

Identification of nuclear export signals and structural analysis of nuclear transport complexes

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

Chandini Kadian

Born in Sunabeda, India

August 6th, 2012

I hereby declare that I prepared the PhD thesis “**Identification of nuclear export signals and structural analysis of nuclear transport complexes**” on my own and with no other sources and aids than quoted.

Chandini Kadian

I dedicate this work to my parents

TABLE OF CONTENTS

Abstract	viii
Abbreviations.....	x
1 Introduction	
1.1 Subcellular compartmentalization	1
1.2 The nuclear pore complex.....	2
1.3 Nuclear Transport Receptors:.....	3
1.4 Selective transport across the NPC.....	3
1.5 Nuclear transport depends on the RanGTP gradient across the nuclear envelope.....	4
1.6 RanGTP, RanBP1 and RanGAP.....	5
1.7 Role of NTF2.....	5
1.8 Nuclear GEF RCC1 regenerates RanGTP in the nucleus.....	6
1.9 Nuclear import.....	7
1.10 Nuclear export.....	7
1.11 Recognition of export Signals.....	8
1.12 Nuclear export by CRM1.....	10
1.13 Recognition of export signals by CRM1:.....	10
1.14 Identification of classical NES in different cargoes for export by CRM1 using bioinformatic tools.....	12
1.15 Contribution of the spacer residues in the NES consensus.....	14
1.16 Nuclear export of translation factors.....	15
1.17 Nuclear exclusion of eIF2β by CRM1.....	16
1.18 Nuclear export of RanGAP1 by CRM1.....	17
1.19 Structural analysis of nuclear transport receptors.....	19
1.20 Exportin 4 and its cargo	20
1.21 Eukaryotic initiation factor 5A (eIF5A).....	20
1.22 Smad 3.....	23
1.24 SRY.....	25

2 Materials and methods

2.1 Chemicals used	27
2.2 Instruments used.....	27
2.3 Software	28
2.4 Bacterial growth media - E.coli culture growth media.....	29
2.5 Buffers.....	29
2.6 Protocol for production of electro-competent E.coli cells.....	31
2.7 LB plates to grow colonies of transformed bacteria.....	32
2.8 Bacterial strains used	32
2.9 PCR.....	33
2.10 Mutagenesis.....	34
2.11 Colony PCR.....	35
2.12 Primer design.....	35
2.13 DNA digestion.....	35
2.14 DNA agarose gel electrophoresis.....	36
2.15 Purification of DNA fragments excised from agarose gel.....	36
2.16 DNA ligation.....	36
2.17 Electroporation.....	37
2.18 Mini or midi preps for DNA stocks.....	37
2.19 DNA sequencing.....	37
2.20 Protein expression.....	37
2.21 Protein purification.....	38
2.22 Binding assay.....	39
2.23 Transfections.....	40
2.24 Confocal microscopy.....	42

3 Results

3.1 eIF2 β forms an export complex with CRM1.....	43
3.1.1 Location of eIF2 β export signal using bioinformatic servers...	46
3.2 Does eIF2 β interact with CRM1 in the binding cleft?	46
3.3 Fragment mapping of eIF2 β to locate the residues that contain the export signal for recognition by CRM1.....	48

3.4 Deciphering the NES binding sequence in eIF2 β	56
3.5 Comparison of eIF2 β NES in different species: modifying the <i>Homo sapien</i> NES to match the EIF2 β NES from <i>Xenopus tropicalis</i> ..	64
3.6 Alanine ϕ 1 plays an active role in modulating the eIF2 β NES affinity for CRM1.....	66
3.7 Analysis of the PKI NES ϕ residue mutants: introducing atypical residues at the ϕ positions in the NES.....	66
3.8 Analysis of the atypical ϕ 1 Alanine residue of the eIF2 β NES: understanding which residues can be permitted in this position and allow NES activity.....	70
3.9 Analysis of the spacer residues in the eIF2 β consensus.....	75
3.10 RanGAP 1	98
3.11 NES prediction of the RanGAP 1 sequence using bioinformatic tools.....	99
3.12 Locating the NES by generated truncations of RanGAP 1 and testing for interaction with CRM1 and RanGTP.....	101
3.13 Exportin 4 and its cargo.....	107
3.14 eIF5A1.....	109
3.15 Sox 2 and SRY.....	115
3.16 Co-expression trials.....	119
3.17 SRY.....	120
3.18 Synthetic genes optimised for expression in bacteria.....	121
3.19 Smad3.....	122
3.20 The expression of the HMG box domain of Sox 2 and SRY.....	124

4 Discussion

4.1 Nuclear export by CRM1.....	126
4.2 eIF2 β	128
4.2.1 eIF2 β interacts with CRM1 in the hydrophobic cleft and contains a classical PKI type leucine rich NES.....	129
4.2.2 Deciphering the eIF2 β NES.....	129

4.2.3 An atypical ϕ 1 hydrophobic residue was discovered in the eIF2 β NES.....	130
4.2.4 Defining the spacer residues in the eIF2 β NES.....	131
4.3 RanGAP 1.....	133
4.3.1 Bioinformatic tools to predict the RanGAP 1 NES.....	134
4.3.2 Trials to locate an NES containing fragment of RanGAP 1.....	134
4.4 Structural analysis of Exportin 4 and its cargo.....	136
4.4.1 Hypusination strategy for eIF5A1.....	137
4.4.2 Expression and purification of Smad 3.....	138
4.4.3 Isolating a stable HMG domain of Sox 2 and SRY for interaction with Exportin 4.....	139

5 Appendix:

5.1 Figures A1, A2 and A3: Tandem purification of dual tagged Sox 2 to isolate full length protein.....	141
---	-----

Bibliography.....	144
Acknowledgement.....	153
Curriculum Vitae.....	154

Abstract

The nucleus is a vital sub cellular compartment and it stores the genome. Transport across the nuclear envelope occurs through Nuclear Pore Complexes (NPC) and cargo traffic across the NPCs is a highly specific and controlled process. Facilitated transport processes are executed by Nuclear Transport Receptors (NTR). The goals of my doctoral work have been divided in two broad themes, which have been summarized below.

Identification of Nuclear Export Signals for CRM1 mediated export.

The nucleus is specialised in transcription and the cytosol is specialised in translation. Nuclear translation is not permitted, because translation of unspliced transcripts could lead to a formation of misfolded or possibly lethal proteins. Exclusion of translation factors is one of the methods employed by the cell to prohibit nuclear translation. eIF2 β is a translation factor and it was assumed that it can passively diffuse across the permeability barrier owing to its small size. CRM1 has been shown to exclude eIF2 β and Leptomycin B inhibition of CRM1 caused rapid accumulation of eIF2 β in the nucleus. The rate of accumulation of eIF2 β was too fast to be sustained by passive diffusion and suggested active import. Thus CRM1 could be correcting this accidental import of eIF2 β . The exact recognition motif of eIF2 β that is recognised by CRM1 was unknown. A major focus of this work was dedicated to deciphering this export signal. NES prediction tools were employed to detect a Leucine rich classical NES in eIF2 β . However, no positive hits were detected. This hinted that either the prediction was imprecise, or, opened the possibility that eIF2 β possessed a novel recognition motif for export by CRM1. An unbiased approach was taken to find regions of eIF2 β that could contain the export signal and it was identified to be a PKI type NES with an atypical Φ 1 residue. The NES spacer residues were analysed and it was shown that they play a vital role in recognition of the NES by CRM1. The range of residues permitted at these spacer positions was defined. The PKI type NES consensus was examined and the spacer residue adjacent to the last Φ position of the NES was also shown to be important for CRM1

interaction. We also addressed the issue of incorporation of atypical residues at Φ positions of the NES. This information can be extrapolated to refine the algorithms used to detect classical Leucine rich NESs in CRM1 cargoes.

The directionality of transport across the nuclear membrane depends on the RanGTP gradient. This gradient is maintained by the concerted action of cytosolic proteins RanGAP1, RanBP1 and nuclear GEF - RCC1. RanGAP1 also plays a role in disassembly of export complexes as they exit the nucleus. It can enter the nucleus accidentally and CRM1 is responsible for nuclear exclusion, however the export signal was unknown in this case as well. Prediction of the NES using bioinformatic tools does not suffice since there is a strong tendency for false positives. This occurs because RanGAP1 has Leucine rich repeats in its N terminal domain. Thus, an unbiased approach was taken to identify the NES and we have located some sequences with potential to be NESs in the Leucine Rich Repeats at the N terminal region of RanGAP 1. Another goal is to understand how CRM1 can export RanGAP1 in a manner such that RanGAP1 does not disassemble its own export complex by helping in hydrolysing RanGTP. To answer this, the aim is to crystallize the RanGAP1-RanGTP-CRM1 export complex.

Structural analysis of Exportin 4 transport complexes.

There are some remarkable NTRs, which can perform both import and export functions - e.g. Exportin 4 has been shown to export eIF5A1 and Smad 3 and import Sox2 and SRY. We wanted to elucidate the mechanism of action of Exportin 4 by crystallizing it in its export and import conformation. By seeing the crystal structure of the export complex, one can also understand the influence of RanGTP on the structure and consequently, function of Exportin 4. The major challenge of the project was to get stable cargo proteins. Sox 2 and SRY degraded heavily and eIF5A is the only known protein to have a post translational modification called Hypusine. Through this work, we worked to develop an effective strategy for complete hypusination of eIF5A1 and established purification of stable HMG domains of Sox 2 and SRY. The next steps will be aimed at isolating stable and homogenous complexes of Exportin 4 in both its import and export conformation for crystallographic trials.

