1	Depleted	0.00065386
IPI00465361	60S ribosomal protein L13	RPL13	Depleted	0.00067496
IPI00030179	60S ribosomal protein L7	RPL7	Depleted	0.00068216
IPI00018196	Notchless protein homolog 1	NLE1	Depleted	0.00073918
IPI00412579	60S ribosomal protein L10a	RPL10A	Depleted	0.00082489
IPI00977912	60S ribosomal protein L36a	RPL36A	Depleted	0.0010053
IPI00306043	YTH domain family protein 2	YTHDF2	Depleted	0.0011536
IPI00470528	60S ribosomal protein L15	RPL15	Depleted	0.0014519
IPI00914529	cDNA FLJ57954, highly similar to 60S ribosomal protein L28	RPL28	Depleted	0.0014932
IPI00790342	60S ribosomal protein L6	RPL6	Depleted	0.0022477
IPI00010153	60S ribosomal protein L23	RPL23	Depleted	0.0025651
IPI00026202	60S ribosomal protein L18a	RPL18A	Depleted	0.0026554
IPI00015905	Exosome complex exonuclease RRP4	EXOSC2	Depleted	0.0027018
IPI01020905	60S ribosomal protein L18	RPL18	Depleted	0.0029715
IPI00299573	60S ribosomal protein L7a	RPL7A	Depleted	0.0030588
IPI00011253	40S ribosomal protein S3	RPS3	Depleted	0.0033832
IPI00025329	60S ribosomal protein L19	RPL19	Depleted	0.0035507
IPI00306332	60S ribosomal protein L24	RPL24	Depleted	0.0036235
IPI00170972	UPF0553 protein C9orf64	C9orf64	Depleted	0.0039927
IPI00304612	60S ribosomal protein L13a	RPL13A	Depleted	0.0041142
IPI00215637	ATP-dependent RNA helicase DDX3X	DDX3X	Depleted	0.0041729
IPI00294610	DnaJ homolog subfamily A member 3, mitochondrial	DNAJA3	Depleted	0.0046503

IPI00219160	60S ribosomal protein L34	RPL34	Depleted	0.0048881
IPI00247583	60S ribosomal protein L21	RPL21	Depleted	0.0052293
IPI00219156	60S ribosomal protein L30	RPL30	Depleted	0.0055755
IPI00169400	28S ribosomal protein S5, mitochondrial	MRPS5	Depleted	0.0057698
IPI00021537	Opioid growth factor receptor	OGFR	Depleted	0.0058033
IPI00947070	60S ribosomal protein L22-like 1	RPL22L1	Depleted	0.005911
IPI00004459	Probable dimethyladenosine transferase	DIMT1L	Depleted	0.0062645
IPI00022373	RNA-binding protein NOB1	NOB1	Depleted	0.0063292
IPI00216237	60S ribosomal protein L36	RPL36	Depleted	0.0066395
IPI00456758	60S ribosomal protein L27a	RPL27A	Depleted	0.0066579
IPI00027032	Acyl-protein thioesterase 2	LYPLA2	Depleted	0.0072736
IPI00021389	Copper chaperone for superoxide dismutase	CCS	Depleted	0.0075411
IPI00002186	Brefeldin A-inhibited guanine nucleotide-exchange protein 2	ARFGEF2	Depleted	0.0075965
IPI00018098	Pre-mRNA-splicing factor 38B	PRPF38B	Depleted	0.0079721
IPI00646899	19 kDa protein	19 kDa protein	Depleted	0.0082803
IPI00910088	cDNA FLJ14622	AP1M1	Depleted	0.0083172
IPI00927658	60S ribosomal protein L32	RPL32	Depleted	0.0084463
IPI00029656	Non-syndromic hearing impairment protein 5	DFNA5	Depleted	0.0084469
IPI00550021	60S ribosomal protein L3	RPL3	Depleted	0.0089571
IPI00418248	cDNA FLJ56351, highly similar to Mps one binder kinase activator-like 2	MOB2	Depleted	0.009235
IPI00604599	Transmembrane emp24 domain-containing protein 3	TMED3	Depleted	0.0092422
IPI00023832	SH3 and PX domain-containing protein 2B	SH3PXD2B	Depleted	0.09048