Abbreviations

CAS	cellular a poptosis susceptibility (= Exportin 2)
C-terminus	carboxy-terminus
CRM1	ch romosomal r egion m aintenance 1 (= Exportin 1/ Xpo1p)
DTT	d ithiothreitol
eIF2 β	e ukaryotic translation i nitiation f actor 2 β
eIF5A	e ukaryotic translation i nitiation f actor 5A
E. coli	Escherichia coli
EDTA	e thylene d iamine t etraacetic a cid
Exp	exportin
FG repeat	phenylalanine/glycine repeat
GAP	G TPase- a ctivating p rotein
GDP	g uanosine 5'- d iphosphate
GFP	g reen f luorescent p rotein
GST	g lutathione S -transferase
GTP	g uanosine 5'- t riphosphate
GTPase	GTP hydrolase
HEAT repeat	class of protein repeats (h untingtin, e longation factor 3, the PR65/A subunit of protein phosphatase 2A (PP2A), lipid kinase TOR)
HIV	h uman i mmunodeficiency v irus
HMG	h igh- m obility g roup
hnRNP	h eterogeneous n uclear r ibonucleoprotein
H. sapiens	Homo sapiens
IgG	immunoglobulin G
Imp	importin
IPTG	isopropyl- β -D-thiogalactopyranoside
LB	L uria- B ertani (lysogeny broth, medium)
LMB	l eptomycin B
K _d	dissociation constant
kDa	k ilodalton
MDa	m egadalton
N-terminus	amino-terminus
NE	n uclear e nvelope
NES	"Leucine-rich" n uclear e xport s ignal
NLS	n uclear l ocalization s ignal
NMR	n uclear m agnetic r esonance
NPC	n uclear p ore c omplex
NTA	n itrilotriacetic a cid
NTF2	n uclear t ransport f actor 2
NTR	n uclear t ransport r eceptor
Nup	n ucleoporin (NPC protein)
OD	o ptical d ensity
PBS	p hosphate- b uffered s aline
PEG	p olyethylene g lycol
PKA	p rotein k inase A (cAMP-dependent protein kinase)
PKI	p rotein k inase A i nhibitor
pre-miRNA	pre-micro RNA
Ran	R as-related n uclear a ntigen
RanBD	Ran- b inding d omain
RanBP	Ran- b inding p rotein
RanGAP 1	RanGTPase- a ctivating p rotein
RanGEF	Ran g uanine nucleotide e xchange f actor (RCC1)
Ras	R at s arcoma
RCC1	r egulator of c hromosome c ondensation 1 ("RanGEF")
RNP	r ibonucleoprotein (particle)
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>

SDS-PAGE

S. pombe

SPN1

SV40

TEV

Tris

UsnRNP

v/v

w/v

X. laevis

YT / 2YT

z (in "zz")

sodium **d**odecyl sulfate-**p**olyacrylamide **g**el
electrophoresis

Schizosaccharomyces pombe

Snurportin1

simian **v**irus 40

tobacco **e**tch **v**irus

Tris(hydroxymethyl)aminomethane, 2-amino-2-
hydroxymethyl-1,3-propanediol

uridine-rich **s**mall **n**uclear **r**ibonucleo**p**rotein
(particle)

volume per volume

weight per volume

Xenopus laevis

yeast extract **t**ryptone medium

IgG-binding domain of the *Staphylococcal* protein A

1 Introduction

1.1 Subcellular compartmentalization:

A remarkable feature of eukaryotic cells is the presence of different sub-cellular compartments. This segregation serves the purpose of maintaining the internal environment specific to each compartment by controlling the contents present (Leighton et al., 1968). This feature gives eukaryotic cells an added level of complexity with regard to regulation of the various cellular processes (reviewed in Güttler and Görlich, 2011).

The nucleus is one of the largest compartments of the cell and it stores the genome (Kaffman and O'Shea, 1999). Within the nucleus, genes are transcribed into pre-mRNA that undergo splicing to generate mature mRNA. Splicing is a process by which non-coding sequences, so called “introns”, are removed from coding sequences, called “exons”, in various patterns by the spliceosome. According to this pattern, multiple transcripts can be generated from a single gene. (Hastings & Krainer, 2001, Brow et al., 2002).

Fully mature transcripts exit the nucleus and they undergo a process of surveillance for premature stop codons by Nonsense Mediated mRNA Decay in the cytosol. This ensures that only fully spliced transcripts are being translated (Maquat LE. 2004). Thus, in contrast to prokaryotes, transcription and translation are spatially segregated by the nuclear envelope (discussed in Bohnsack et al., 2002).

Taking into consideration the various functions of the Nuclear Envelope, it has been suggested that it has played a significant role in evolution and in increasing the diversity of proteins being synthesised by enabling splicing, controlling access of transcription factors to genes, complex regulation of multiple processes and allowing storage of much more genetic material (reviewed in Güttler & Görlich, 2011 and Conti & Izaurralde, 2001; Gilbert, 1978).

Constitutive and regulated traffic of several macromolecules occurs across the nuclear membrane. Occasionally, macromolecules without a nuclear destination enter the nucleus accidentally e.g. during cell division or due to passive diffusion because of small size. These molecules need to be constantly exported out of the nucleus. One example of this is the exclusion of translation factors from the nucleus. (Bohnsack et al., 2002).

Exchange of macromolecules across the nuclear envelope occurs through the nuclear pore complexes (NPCs) (Feldherr, 1962) and includes different proteins like e.g. transcription factors, protein complexes like the ribosomal proteins, proteins involved in the splicing machinery, mRNA and tRNA (reviewed in Görlich & Kutay, 1999; Cook A et al., 2007; Kuersten et al., 2001; Macara et al., 2001; Paine and Feldherr, 1972).

1.2 The nuclear pore complex:

The nuclear envelope has pores with a special membrane topology. These pores are stabilized by Nuclear Pore Complexes (Bahr and Beermann, 1954) and allow selective passage of macromolecules across the nuclear envelope. The NPCs consist of so called Nucleoporins (Nups), which interact with each other in a specific manner and so far, about 30 different Nucleoporins have been found (Rout et al., 2000; Cronshaw et al., 2002). These Nucleoporins form a highly evolved complex, which repeats in an eight-fold rotational symmetry in the NPC (Gall, 1967). The approximate size of the NPC in higher eukaryotes is 125MDa (Reichelt et al., 1990).

Nucleoporins or Nups can be divided into 3 categories. There are membrane integral Nups, Nups that form the rigid scaffold and FG domain containing Nups. (Cronshaw et al., 2002, Hetzer et al., 2005). The FG domains are required to allow selective transport across the Nuclear Pore Complex and this will be discussed later. (Yamada et al., 2010; Frey et al., 2006; Patel et al., 2007; Frey and Görlich, 2009; Ader et al., 2010; Yamada et al., 2010, Iovine et al., 1995; Radu et al., 1995; Bayliss et al., 1999, Bayliss et al., 2000).

1.3 Nuclear Transport Receptors:

Transport of different cargo molecules across the NPC is mediated by soluble receptors called Nuclear Transport Receptors (NTRs). These belong to the Importin β superfamily (reviewed in Güttler and Görlich, 2011 and Cook et al., 2007).

Nuclear Transport Receptors (NTRs) have a unique and conserved architecture - however at the sequence level conservation is low. They are large proteins and range from approximately 90 -150 kDa in molecular weight (Görlich et al., 1997; reviewed in Güttler and Görlich, 2011; reviewed in Cook et al., 2007)

NTRs are built from HEAT repeat motifs that consists of two alpha helices - stacked anti-parallel to each other. The two helices, A and B, are connected by a loop and placed at a certain angle to each other. These helices are found all along the NTRs and form a continuous structure without sub-domains. Due to the continuous arrangement of HEAT repeats, the NTRs have an overall 'twist' in structure and contribute to the intrinsic flexibility of NTRs. (Cook and Conti, 2010; Güttler and Görlich, 2011).

1.4 Selective transport across the NPC:

Different models have been proposed to explain the selective transport across the NPC. It has been suggested that the FG domains interact with each other to form a sieve like structure that has a size exclusion limit of 40 kDa or 5 nm diameter for macromolecules. Thus making the transport process across the NPC very selective. (Mohr et al., 2009; Ribbeck & Görlich, 2001; Frey & Görlich, 2006; Frey et al., 2007). Macromolecules smaller than the size of 5nm in diameter can go across the NPC in an unregulated manner passively (Mohr et al., 2009). Transport of macromolecules larger than 5nm diameter or 40kDa is heavily retarded.

Larger macromolecules are carried across the NPC by NTRs. NTRs allow their facilitated transport across the NPC at a much higher rate (Ribbeck and Görlich, 2001; Mohr et al., 2009; Frey et al., 2007, reviewed in Fried and Kutay, 2003).

Thus, FG repeats bind NTRs required for facilitated translocation and formation of the passive permeability barrier (Iovine et al., 1995; Radu et al., 1995; Bayliss et al., 2000, reviewed in Güttler and Görlich, 2011) (fig 1).

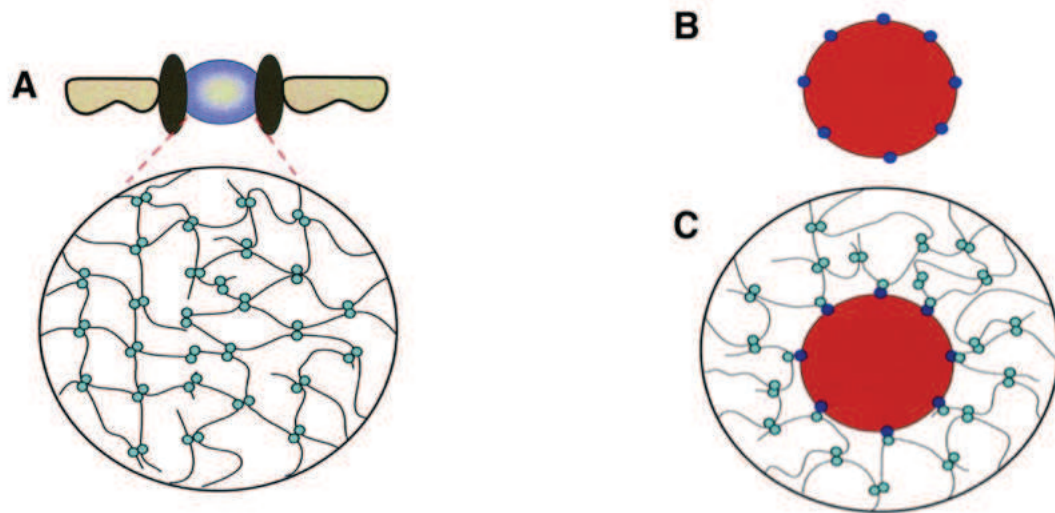


Figure 1: Nuclear transport receptors crossing the mesh-like permeability barrier (Ribbeck & Görlich, 2001).

A - The Permeability Barrier

B - cargo with interaction points on its surface by virtue of which it can interact with the FG repeats

C - The cargo becomes a part of the barrier by interacting with the FG repeats as it crosses it.

1.5 Nuclear transport depends on the RanGTP gradient across the nuclear envelope:

The directionality of NTR-mediated transport depends on the RanGTP gradient across the nuclear envelope (Görlich et al., 1996, Izaurralde E, 1997). The RanGTPase system maintains this gradient and it includes critical components namely - RanGTP, RCC1, RanGAP and RanBP1 (Görlich et al., 2003). This RanGTP gradient fuels directed cargo transport through NPCs, even against a gradient of chemical activity.

1.6 RanGTP, RanBP1 and RanGAP:

RanGTP is a small GTPase of the Ras superfamily (Moore MS, Blobel G; 1994; Avis JM, Clarke PR ,1996; Bischoff and Ponstingl, 1991; Melchior et al., 1993).

RanBP1 and RanGAP1 are found in the cytosol. RanBP1 is the accessory protein that helps in the function of RanGAP1 (Bischoff and Görlich, 1997). RanGAP1 activates RanGTP and this causes the hydrolysis of RanGTP to RanGDP. (Scheffzek et al., 1995; Partridge and Schwartz, 2009; Vetter et al., 1999).

A crystal structure of Ran-RanBP1-RanGAP gave further insight into the disassembly procedure of export complexes (Seewald MJ et al., 2002). RanGAP associates with RanGTP directly via its Leucine Rich Repeats at the N terminus (Seewald MJ et al., 2002).

1.7 Role of NTF2:

The Nuclear Transport Factor 2 (NTF2) and RanGEF (RCC1) help regenerating a high concentration of RanGTP within the nucleus and therefore maintain the RanGTP gradient across the nuclear envelope (fig 2) (Ribbeck et al., 1998; Bischoff and Ponstingl, 1991; Bischoff and Ponstingl, 1995, Smith et al., 1998). NTF2 binds to RanGDP in the cytosol and transports it to the nucleus across the NPC (Ribbeck et al., 1998).

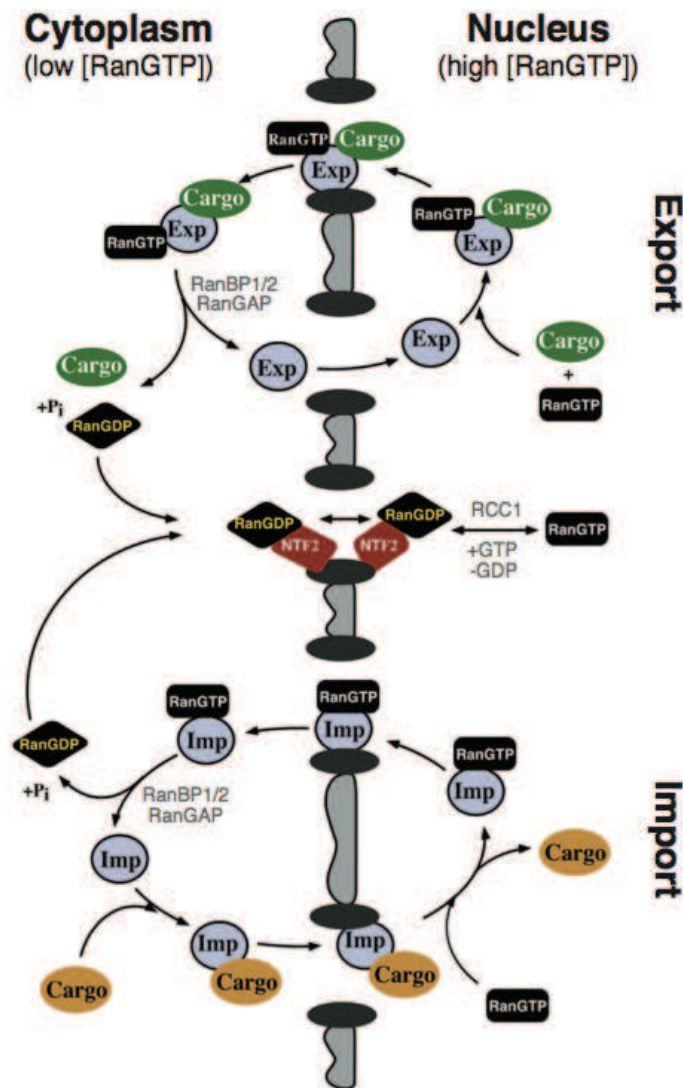


Figure 2: Nuclear import and export; (Görlich & Kutay, 1999).

A simplistic view of the export and import cycles across the NPC. The export complex is formed in the nucleus with a high concentration of RanGTP. It travels to the cytoplasm and it disassembles because of RanGTP hydrolysis to RanGDP. This hydrolysis is achieved by the action of RanBP1 and RanGAP. The import complex is formed in the cytosol and travels to the nucleus where it disassembles because of the presence of RanGTP. RanGTP bound to Importins recycles to the cytosol. Here, RanGTP gets hydrolysed and the Importin is free to bind other cargo. NTF2 carries RanGDP to the nucleus and RCC1 regenerates RanGTP in the nucleus.

1.8 Nuclear GEF RCC1 regenerates RanGTP in the nucleus:

In the nucleus, RCC1 - also known as regulator of chromosome condensation 1 - performs the function of a RanGEF (Guanine Exchange Factor) (Bischoff and Ponstingl, 1991). RCC1 is found immobilized to histones and converts RanGDP to

RanGTP in the nucleus (Bischoff and Ponstingl, 1991; Ohtsubo et al., 1989). Thus, RanGTP gradient is established by a concerted action of nuclear GEF RCC1 (Nemergut et al., 2001), RanGAP1, RanBP1 and NTF2 (Bischoff & Ponstingl, 1991; Bischoff et al., 1995; Ribbeck et al., 1998; Görlich et al., 1996; Izaurralde et al., 1997).

1.9 Nuclear import:

The cargo and its respective Importin form a binary complex in the cytosol that travels across the NPC into the nucleus. In the nucleus itself, RanGTP interacts with the import complex and disassembles it. RanGTP and the Importin form a complex and the respective cargo is released. Then the Importin and RanGTP travel across the NPC to the cytosol where the Importin and RanGTP complex disassemble. The importin is recycled and participates in another round of import (Rexach and Blobel 1995; Görlich et al., 1996; reviewed in Cook et al., 2007; Görlich & Kutay, 1999) (Fig 2).

Nuclear Transport receptors that import cargo recognise specific signals (Dingwall et al., 1982) or structured motifs. The common feature of Nuclear Localisation Signals (NLSs) is the presence of exposed positively charged residues. For instance, the classical NLS represents a short Lysine-rich sequence. Import signals can be monopartite, which contain a single cluster of basic residues - as seen in SV40 Large T antigen (Kalderon et al., 1984)- and bipartite. Bipartite NLSs consist of two clusters of basic residues separated by approximately ten uncharged residues – as seen in nucleoplasmin (Dingwall et al., 1988, Robbins et al., 1991).

1.10 Nuclear export:

Cargo that travels from the nucleus to the cytosol, interacts with its respective Exportin in the presence of RanGTP. This leads to the formation of a trimeric complex, which consists of the exportin bound to its cargo and RanGTP (Kutay et al., 1997; Fornerod et al., 1997). The trimeric complex then travels across the NPC and reaches the cytosol. RanBP1 and RanGAP1 are found in the cytosol and they help in the disassembly of the export complex (Richards et al., 1996, Izaurralde et al., 1997). Following this, the cargo is released and the exportin recycles back to the nucleus to

start over again (Bischoff and Görlich, 1997; reviewed in Kutay and Görlich, 1999; reviewed in Güttler and Görlich, 2011; reviewed in Cook et al., 2007) (fig 2).

Exportins recognise specific nuclear export signals in their cargo. This can be a particular sequence of residues, a three dimensional fold and may even involve adaptor proteins (Fornerod et al., 1997; Fukuda et al., 1997; Ossareh-Nazari et al., 1997). The work in this thesis is focused on nuclear export and some of the identified exportins and their cargo have been shown in Table 1.

1.11 Recognition of export Signals:

The export cargo contain specific 'signals' which are recognised by the respective exportins. This could be a linear sequence or a three dimensional fold. For instance, CAS recognises Importin α and CRM1 recognises Snuporin by a three dimensional fold. CRM1 can also recognise a simple linear sequence, as seen in the classical PKI type Nuclear Export Signal. Nucleic acids are also recognised, as seen in export of tRNA by Exportin t and Exportin 5 (reviewed in Güttler and Görlich, 2011). In some cases, an adaptor is required for the interaction of the cargo with its respective exportin. There are a few cargoes of CRM1 that require adaptors for export. Few examples have been listed here. E.g. This is seen in nuclear export of RNA genome of viruses using adaptor HIV 1 Rev. Another example is when CRM1 exports U snRNAs using adaptors PHAX and the CBC complex (Ohno et al., 2000; Izaurralde et al., 1995). Thus exportins may differ with respect to the signals and cargo they export and in the conformation the export complex.

Exportin	Cargo	Adaptor	References
CRM1 (Exportin 1) <i>Xpo1/Kap124p</i>	Leu rich NE cargoes Translation factors eIF2Bε, eRF1, eEF1B β, eEF2, eIF2 complex and all its subunits, the eIF3 complex, eIF5 SPN1 HIV genomic RNA m7G capped UsnRNAs 60S pre ribosomal subunits	HIV Rev PHAX + CBC NMD3	Fischer et al., 1995; Wen et al., 1995; Fukuda et al., 1997; Stade et al., 1997; Fornerod et al., 1997 Bohnsack et al., 2002 Paraskeva et al., 1999 Fischer et al., 1995 Izaurrealde et al., 1995; Ohno et al., 2000 Ho et al., 2000; Thomas and Kutay 2003
CAS (Exportin 2) <i>*Cse1/Kap109p</i>	Importin αs <i>Srp1p/Kap60p</i>		Kutay et al., 1997; Solsbacher et al., 1998
Exportin t (Xpo t) <i>*Los1p/Kap127p</i>	tRNA		Kutay et al., 1998; Arts et al., 1998; Hellmuth et al., 1998
Exportin 5 (Xpo 5)	tRNA, eEF1A (via aa-tRNA) dsRNA binding proteins (via ds RNA) pre miRNA 60S pre ribosomal subunits		Bohnsack et al., 2002; Calado et al., 2002 Brownawell and Macara (2002) Yi et al, 2003; Bohnsack et al., 2004; Lund et al., 2004 Wild et al., 2010
Exportin 6	Actin profilin complexes		Stüven et al., 2003
Exportin 7	p50RhoGAP, 14-3-3σ		Mingot et al., 2004
Importin 13	Import: Mago Y14, Ubc9 histone fold heterodimers Export:eIF1A		Mingot et al., 2001; Kahle et al., 2005; Walker et al., 2009 Mingot et al., 2001
Exportin 4	Import: Sox2, SRY Export:eIF5A, Smad3		Gontan et al., 2009 Lipowski et al., 2000; Kurisaki et al., 2006
<i>*Msn5p/Kap142p</i>	Import: RPA Export: tRNA, dsRNA, phosphorylated Pho4p		Yoshida and Blobel 2001 Kaffman et al., 1998; Shibata et al., 2006

Table 1: List of exportins and their respective cargo (adapted from Güttler and Görlich, 2011).
(*yeast orthologs and their cargo are marked in blue).

1.12 Nuclear export by CRM1

CRM 1 - also known as Exportin 1 - is a nuclear transport receptor with an extremely wide range of export cargo (Fornerod et al., 1997; reviewed in Güttler and Görlich, 2011). It was identified in *Schizosaccharomyces pombe* and was shown to be important for nuclear export of cargo (Kudo et al., 1998; Stade et al., 1997; Adachi and Yanagida et al., 1989). The HIV Rev NES and the PKI NES serve as prototypic examples of NESs recognised by CRM1 (Fischer et al., 1995; Wen et al., 1995). It exports cargo of a varied nature, e.g. many proteins of different sizes and structure, including various viral particles (e.g. the HIV-1 genomic RNA using HIV-Rev as an adaptor), both small and large ribosomal subunits, various translation factors (e.g. eIF2 β) and RanBP1 (Fornerod et al., 1997; Bohnsack et al., 2002; Fischer et al., 1995; Richards et al., 1996; Conti and Izaurralde, 2001). The crystal structure of CRM 1 bound to Snuportin and RanGTP has been solved and this gives valuable insights into the export mechanism of CRM1 (Monecke et al., 2009).