Appendix 3 Protein IDs that were significantly detected as a RanGTP dependent CRM1 interaction partners from *X. laevis*

gi 113971	gi 80476543	gi 147900237	gi 148222186	gi 148229563	gi 148236444	gi 351542203
gi 116163	gi 80477771	gi 147900404	gi 148222264	gi 148229675	gi 148236529	gi 428697039
gi 119140	gi 83318456	gi 147900456	gi 148222316	gi 148229797	gi 148236643	gi 512809468
gi 129700	gi 89267912	gi 147900496	gi 148222523	gi 148229808	gi 148236849	gi 512809930
gi 134026	gi 102991831	gi 147900544	gi 148222563	gi 148229874	gi 148236867	gi 512810033
gi 138532	gi 109893256	gi 147900604	gi 148222968	gi 148229929	gi 148236875	gi 512811987
gi 214027	gi 110645447	gi 147900612	gi 148222997	gi 148230036	gi 148237044	gi 512813733
gi 464252	gi 110645571	gi 147900624	gi 148223299	gi 148230116	gi 148237173	gi 512814118
gi 1095557	gi 111185522	gi 147900626	gi 148223435	gi 148230292	gi 148237404	gi 512814579
gi 1169421	gi 111305473	gi 147900702	gi 148223451	gi 148230354	gi 148237426	gi 512814896
gi 1220554	gi 111305973	gi 147900740	gi 148223485	gi 148230517	gi 148237458	gi 512815274
gi 1350785	gi 111306191	gi 147900810	gi 148223495	gi 148230653	gi 148237470	gi 512817311
gi 1439611	gi 111309077	gi 147900951	gi 148223501	gi 148230675	gi 148237484	gi 512819522
gi 1582214	gi 111598547	gi 147901017	gi 148223539	gi 148230709	gi 148237502	gi 512820710

gi 1890239	gi 112419036	gi 147901031	gi 148223553	gi 148230743	gi 148237649	gi 512821260
gi 11493738	gi 112419059	gi 147901095	gi 148223601	gi 148230798	gi 148237753	gi 512821422
gi 12004634	gi 112419083	gi 147901107	gi 148223685	gi 148230847	gi 148237846	gi 512822546
gi 12313579	gi 112419355	gi 147901141	gi 148223736	gi 148230885	gi 148237916	gi 512823295
gi 18157370	gi 113931396	gi 147901155	gi 148223922	gi 148231023	gi 148237966	gi 512825542
gi 27503841	gi 113931440	gi 147901420	gi 148223932	gi 148231213	gi 148238016	gi 512827234
gi 27769162	gi 114107782	gi 147901482	gi 148224144	gi 148231392	gi 148238092	gi 512827934
gi 27881813	gi 114107786	gi 147901600	gi 148224387	gi 148231408	gi 148238223	gi 512828182
gi 27882190	gi 114107812	gi 147901618	gi 148224397	gi 148231524	gi 150416158	gi 512829494
gi 27882192	gi 114107881	gi 147901729	gi 148224411	gi 148231554	gi 153791802	gi 512835953
gi 27924365	gi 114107983	gi 147901896	gi 148224469	gi 148231583	gi 158254143	gi 512840815
gi 27924367	gi 114108009	gi 147902004	gi 148224598	gi 148231603	gi 158254243	gi 512843052
gi 28277340	gi 114108106	gi 147902111	gi 148224642	gi 148231631	gi 161169050	gi 512843377
gi 28386066	gi 114108292	gi 147902158	gi 148224658	gi 148231645	gi 161611524	gi 512844000
gi 28422157	gi 115174281	gi 147902248	gi 148224668	gi 148231847	gi 161611735	gi 512844424
gi 28422233	gi 115292090	gi 147902254	gi 148224768	gi 148231875	gi 161611793	gi 512846207
gi 29126980	gi 115312897	gi 147902284	gi 148224852	gi 148231929	gi 163915671	gi 512851345
gi 38174108	gi 115527861	gi 147902441	gi 148224857	gi 148231939	gi 163916440	gi 512853326
gi 38174128	gi 115527893	gi 147902509	gi 148224874	gi 148232020	gi 165970918	gi 512855355
gi 38174419	gi 115528674	gi 147902633	gi 148224927	gi 148232048	gi 165971165	gi 512856132
gi 38174727	gi 115528688	gi 147902637	gi 148225212	gi 148232062	gi 166796313	gi 512862061
gi 38181707	gi 115528802	gi 147902718	gi 148225274	gi 148232090	gi 169642646	gi 512865119
gi 38494291	gi 116063374	gi 147903018	gi 148225316	gi 148232132	gi 170284630	gi 512866731
gi 38511935	gi 116063392	gi 147903409	gi 148225470	gi 148232140	gi 170285117	gi 512866845
gi 38648975	gi 116063404	gi 147903453	gi 148225598	gi 148232198	gi 183985620	gi 512871512
gi 38649187	gi 116063410	gi 147903585	gi 148225648	gi 148232220	gi 183986332	gi 512871922
gi 39794491	gi 116063428	gi 147903617	gi 148225695	gi 148232254	gi 183986364	gi 512877650
gi 39795834	gi 116063456	gi 147903721	gi 148225744	gi 148232272	gi 183986453	gi 512879178
gi 39849973	gi 116284131	gi 147903791	gi 148225746	gi 148232323	gi 187469513	gi 512881957
gi 39850034	gi 116487464	gi 147903819	gi 148225809	gi 148232383	gi 189441581	gi 512882882
gi 45708834	gi 117557968	gi 147903835	gi 148225885	gi 148232439	gi 189441629	A0A8M2
gi 45709737	gi 118763560	gi 147903912	gi 148225893	gi 148232630	gi 189441658	A0JMU3
gi 46250287	gi 118763640	gi 147904080	gi 148225911	gi 148232661	gi 189441783	A1DPK7
gi 46362426	gi 119352603	gi 147904100	gi 148225959	gi 148232750	gi 189441952	A1DPK8
gi 47123230	gi 120537910	gi 147904174	gi 148226062	gi 148232754	gi 189442595	A1DPL3
gi 47125185	gi 123892785	gi 147904248	gi 148226152	gi 148232768	gi 195539778	A1L3E3
gi 47937658	gi 123896325	gi 147904296	gi 148226246	gi 148232878	gi 195540121	A1L3J7
gi 49115369	gi 123898906	gi 147904320	gi 148226493	gi 148232948	gi 197245526	A2BDB7
gi 49115756	gi 123899794	gi 147904334	gi 148226573	gi 148232966	gi 197246573	A2RV61
gi 49115770	gi 123900457	gi 147904425	gi 148226749	gi 148233066	gi 197246697	A2RV65
gi 49119460	gi 123903246	gi 147904539	gi 148226825	gi 148233094	gi 197246701	A2RV69
gi 49250390	gi 124107585	gi 147904649	gi 148226837	gi 148233173	gi 197246801	A2VD00
gi 49250519	gi 133737022	gi 147904708	gi 148226986	gi 148233274	gi 206984811	A2VD56
gi 49250860	gi 134023835	gi 147904714	gi 148227020	gi 148233332	gi 206984831	A2VD92