RanGTP binds on the inner surface (Monecke et al., 2009) and perhaps strains CRM1 and as a consequence, CRM1 adopts its nuclear conformation. This leads to opening of the hydrophobic cleft where the Nuclear Export Signal (NES) can bind. Hence, it is suggested that the binding of CRM1 to its cargo occurs by an allosteric mechanism, which involves cooperative binding of RanGTP and cargo. Thus, the strain caused by binding of RanGTP is relieved by binding of cargo in the hydrophobic cleft (Monecke et al., 2009). This mechanism for CRM1 export of its cargo varies in comparison to Exportin t or CAS for instance (Cook et al., 2009; Matsuura and Stewart, 2004). In the latter, RanGTP is in direct contact to the transport receptor and the cargo. These interactions contribute to the stability of the export complex by virtue of the binding energies released (Monecke et al., 2009; reviewed in Cook et al., 2007; Matsuura and Stewart, 2004, reviewed in Güttler and Görlich, 2011).

1.13 Recognition of export signals by CRM1:

CRM1 recognises its cargo via specific export signals. Initially this export signal was defined to contain the consensus: Φ 1 x(2-3) Φ 2x(2-3) Φ 3x Φ 4, where Φ represents a hydrophobic residue, x represents the residues that intersperse the Φ positions and are

referred to as the spacer residues (Kutay & Guttinger, 2005; Fischer et al., 1995). Based on previous work, Leucine was shown to be favoured for the Φ positions, but after systematic mutagenesis trials, it was shown that Isoleucine, phenylalanine, methionine and valine may also be present at Φ positions. Based on the crystal structure of CRM1 bound to SNP1 and RanGTP, it was seen that the N terminal SNP1 region, which binds CRM1 in the hydrophobic cleft had a residue in addition to the defined 4 Φ residues of the NES consensus which contributed to binding (Monecke et al., 2009; Güttler et al., 2010). Following this structural analysis, the NES was examined more closely to identify what was the role of this additional Φ residue. This additional Φ residue was defined as the $\Phi 0$ position in addition to the 4 Φ hydrophobic residues. Based on this analysis the NES consensus was redefined as $x \Phi 0 x(2)\Phi 1x(3) \Phi 2 x(2-3)\Phi 3x \Phi 4$ for a PKI type NES and is shown in the fig 3 (Monecke et al., 2009, Güttler et al., 2010).

Further structural analysis was done to understand how the Rev- NES and PKI-NES fit into the CRM1 hydrophobic cleft in the export complex. It was clear that the hydrophobic residues of the NES need to be placed at accurate points in the binding cleft of CRM1 and this is achieved by adopting different conformations with the help of the spacer residues e.g. as seen in SNP1 where the NES adopts a three-dimensional fold to correctly align the Φ residues in the binding cleft. This is clearly depicted in the fig 3 (Güttler et al., 2010).

There is a range of different residues that are permitted at the Φ positions and these can also have different spacings. There are some other factors that explain why many CRM1 cargoes can bind the hydrophobic cleft even though the residues do not fit the perfect consensus. The hydrophobic residues need to be placed at specific points in the NES binding cleft. Variation in the spacing of the Φ positions in the different functional NESs is compensated by different conformations of the NES backbone in the binding cleft. There are other reasons why CRM1 can allow export of so many cargoes.

By studying the crystal structure of CRM1-SNP1-RanGTP it was shown that the cargo bind on the outer surface of CRM1, thus can have a greater freedom in shape and size, since CRM1 does not wrap around the cargo. Moreover, there is no direct

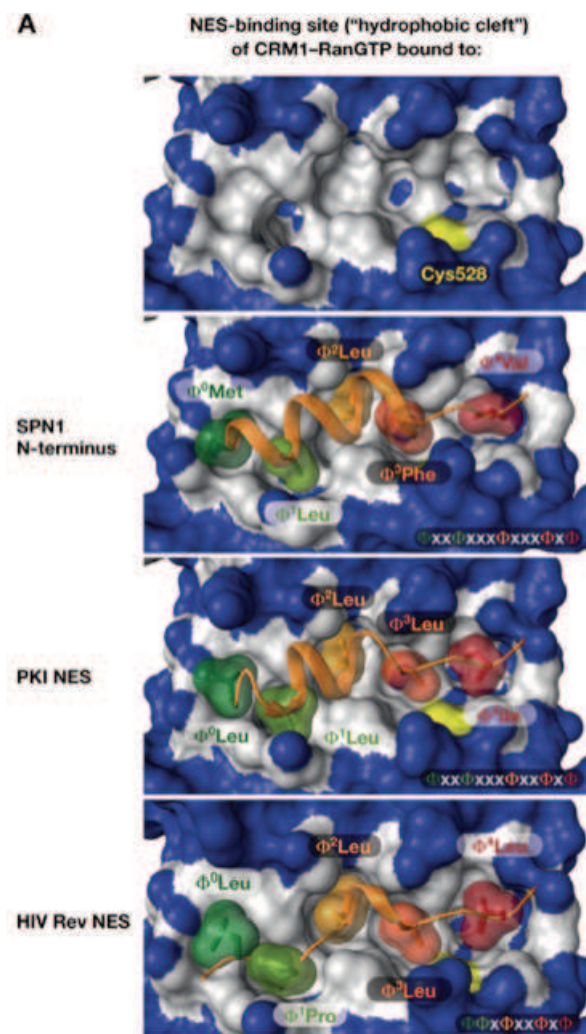
contact with RanGTP, so the cargo has less restrictions imposed by this interaction as well (reviewed in Güttler and Görlich, 2011).

So far, the major emphasis has been put on defining the hydrophobic residues of the Leucine-rich NES consensus. The hydrophobic residues have been regarded as key points in influencing of affinity of the NES of the cargo to CRM1. This is a rather incomplete approach since the spacer residues also have a very important role in influencing the affinity of the NES to CRM1 and this issue needs to be addressed.

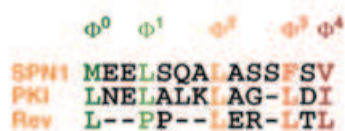
1.14 Identification of classical NES in different cargoes for export by CRM1 using bioinformatic tools

The NES consensus recognized by CRM1 has been used to create various bioinformatics tools where the primary sequence of the protein of interest can be submitted. The software then scans the protein for a NES-like sequence based on a defined consensus. However, these results are often inaccurate and give false positives or false negatives hits. Some examples of NES search softwares are NetNES and NESsential (la Cour et al., 2004; Fu et al., 2011).

When these algorithms are designed, the parameters usually used belong to the 'perfect' NES definition. In this scenario, all the Φ residues should occupy a specific position and constitute specific residues for maximum strength in binding. Sequence based predictions can often be misleading since the structure or disorder of the sequence and its accessibility to the solvent should also be taken into account. If a crystal structure is available, then this information should also be considered. A sequence hidden in a structured fold has a less likely chance to act as a functional NES (Fu et al., 2011; la Cour et al., 2004). Usually these are found in exposed regions or at the ends of the structured domains. Natural NESs bind CRM1 more weakly than a perfect consensus NES would do. Biologically this makes sense, since an export signal needs to bind for effective export and then allow disassembly of the complex in the cytosol (Kutay et al., 2005). This attenuation in the binding strength of an NES could occur because of incorporation of less favourable residues in the Φ positions, less optimal spacings of the Φ residues or skipping of Φ positions altogether.



B Structure-based sequence alignment of NESs



C NES consensus

PKI-class NESs	Rev-class NESs
<p>x Acidic residues preferred</p> <p>Φ⁰ I = V = M > L > A = Y > F = W > P</p> <p>x₂ Acidic residues preferred</p> <p>Φ¹ L > I > V = M > F > A > W (P) (Y)</p> <p>x₃ F = M > L > I = V > Y > W (P)</p> <p>x₂ > x₃ L = M > I > V > F > W = A (P) (Y)</p> <p>x L > I > M > V > F (P) (A)</p>	<p>Φ⁰ -</p> <p>Φ¹ = P >> L</p> <p>x</p> <p>Φ²</p> <p>x₂</p> <p>Φ³</p> <p>x</p> <p>Φ⁴</p>
Number of Φ residues required for CRM1 binding: 4 / 5	5 / 5

Figure 3: Structural analysis of the CRM1 binding cleft (Güttler and Görlich, 2011):

A) The hydrophobic cleft with different NESs bound is shown.

B) give sequence details and show how the NES are aligned in the hydrophobic cleft of CRM1.

C) shows the NES consensus and describes which residues are preferred at particular hydrophobic positions by mutagenesis in the PKI NES.

In the bigger picture, the Φ residues may have multiple functions and this could also influence the NES sequence. All these possibilities introduce a lot of diversity in the NESs for CRM1 cargo (reviewed in Güttler and Görlich, 2011).

Recently a new NES prediction algorithm has been introduced called NESdb database. Here the consensus followed is $\Phi 1x(3) \Phi 2x(2) \Phi 3x \Phi 4$ (Xu et al., 2012). Thus, different sources follow different NES consensus definitions and introduce a bias of their own. The structure of the PKI-NES and Rev NES suggests the consensus $x \Phi 0 x(2)\Phi 1x(3) \Phi 2 x(2-3)\Phi 3x \Phi 4$. In our opinion, this is a validated and more reliable definition of the NES consensus which has been referred to in this work.

An important issue regarding the NES consensus is the spacer residues. These have been not been defined yet. Spacer residues of the NES play a crucial role in modulating the secondary structure of the NES for correct fitting and also interact with the CRM1 backbone. Hence, these residues should also be analysed deeper and defined.

These parameters should be included in designing algorithms for NES prediction to refine the search and get reliable hits.

1.15 Contribution of the spacer residues in the NES consensus:

The spacer residues help the NES to attain the conformation, which enables the correct fitting of the hydrophobic residues in the binding cleft of CRM1. In some cases, nature of these residues in the spacer region also contribute to the binding e.g. in the S-PKI NES, when the spacer residues between $\Phi 0$ - $\Phi 1$ were made more acidic, the affinity for CRM1 increased dramatically. These spacers interact with the Lys514 and Lys 560 in the binding cleft of CRM1 (Güttler et al., 2010). When the neutral spacer residues between $\Phi 0$ - $\Phi 1$ were made acidic in An3 NES, this caused a significant increase the binding to CRM1 as well. They interact with Lys522 and Lys560 of CRM1 (Monecke et al., 2009; Güttler et al., 2010). These acidic residues may also position the NES Φ in the binding cleft by helping to form a secondary structure. Residues like Methionine, Alanine, Leucine, uncharged Glutamate, and Lysine have especially high helix-forming propensities and could enable better fitting

in the binding cleft. There is a preference of specific amino acids at the spacer positions and this needs to be investigated.

Hence, to have a more complete and accurate definition of the NES, information regarding the spacer residues needs to be taken into consideration. If these parameters are taken altogether, a more reliable NES search portal can be developed for the identification of CRM1 cargo and their export signals.

There are two CRM1 cargoes of interest where the Nuclear Export Signals have not been identified yet - eIF2 β and RanGAP1. During my doctoral studies, I focused on locating the NESs in these cargo proteins.

1.16 Nuclear export of translation factors:

Compartmentalization in eukaryotes avoids translation of unspliced transcripts, which can lead to formation of misfolded or lethal mutants (Bohnsack et al., 2002). In spite of the presence of a physical barrier, translation factors can enter the nucleus accidentally during reformation of the nuclear envelope during cell division. If the size of the translation factors is below the exclusion limit of the permeability barrier of the NPC, they can passively diffuse into the nucleus as well (Bohnsack et al., 2002). Moreover, key components for the translation machinery are generated in the nucleus e.g. tRNA, mRNA, ribosomal RNA, ribosomal subunits and these need to be exported into the cytosol (reviewed in Görlich and Kutay, 1999). Hence, there is constant traffic of cargoes across the nuclear pore complexes to ensure that the translation machinery is fully functional in the cytosol.

With issues pertaining to the accidental entry of translation factors (40kDa or below in size), many transport receptors have been identified to constantly exclude them and ensuring that they are present in their destined location - the cytosol. For instance, eIF2B ϵ , eRF1, eEF1B β , eEF2, eIF2 complex and all its subunits, the eIF3 complex, eIF5 etc. are some of the translation factors, which have been shown to be excluded by CRM1 (Bohnsack et al., 2002). Exp5 was shown to export eEF1A along with aa-tRNA. (Bohnsack et al., 2002). The final stages of ribosome maturation that occurs in

the cytosol (Zurdo et al., 2000) and nuclear exclusion of translation factors are multiple mechanisms used by the cell to restrict translation in the cytosol.

1.17 Nuclear exclusion of eIF2 β by CRM1:

eIF2 β is a translation factor which belongs to the eIF2 complex and this has been shown to be excluded by CRM1 (Bohnsack et al., 2002). eIF2 complex is crucial in transferring Met-tRNA^{met} to the 40S subunit of the ribosome and consequently forming the pre initiation complex. Upon binding of the initiator tRNA to the AUG codon, eIF2 complex undergoes GTP hydrolysis (Pestova et al., 2001; Price and Proud, 1994).

The N terminal part of the protein is unstructured and highly charged. It has three conserved K boxes, which consist of poly Lysine tracks. The C terminal part is structured and this consists of a conserved Zn finger domain and is found in archaea and eukaryotes. (Laurino et al., 1999). The NMR structure of the archaeal aIF2 β has been solved (Cho et al., 2002).

Small sized translation factors like eIF2 β can accidentally enter the nucleus via passive diffusion across the permeability barrier. Another way of misallocation is during dissolution and reformation of the nuclear envelope during cell division. However, the nuclear localisation of eIF2 β is rather fast and strong, as seen when CRM1 function is blocked by Leptomycin B. This rapid rate of accumulation in the nucleus could not be sustained by passive diffusion. This suggests that eIF2 β may be actively imported. The basic motifs present in eIF2 β could be mistaken for sequences involved in nuclear localisation. The sequence of eIF2 β is so charged, that maybe an importin binds the charged region accidentally to mask the charges and prevent aggregation - thus acting like a chaperone. Nuclear export eIF2 β may also be a way to correct this 'accidental import' as well.

The mode of recognition of eIF2 β by CRM1 is unknown. When the sequence was submitted to various bioinformatics servers to predict an NES for CRM1 binding, no hits were found. Hence, the goal of the project was to elucidate how CRM1 recognises eIF2 β for nuclear export.

1.18 Nuclear export of RanGAP1 by CRM1:

RanGAP1 is the functional homolog of *Saccharomyces cerevisiae* protein Rna1p. It shares little sequence conservation (42%) with its yeast homolog but it has functional similarities that have been characterized (Melchior et al., 1993; Bischoff FR et al., 1995). Rna1p was shown to be exported by CRM1/Xpo1p and an NES and NLS were reported to be identified (Feng W 1999). Truncated constructs of Rna1p were analysed for interaction with CRM1 for export, and the residues 316 to 357 were suggested to have 2 NESs (320 to 328 and 329 to 357) (Feng W et al., 1999). When the crystal structure of Rna1p was solved (Hillig RC et al., 1999), the suggested NES were located in a structured region and the critical hydrophobic residues were not exposed. In our opinion, it is improbable that this structured sequence could serve as an export signal and fit inside the narrow CRM1 binding cleft. Maybe creating truncations disrupted secondary structures of the protein or exposed hidden regions of Rna1p, which acted as false positives for nuclear export in the biochemical assays.

As mentioned before, the RanGTP gradient across the nuclear envelope is crucial for directional transport across the Nuclear Pore Complexes (Gorlich & Kutay, 1999). In the cytosol, RanGAP1 and Ran BP1 help in promoting the hydrolysis of RanGTP to RanGDP. Thus, they play a role in maintaining the RanGTP gradient and disassemble export complexes while exiting the nucleus. RanGAP1 has been shown to be involved in cell cycle regulation in mitosis and assembly of the spindle complex as well (Arnaoutov & Dasso, 2003).

The N terminal region of RanGAP1 consists of LRR or Leucine Rich Repeats and the C terminal has the RanGAP1 domain. RanGAP1 is sumoylated at the C terminus in eukaryotes, which enables it to be associated with the cytosolic side of the NPC by interacting with RanBP2/Nup358 (Bernier-Villamor et al., 2002). The N terminal domain with the LRR is more conserved when compared to the C terminal domain - the LRR region is found in vertebrates and fungal proteins (Hillig et al., 1999). The crystal structure of the yeast homolog, Rna1p has been solved (Fig 4).

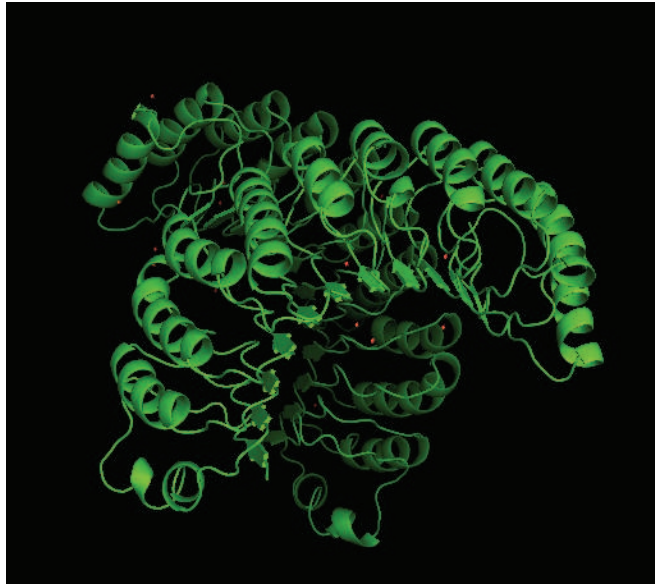


Figure 4: Crystal structure of Rna1p (Hillig et al., 1999; PDB: 1YRG).

When the sequences of RanGAP1 homologs were compared the lack of conservation of residues was evident. Moreover an additional C terminal domain was also added and it has a sumoylation site. This allows RanGAP1 to be tethered to the NPC in vertebrates. (Bernier-Villamor et al., 2002). A figure showing the alignment of the domains has been shown in fig 5.

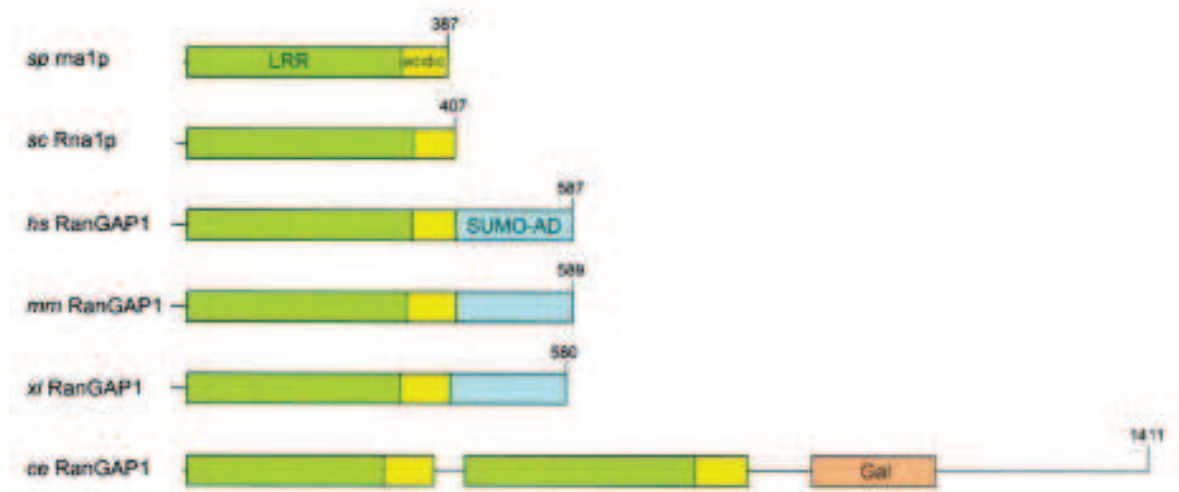


Figure 5: RanGAP1 homologs with the domains aligned (Hillig et al., 1999).

It would be very interesting to see how RanGAP1 is recognized by CRM1. Our immediate goal is to locate the sequence recognized by CRM1 for export.

Another issue we want to address is how RanGAP1 can get exported without hydrolysing its own export complex? Solving the crystal structure of CRM1-RanGAP1-RanGTP complex will help us elucidate the mechanism for interaction of these three components for successful nuclear export.

Previous attempts to locate an NES using bioinformatic tools proved to be imprecise. This is a Leucine rich protein which gives several false positive hits when searched for the classical Leucine rich NES. Moreover the conservation of the residues is rather less, thus, extrapolating information from one homolog to another to locate the NES may also give misleading results. Thus, the best approach is to use an unbiased and independent approach to locate the export sequence of RanGAP1 for recognition by CRM1.

1.19 Structural analysis of nuclear transport receptors:

An X-ray Crystallography approach of protein complexes depends on certain crucial factors: Size, domain structure and rigidity of the investigated proteins are just a few important parameters to be considered. A highly pure, homogenous, small sized and rigid protein represents an ideal candidate. NTRs are usually above 100 kDa in size and they have an intrinsic flexibility by virtue of the HEAT repeats. Their flexibility makes them a challenging target to crystallize. Usually large proteins are truncated to form shorter proteins with domains being purified separately. This approach cannot be undertaken with NTRs as they form a continuous structure with HEAT repeats. A promising strategy to make these proteins rigid enough for a successful x-ray crystallography trial is to investigate them when they are bound to their respective cargo.