gi 49256167	gi 134024250	gi 147904776	gi 148227097	gi 148233378	gi 209969740	A2VD97
gi 49257246	gi 134024331	gi 147904798	gi 148227244	gi 148233470	gi 212276428	A2VDA5
gi 49257921	gi 134025743	gi 147904874	gi 148227248	gi 148233630	gi 213623790	A3KMI0
gi 49522851	gi 134025950	gi 147904918	gi 148227437	gi 148233734	gi 213624174	A4FVE0
gi 49899997	gi 134025964	gi 147905097	gi 148227449	gi 148233872	gi 213625406	A5D8N1
gi 49900031	gi 134254220	gi 147905185	gi 148227482	gi 148233978	gi 213625562	A5D8N2
gi 49903584	gi 134254226	gi 147905211	gi 148227542	gi 148234316	gi 213626632	A5D8P6
gi 49903747	gi 134254231	gi 147905321	gi 148227614	gi 148234336	gi 213627354	A5H447
gi 50370198	gi 134254286	gi 147905372	gi 148227640	gi 148234401	gi 224495945	A5J090
gi 50414778	gi 138519822	gi 147905374	gi 148227694	gi 148234413	gi 259469342	A5PKQ6
gi 50416640	gi 138519856	gi 147905558	gi 148227710	gi 148234435	gi 284507299	A6H8I5
gi 51258855	gi 140833060	gi 147905658	gi 148227760	gi 148234472	gi 284521569	A7E222
gi 51258906	gi 146327079	gi 147905704	gi 148227794	gi 148234565	gi 291290931	A8E5Z0
gi 51703760	gi 147898538	gi 147905764	gi 148227832	gi 148234611	gi 301603698	A9JS75
gi 51703787	gi 147898570	gi 147905844	gi 148227862	gi 148234660	gi 301605058	A9JS80
gi 51704122	gi 147898931	gi 147905967	gi 148227874	gi 148234819	gi 301605382	A9JS90
gi 51950233	gi 147899079	gi 147905979	gi 148227978	gi 148234849	gi 301605686	A9ULY1
gi 52354820	gi 147899223	gi 147906013	gi 148228018	gi 148235048	gi 301607017	B1H1P6
gi 54311344	gi 147899475	gi 147906132	gi 148228028	gi 148235554	gi 301607598	B1H1Q0
gi 56269257	gi 147899497	gi 147906584	gi 148228088	gi 148235596	gi 301611338	B1H1O8
gi 56540896	gi 147899539	gi 147906655	gi 148228126	gi 148235681	gi 301612969	B1WBA6
gi 56972258	gi 147899668	gi 147906757	gi 148228261	gi 148235725	gi 301614708	B1WBD3
gi 60552042	gi 147899700	gi 147906867	gi 148228370	gi 148235727	gi 301614835	B2GTY0
gi 60618407	gi 147899826	gi 147906994	gi 148228420	gi 148235751	gi 301622003	B3DLL9
gi 60649740	gi 147899972	gi 147907002	gi 148228553	gi 148235805	gi 301622111	B7ZPG0
gi 62531233	gi 147900009	gi 147907300	gi 148228663	gi 148235853	gi 301624752	B7ZR49
gi 62858085	gi 147900065	gi 148222005	gi 148228673	gi 148235959	gi 313747477	B7ZR96
gi 62859383	gi 147900069	gi 148222043	gi 148228859	gi 148235981	gi 349585162	B7ZRW4
gi 62859649	gi 147900121	gi 148222063	gi 148228908	gi 148236035	gi 349585432	B9UR70
gi 66269747	gi 147900167	gi 148222140	gi 148229130	gi 148236153	gi 350529312	C5H606
gi 67677978	gi 147900179	gi 148222178	gi 148229521	gi 148236235	gi 350606386	E3WDQ9
gi 68533910	gi 147900211	gi 148222182	gi 148229525	gi 148236355	gi 351542175	B7ZR20
gi 77748167						