The ultimate goal of solving a structure of a NTR bound to its cargo is to deduce the mechanism of how the proteins interact, form a complex and possibly identify a recognition motif. This recognition motif can be used to identify other cargo of the NTR being studied. Moreover, for export complexes, the crystal structure will reveal how RanGTP influences the structure and sheds light to the consequent behaviour of the NTR.

Amongst the identified nuclear transport receptors, there are certain NTRs, which have the remarkable feature of transporting cargo in both directions (Mingot et al., 2001, Gontan et al., 2009; Lipowski et al., 2000, Kurisaki et al., 2006). These represent therefore interesting candidates to characterize structurally.

For this project, the focus was set to analyse Exportin 4 bound to its export and import cargo.

1.20 Exportin 4 and its cargo

Exportin 4 belongs to the Importin β family of nuclear transport receptors. It is 120 kDa in size and has approximately 19-20 HEAT repeats. It was initially discovered as an NTR responsible for the nuclear export of eIF5A1 (Lipowski et al., 2000). Later on, Exportin 4 was shown to export Smad 3 (Kurisaki et al., 2006). In 2009, it was proven that Exportin 4 has a dual function by also acting as an Importin that carries Sox2 and SRY into the nucleus (Gontan et al., 2009).

1.21 Eukaryotic initiation factor 5A (eIF5A):

Eukaryotic initiation factor 5A or eIF5A was initially identified to play a role in the initiation of translation. The exact function of eIF5A is under speculation (Zanelli & Valentini, 2007). eIF5A has two distinct isoforms - eIF5A1 and eIF5A2, which have a similarity of 84% at the sequence level, but the functions of the two isoforms can vary (Caraglia et al., 2011).

eIF5A is important for the initiation in translation of specific mRNAs and has also been shown to play a role in cell apoptosis, cell cycle progression and cell proliferation. Apart from that is considered to be of interest in oncologic treatment (Paz et al., 2011).

It is also shown that eIF5A is involved in modulation of cytokine signalling and it is being considered as a therapeutic target for controlling diabetic inflammation (Maier et al., 2010). Other research work shows involvement of eIF5A in various diseases

like malaria, several cancer types, HIV-1 infection etc. (Kaiser, 2012). Due to its involvement in this wide range of diseases, eIF5A is considered a linker for drug targeting and therefore a possible avenue for therapy (Kerscher et al., 2010).

A remarkable feature of eIF5A is that it is the only known protein with the unusual amino acid Hypusine (N (ε)-(4-amino-2-hydroxybutyl)-lysine) (Kaiser, 2012). Hypusine is formed by a post-translational modification called hypusination, which is executed by the two enzymes deoxyhypusine synthase (DHPS) and deoxyhypusine hydroxylase (DOHH) (Park et al., 2010). Interestingly, the maximum conservation is seen around the so called hypusination sites (Park et al., 2010; Kaiser, 2012). This is a two step catalytic procedure, where an aminobutyl moiety is attached to a Lysine (Lysine 50 in homo sapiens) and modified to achieve the mature and active eIF5A. The lysine of eIF5A1 is acted upon by deoxyhypusine synthase and converted into an intermediate state. This intermediate is consecutively converted to the mature and active form by deoxyhypusine hydroxylase. The scheme explaining this process is shown below (Fig 6). Hypusination is essential for viability and this is found only in eukaryotes and certain species of archea. DHPS and DOHH are highly conserved enzymes and specific in their activity, hence they are also targets for therapeutic intervention (Park et al., 2010).

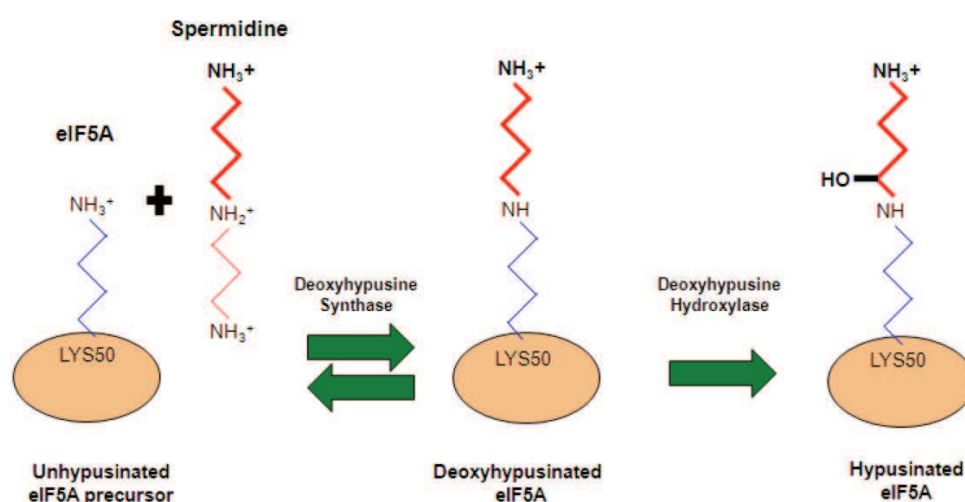


Figure 6: Modification of eIF5A by the hypusinating enzymes (Maier et al., 2010).

The crystal structure of eIF5A from Homo sapiens, hyperthermophilic archaeobacterium Pyrococcus horikoshii OT3 (Yao et al., 2003), Methanococcus jannaschii (Kim et al., 1998) and Pyrobaculum aerophilum (Peat et al., 1998) has been elucidated. Fig 7 shows the crystal structure of Hs eIF5A (PDB:3cpf) (Tong et al., 2009).

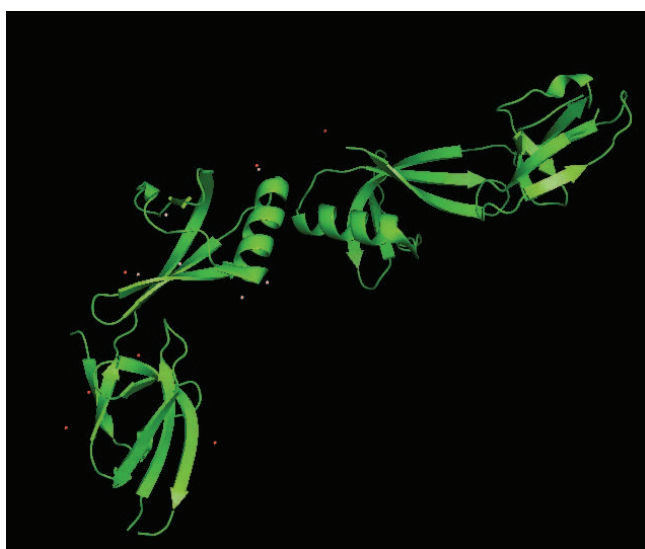


Figure 7: Crystal structure of human eIF5A1; PDB: 3cpf (Tong et al., 2009).

The interaction strength of Exportin 4 and eIF5A depends on the presence of this hypusine. The strength of interaction of Exportin 4 and RanGTP with unmodified, intermediate and fully modified eIF5A was measured. It was found to be the highest with the fully modified variant. (Lipowsky et al., 2000) This has been shown in Table 2 (shown below).

eIF-5A species	Apparent K_D for dissociation from eIF-5A–Exp4–RanGTP complex	Relative affinity for Exp4 (native eIF-5A = 100)
Purified from HeLa cells, fully hypusinated	2 nM	100
Recombinant, non-modified	75 nM	2.5
Recombinant, deoxyhypusinated	25 nM	8
Recombinant domain I (residues 1–83), deoxyhypusinated	~ 1–2 μ M	0.1
Recombinant domain II	>10 μ M	<0.01

Table 2: Exportin 4 interaction with eIF5A in different states of modification; (Lipowsky et al., 2000). Exportin 4 has highest affinity for the fully mature eIF5A.

Thus, to crystallize a stable export complex of Exportin 4 with eIF5A and RanGTP, a large amount of fully hypusinated and pure eIF5A is required.

1.22 Smad 3:

Smad 3 (Mothers against decapentaplegic homolog 3) is involved in cell signalling in the cytosol and acts as a transcription factor in the nucleus. Smad proteins are important intracellular mediators of signalling activated by the transforming growth factor- β (TGF- β) family cytokines. Smad proteins execute processes like embryonic development, arrest of cell growth and are involved in regulation of tissue homeostasis (Heldin & Moustakas 2012). Smad 3 belongs to the R Smads and they contain the mad homology domains MH1 and MH2. MH1 and MH2 structured and are connected by a flexible linker (Chai et al., 2003; Heldin and Moustakas, 2012).

Smad 3 uses the β hairpin in its MH1 domain to interact with DNA and other transcription factors (Shi, 2001; Heldin and Moustakas, 2012; Shi & Massagué, 2003). Smad 3 interacts with Type I receptors like the TGF β type I receptor via its MH2 domain (Qin et al., 2002). On its extreme C terminal region, it gets phosphorylated by activated TGF β (Heldin and Moustakas, 2012). The crystal structures have been depicted in the fig 9 and fig 10.

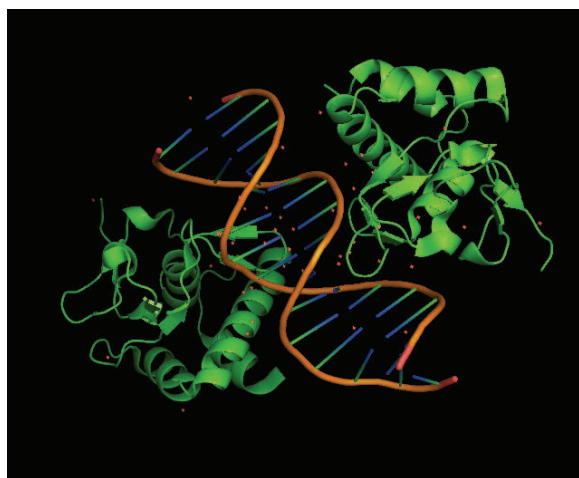


Figure 9: Crystal structure of Smad3 MH1-DNA complex; (Chai et al., 2003) PDB: 1ozj.



Figure 10: MH2 domain of transcriptional factor SMAD3; (Qin et al., 2002). PDB: 1mjs.

The motif Lys(40-)Lys-Leu-Lys-Lys(44) was found in its N terminal MH1 domain and shown to be imported by Importin β . Upon ligand binding TGF β receptor phosphorylates the N terminal domain of Smad 3 and this increases its affinity for Importin β (Xiao et al., 2000; Kurisaki et al., 2001).

The interaction of Smad 3 with Exportin 4 has been mapped previously. This work suggests that Smad 3 has high affinity patches along its surface for its interaction with Exportin 4 (Kurisaki et al., 2006).

A structural analysis of Exportin 4 bound to RanGTP and Smad 3 will give clear insight into the export mechanism. This will also help us identify the recognition sequence in Smad3 for its nuclear export.

1.23 Sox 2:

Sox 2 or SRY (sex determining region Y)-box 2 is a transcription factor involved in the proliferation of undifferentiated embryonic stem cells. Along with POU5F1/OCT4, MYC/c-Myc and KLF4, SOX2 forms the Yamanaka factors and these proteins can act in a coherent manner to generate iPS or induced pluripotent cells by converting differentiated cells into cells in undifferentiated embryonic-like cells (Takahashi et al., 2007). Sox2 has also been shown to be involved in various cell

signalling pathways e.g. the canonical Wnt receptor signalling pathway. (Kelberman et al., 2008).

The crystal structure of Sox2 (HMG domain) and Oct4 (POU domain) has been solved - shown in fig 11.

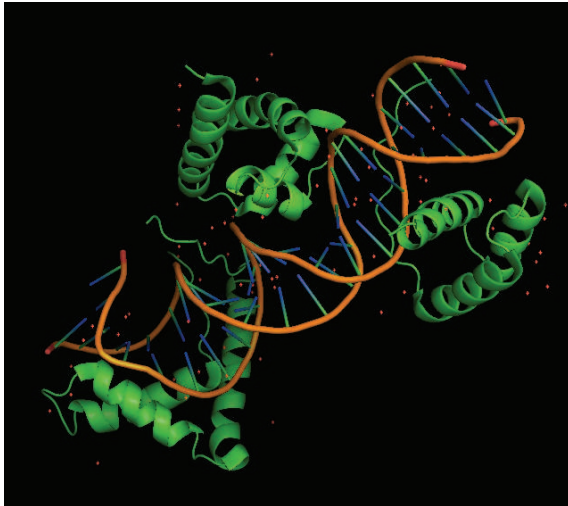


Figure 11: Crystal Structure of a POU/Hmg/DNA Ternary Complex Suggests Differential Assembly of Oct4 and Sox2 on Two Enhancers; (Remenyi et al., 2003) PDB: 1gt0

Sox family of proteins contain an HMG domain which has been shown to interact directly with the target genes to control their expression (Mertin et al., 1999; Remenyi et al., 2003).

In 2009, it was shown that Exportin 4 can import Sox 2 into the nucleus (Gontan et al., 2009).

1.24 SRY:

Sex-determining region Y protein or SRY protein is a transcription factor and it also belongs to the Sox family of transcription factors. It uses its HMG domain to interact with the genes whose expression it regulates in order to support male sex determination and embryonic development (Koopman P, 1999; Kashimada and Koopman, 2010).

The function of SRY crucially depends on its nuclear import. Two NLSs have been shown to flank the HMG domain - the C terminal region has a β -NLS, which interacts with Importin β for nuclear import and the N terminal region has a CaM-NLS, which recognises Calmodulin (calcium-binding protein CaM) (Smith et al., 2004; Kaur et al., 2010). Both these NLSs have been shown to be important for nuclear import of SRY (Kaur et al., 2010; Kaur and Jans, 2011).

The structure of the HMG domain of SRY solved by NMR is shown in fig 12, where it is clearly shown that the SRY HMG domain binds in the minor groove. It causes a dramatic bending of the target gene sequence and as a consequence SRY can regulate the expression. (Murphy et al., 2001).

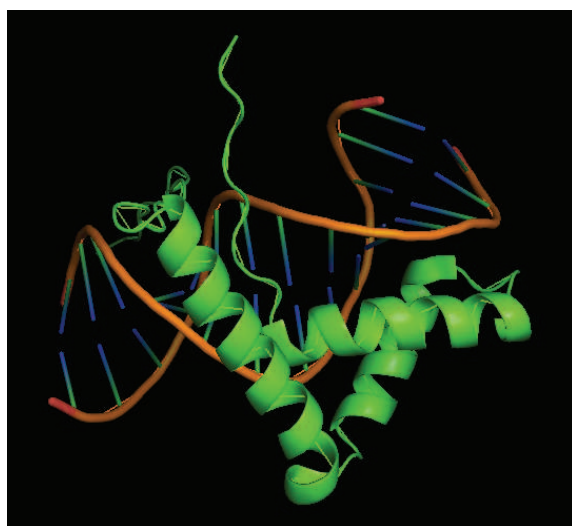


Figure 12: Structure of the HMG domain of SRY (PDB:1j46); (Murphy et al., 2001).

Both Sox2 and SRY were shown to be imported by Exportin 4 (Gontan et al., 2009). The mode of recognition of these transcription factors by Exportin 4 is unknown. Does it recognise a sequence or a three dimensional conformation? Is there a consensus followed by Exportin 4 to recognise its import and export cargo in the presence of RanGTP? The crystal structure of Exportin 4 in its export and import conformation will give valuable mechanistic insights. This information can be used further to identify more cargo of Exportin 4 by using bioinformatic tools and searching similar modules in different proteins and then validating the information in vivo by cell biological assays and in vitro by direct interaction analysis.

2 Materials and Methods

2.1 CHEMICALS USED:

The reagents used in the laboratory were purchased from, the following companies:

- Calbiochem (San Diego, CA, USA)
- GibcoBRL-Life Technologies (Paisley, UK)
- Merck (Darmstadt, Germany)
- MoBiTech (Göttingen, Germany)
- Promega (Madison, WI, USA)
- Roche (Mannheim, Germany)
- Carl Roth GmbH (Karlsruhe, Germany)
- Sigma-Aldrich (St. Louis, MO, USA)

Some chemicals had been specifically ordered for some experiments and will be mentioned separately.

2.2 INSTRUMENTS USED:

SensoQuest lab cycler	SensoQuest, Göttingen, Germany
Perfection V700 Photo Scanner	Epson Perfection V700 Photo
TCS SP5 microscope	Leica Microsystems, Wetzlar, Germany
Cooled tabletop centrifuge	Eppendorf 5417R
Tabletop centrifuge	Eppendorf 5424
WX Ultra centrifuge	Sorvall
Discovery M120	Sorvall
RC6 plus centrifuge	Sorvall
NanoDrop ND-2000	PeqLab, Germany
Sonifier 450	Branson, UK

Pipetman pipettes	Gilson, Middleton, WI, USA
Intas Gel Documentation System	Intas Science Imaging Instruments GmbH, Germany
Incubator/Climo - Shaker ISF1-X	Kuhner Shaker

2.3 SOFTWARE:

Computer Operating system:

- MacOSX 10.6.8

Computer programs employed for lab work

For image analysis:

- Adobe Photoshop CS4 for Macintosh
- Adobe Illustrator CS4 for Macintosh

For designing of primers for cloning and mutagenesis:

- DNA Star software Lasergene 8
- Oligo 6.8 software for designing Oligos
- Gene Designer

Literature compilation and assembling bibliography:

Papers version 1.9.4

Documentation, presentations:

- Microsoft Office 2008, 2011 for Macintosh
- Adobe Acrobat Professional for Macintosh

Literature review and internet searches:

- Research articles and reviews: Pubmed, Google
- Gene sequences: Uniprot, NCBI

- Homology search: BLAST
- Protein domain prediction: SMART
- Various bioinformatic servers from ExPASy were also used often.

2.4 BACTERIAL GROWTH MEDIA - E.COLI CULTURE GROWTH MEDIA.

Media	Composition
2YT medium	16g Tryptone 10g Yeast extract 5g NaCl 20mL Glycerol 85% ddH ₂ O to 1 L
2YT medium for plates	16g Tryptone 10g Yeast extract 5g NaCl 20mL Glycerol 85% 14g Agar ddH ₂ O to 1 L
LB medium	10g Tryptone 5g Yeast extract 10g NaCl ddH ₂ O to 1 L

2.5 BUFFERS

Buffer	Composition
SDS Page:	10x Elpho (SDS-PAGE) Running

	buffer: 60 g Tris Base 300 g Glycine 20 g SDS (good quality)
SDS Sample Buffer:	125 mM Tris pH 6.8 50 mM DTT 3% SDS 1.0 M Sucrose
Bromophenol-blue	Coomassie Stock Solution: 2 % (weight/volume or w/v) Coomassie Brilliant Blue G250
Agarose gel electrophoresis:	50xTAE Running buffer: 242 g Tris Base 57.1 ml Acetic acid 100 ml 0,5M EDTA pH 8,0
Orange G Sample Buffer:	10 mM TrisHCl pH 8.0 10 mM EDTA pH 8.0 50 % (w/v) Glycerin 25 % (w/v) Orange G
DNA-ladder:	50 ng/μl 1kb-Ladder or 50 ng/μl 100kb-Ladder in Orange Sample Buffer (Gibco)
E-Mix (20x)	200 mM Creatine Phosphate (Sigma; MW 255.1 g/mol) + 10 mM ATP (Sigma; Mw 551.1g/mol) + 10 mM GTP

	(Sigma; MW 523.2 g/mol)) + as powder were weighed separately and dissolved in water. After this they were mixed. 20-50 mM Hepes pH7.5 + 250 mM Sucrose Creatine kinase 1mg/ml (powder; Roche) were added and dissolved. This solution was aliquoted and stored at -80°C.
50x TAE (agarose gel electrophoresis)	242 g Tris Base 57.1 ml Acetic Acid 100 ml 0.5M EDTA pH 8.0 add 1L ddH ₂ O

All Tris buffers were prepared using the Henderson-Hasselbach equation.

2.6 PROTOCOL FOR PRODUCTION OF ELECTRO-COMPETENT E.COLI CELLS - BY G. KOPP, KORAY KIRLI.

Preculture: The E.Coli strain is grown in liquid SOB medium with appropriate antibiotics at 37°C in a shaker.

The next day in the morning the pre-culture is diluted with SOB medium (warmed to 37°C) and grown further at 37°C till the OD reaches 0.8.

Harvesting: The bacterial culture is sedimented at 3500 rpm, 10 min at 4°C. Following this the supernatant is discarded.

Resuspension and wash : The harvested bacteria is resuspended in 100 ml 1.4% glycerol in water (w/v). These are again harvested.

The previous step is repeated – resuspension in the glycerol solution and harvesting.

Resuspension and aliquoting: The harvested bacteria is resuspended in twice their amount in 14% glycerol in water.

Following this, the resuspended E.Coli are aliquoted and snap frozen in liquid nitrogen for storage at – 80°C.

2.7 LB PLATES TO GROW COLONIES OF TRANSFORMED BACTERIA.

The transformed bacteria adopt the resistance of the plasmid that is transformed into them – hence for selection they are plated on LB plates with required antibiotics with the given concentrations:

Ampicillin (100µg/ml)

Kanamycin (25µg/ml)

Chloramphenicol (100µg/ml).

The bacterial cultures can also be grown in liquid cultures using the mentioned antibiotic concentrations.