Curriculum Vitae

Personal details

Name: Samir Karaca
Date of birth: 20 August 1985
Place of birth: Kardzhali, Bulgaria
Nationality: Turkish and Bulgarian

Education

09.2010 - 09.2014 Doctoral studies in Bioanalytical Mass Spectrometry Group,
Title: MS-based quantitative analysis of the CRM1 export pathway
and spatial proteomics of the *Xenopus laevis* oocyte
IMPRS Molecular Biology, Göttingen
Max Planck Institute for Biophysical Chemistry, Göttingen
(Germany) and Georg-August-Universität Göttingen (Germany)

10.2009-09.2010 MSc study in Molecular Biology
IMPRS Molecular Biology, Göttingen (Germany)

09.2005-07.2009 BSc in Molecular Biology and Genetics
Middle East Technical University, Ankara (Turkey)

Laboratory Skills

- Quantitative proteomics approaches: SILAC, iTRAQ, label-free quantification.
- Hands on experience with LC-MS systems; e.g. Orbitrap Velos, Orbitrap Q-exactive.
- MS data analysis and bioinformatics: MaxQuant, Perseus, Mascot, MassMatrix, Cytoscape, R scripting and bioinformatics tools (IPA, KEGG, GO).
- Protein and peptide characterization: SDS- PAGE, peptide isoelectric focusing, PTM analysis, chromatographic separations of peptides and proteins.
- Recombinant protein expression & purification, and mammalian cell culture.

Professional Activities

- Co-organizer, 8th and 9th International Horizons in Molecular Biology Symposium, October 2012, Göttingen, Germany.
 - Responsible for communication with speakers, moderating the sessions, logistics.
- Co-organizer, "Bridging the Cultures: Turkish Culture Night", IMPRS in Molecular Biology, April 2010, Göttingen, Germany.
- Supervisor, supervision of rotation students, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany, 2011&2013.
 - Designing and mentoring research project.

Free-time Activities:

Passionate swimmer and runner:

- Cross-Continental Swimming Race, Bosphorus-Istanbul, 2013, Turkey.
- Strongman Run, Fisherman's Friend, 2014, Ferropolis, Germany.
- Half- Marathon runner, TUI Marathon, 2014, Hannover, Germany.

Awards & Distinctions:

- Scholarship of the Excellence Foundation for Promotion of Max Planck Society, 2009-2011.
- Graduated as an honor student, METU, 2004-2009, Turkey.
- Republic of Turkey General Directorate of Higher Education Credit and Hostels Institution Scholarship, 2004-2009, Turkey.
- Mehmet Zorlu Association Higher Education Scholarship, 2004-2009, Turkey.

Publications

Thakar K*, Karaca S* , Port SA, Urlaub H, Kehlenbach RH. "Identification of CRM1-dependent nuclear export cargoes using quantitative mass spectrometry." *Molecular & Cellular Proteomics* 2013; 12 (3), 664-678.

*Equal contribution

Kimura M, Thakar K, Karaca S, Imamoto N, Kehlenbach RH. "Novel approaches for the identification of nuclear transport receptor substrates." *Methods in Cell Biology* 2013; 122, 353-378.

Becker J, Barysch SV, Karaca S, Dittner C, Hsiao HH, Diaz MB, Herzig S, Urlaub H, Melchior F. "Detecting endogenous SUMO targets in mammalian cells and tissues." *Nature Structural & Molecular Biology* 2013; 20 (4), 525-531

Wirth M, Karaca S, Wenzel D, Ho L, Tishkoff D, Lombard DB, Verdin E, Urlaub H, Jedrusik-Bode M, Fischle W. "Mitochondrial SIRT4-type proteins in *Caenorhabditis elegans* and mammals interact with pyruvate carboxylase and other acetylated biotin-dependent carboxylases." *Mitochondrion* 2013; 13 (6), 705-720