2.8 BACTERIAL STRAINS USED

Strain	Genotype	Company
TOP10 F'	F' {lacIq Tn10 (TetR)} mcrA Δ(mrr-hsdRMS- mcrBC) Φ80lacZΔM15 lacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG	Invitrogen (Karlsruhe, Germany)
NEB5α F'	F' {proA+B+ lacIq Δ(lacZ)M15 zzf::Tn10 (TetR)} fhuA2Δ(argF- lacZ)U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	New England Biolabs (Beverly, MA, USA)
BLR	F-ompT hsdSB(rB-mB-) gal dcm Δ(srl-	Invitrogen (Karlsruhe)

	recA)306::Tn10 (TetR)	
TOP10F'	F'[lacQ, Tn10 TetR] mcrA Δ(mrr-hsdRMS-mcrBC, 80lacZΔM15 ΔlacX74, deoR, recA1, araD139 (ara-leu) 7679, galU, galK, rpsL (StrR), endA1, nupG	Invitrogen (Karlsruhe)
BL21 DE3 Star	F-ompT gal dcm lon hsdSB(rB-mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Invitrogen (Karlsruhe)
NEB Express	fhuA2 [lon] ompT gal sulA11 R(mcr- 73::miniTn10— TetS)2 [dcm] R(zgb-210::Tn10— TetS) endA1 Δ(mcrC- mrr)114::IS10	Invitrogen (Karlsruhe)

Methods:

All standard methods in Molecular Biology were performed as indicated in Sambrook and Russell 2001 and some optimisations.

Oligonucleotides for cloning were ordered from Sigma-Aldrich (St. Louis, MO, USA)

2.9 PCR

This was used to amplify a defined DNA fragment or the entire vector with specific mutations introduced by the forward primer. These procedures are done using specific primer pairs. The enzyme used is PfuS Tripple Mix (100ng/μL PfuS

DNA polymerase, 15ng/μL Pab pyrophosphatase, 2,5ng/μL Pab dUTPase; generated by Steffen Frey) which contains the thermostable polymerase, a thermostable pyrophosphatase and dUTPase.

The PCR mix is as follows: A 50 μl PCR reaction contains 50 ng template DNA, 5 μl DMSO (2% v/v), 1μl PfuS Tripple Mix (2U/μl), 5μl of 10x dNTPs (2.5 mM each) and 5 μl of 10x PfuS buffer (200 mM Tris/HCl pH9, 250 mM KCl, 15 mM MgSO₄, 10 mM (NH₄)₂SO₄, 1% Tween-20, 1 mg/ml BSA), 0.5 μl each of the forward and reverse oligoes (100 μM each), and ddH₂O to add up to 50μl volume.

PCR steps on the SensoQuest Lab Cycler (London, UK) PCR machine:

initial denaturation: 94°C for 5 minutes,

denaturation 94°C for 10 seconds,

annealing temperature varies for each construct - for 10 seconds,

extension 72°C, 1 minute per 3 kb of PCR product length,

final extension 72°C for 10 minutes.

Steps 2 to 4 (denaturing to extension) were repeated 25 times.

The annealing temperature was calculated using the Oligo 6.8 software.

References: Mullis *et al*; 1986 (PCR).

2.10 MUTAGENESIS:

This was done using PCR mutagenesis where the plasmid was amplified with apposing primer and the forward primer carried the mutation of interest - deletion, point mutation etc. The primers have to be HPLC purified and phosphorylated. The PCR was performed as previously described and following this the template was put for digestion with Dpn1 at 37°C over night or for 2-3 hours (for 100μl PCR mix, 5.5 μl of NEB buffer 4 and 2 μl of Dpn1 enzyme was added). Following this the DNA was purified on agarose gel, ligated over night at

16°C, transformed into an appropriate bacterial strain and plated on agar plates with the correct antibiotic. These steps have been described below.

2.11 COLONY PCR:

The PCR was carried out as previously described but with the Tth polymerase (100ng/μL Tth DNA polymerase, 15ng/μL Pab pyrophosphatase; generated by Steffen Frey) was used for colony-PCR. The colony was picked from the agar plate and resuspended in 10μl distilled water. 1μl of this was used for the colony PCR reaction. The results were analysed by running PCR product on agarose gel.

2.12 PRIMER DESIGN:

Oligo 6.8 software was used to design primers and SeqBuilder was used to make vector maps for virtual cloning. These belong to the DNASTar software package.

The primers were ordered from Sigma-Aldrich, Germany.

2.13 DNA DIGESTION:

All restriction endonucleases were purchased from NEB (New England Biolabs, Beverly, MA, USA). The digestion was performed according to the manufacturer's instruction.

Thereby, the enzyme was used for an 1 to 1.5 hr digestion at 37°C.

The plasmid DNA with sticky ends was additionally de-phosphorylated using FastAP by Fermentas (2μl per 100μl digestion mix at 37°C for 30 minutes).

2.14 DNA AGAROSE GEL ELECTROPHORESIS:

Following PCR or restriction digestion the ds DNA was run on an agarose gel for separation of the fragments generated. The agarose gel concentration varies from 1-2% and depends on the length of the DNA fragments to be separated.

The agarose gel of a selected percentage was made with 3 μ l of a 20 mg/ml ethidium bromide solution per 100 ml agarose gel. 1xTAE buffer was used for gel preparation and was also used to run the gel. The DNA samples were loaded with 1/8 (v/v) of Orange G loading buffer and visualised in UV light.

2.15 PURIFICATION OF DNA FRAGMENTS EXCISED FROM AGAROSE GEL:

Zymoclean DNA Recovery Kit (Zymo research CA; USA) was used to purify DNA fragments excised from agarose gels and used according to the manufacturer's instructions.

The concentrations of the purified DNA was measured using a NanoDrop 2000C spectrophotometer (Thermo Fisher, Ohio, USA).

2.16 DNA LIGATION:

To ligate the digested and purified DNA fragments first the vector was dephosphorylated right after digestion using FastAP (Fermentas). Then a ligation was performed using T4 ligase (100ng/ μ L; generated by S. Frey)) -1 μ l, ligase buffer - 1 μ l, DNA - insert and vector were added according to the calculation - insert 40ng/kb and vector 20ng/kb. The ligase buffer was also made in house - 10 x buffer: 500 mM Tris pH=7.5, 100 mM MgCl₂, 100 mM DTT, 10 mM ATP, 250 μ g/ml BSA. The volume was built to 10 μ l with water. This was kept o/n at 16°C for a blunt end ligation and at 37°C for 1 hour for stick end ligation before being transformed into E.Coli.

2.17 ELECTROPORATION:

Electro-competent E.coli cells (50 µl) were mixed with 50 to 80 ng plasmid DNA in an electroporation cuvette (Biorad, Hercules, CA, USA) on ice. The DNA was electroporated into the bacteria using a GenePulser (BioRad, Burlington, USA) according to the manufacturer's instruction.

Then the bacteria were allowed to recover for 1 hour at 37°C with mild shaking in 2YT media without the antibiotics.

Following this the bacteria was plated on agarose plates with the correct antibiotic and kept at 37°C o/n. The next day a colony was picked to grow a pre culture in a liquid medium or for other purposes e.g. colony PCR.

2.18 MINI OR MIDI PREPS FOR DNA STOCKS:

Plasmid DNA was either prepared in an analytical scale (mini-preps) by using the NucleoSpin Plasmid Kit (Marcherey Nagel, Germany) or in a preparative scale (midi-preps) using the NucleoBond PC100 Kit (Marcherey Nagel, Germany), in each case according to the manufacturer's instructions.

2.19 DNA SEQUENCING:

The constructs were sent to Seq-Lab Göttingen, Germany for sequencing. The sequencing results were analysed using SeqMan software (DNASTar software package).

2.20 PROTEIN EXPRESSION:

The proteins were cloned into plasmids for expression in E.Coli. The constructs were transformed into electro competent cells and plated on agarose gels with the appropriate antibiotic. The next day a colony was picked to grow a pre

culture at 37°C over night with the appropriate antibiotic. The next day this pre culture was diluted and the culture was induced to grow at conditions which were optimized for temperature, duration of induction and IPTG (Isopropyl- β -D-thiogalactopyranosid) concentrations in an incubator with a shaker. At the time of harvesting the bacteria – the OD was measured and 1mM PMSF was added to prevent proteases. This was harvested at 5000 rpm for 10 minutes. The pellet was resuspended in resuspension buffer (50 mM Tris, 500 mM NaCl, 2 mM Mg(OAc)₂, pH 7.5). This resuspension was stored at – 80°C until further use.

2.21 PROTEIN PURIFICATION:

For large scale purifications Ni-EDTA-Amid chromatography was used.

The protein expressed has a N or C terminal Histidine tag. These Histidine residues can bind divalent heavy metals such as Ni²⁺. Hence when the bacterial lysate containing the Histidine tag is exposed to Ni²⁺ beads, the recombinantly expressed protein binds to these with high affinity and the other non tagged bacterial proteins are un bound and later washed away.

The principle of nickel-affinity chromatography is based on the ability of histidine

residues to bind divalent heavy metal ions, such as e.g. Ni²⁺, with high affinity.

5%, 500 Å Ni-EDTA-Amid Sepharose were prepared by Prof. Dirk Görlich.

The bacterial resuspension was thawed in warm water and 1 µg/ml lysozyme was added for half an hour on ice. Then using a sonifier (Sonifier 450, Branson, UK) the bacterial cells were sonicated for lysis. After lysis the lysed cells were centrifuged to separate the soluble and insoluble components of this lysed mixture. They were centrifuged at 37,000 rpm for 2 hours. The supernatant was then transferred to a falcon and Ni²⁺ beads (already equilibrated with resuspension buffer) were added to it and incubated for an hour on a slow rotor at 4°C. After this, the solution with the nickle beads was poured into liquid-chromatography column (Luerlock, Sigma) and the flowthrough was allowed to pass through. A small amount of this was taken for SDS page analysis.

(resuspension buffer - 50mM Tris, 500mM NaCl, 2 mM MgOAc₂, 25mM Imidazole pH 7.5).

NOTE: The ratio of beads to lysate was also optimized on a small scale experiment.

Then the beads were washed with purification buffer (the resuspension buffer + 5 mM DTT) with 3 column volumes. The proteins were eluted in elution buffer (50 mM Tris, 500 mM Imidazole, 100 mM NaCl, 5 mM DTT, 5 mM Mg(OAc)₂) with 500µl volume at a time. These samples were analysed for protein amounts by staining on a nitrocellulose membrane with Amido Black solution and the appropriate samples were pooled. To this sucrose was added to a final concentration of 250 mM.

To calculate the concentration – the absorbance was measured by the nanodrop. Then the molecular extinction coefficient (e₂₈₀) was found by using ProtParam bioinformatic server. This was added to a formula: $A = (e_{280}) * \text{concentration (M)} * \text{length (1 cm)}$.

After this an SDS page analysis (Weber and Osborn, 1969) was done with the before and after induction samples, the flow through and washed, and elution samples - loading 40 mOD was loaded per sample and the proteins were stained by Coomassie blue.

2.22 BINDING ASSAY:

Binding assays were performed to check interaction between proteins. To do this, IgG matrix was used and one of the interacting proteins was tagged with a ZZ tag – bait protein. The untagged protein which was to be pulled out if it interacts is untagged and is referred to as the prey. These interactions were done on a small scale and done using MoBi Cols. For each sample 10µl beads IgG were used. The beads are first washed with

2x 0.2 M Glycine pH 2.2 and then twice with the binding buffer (50 mM Tris pH 7.5, 50 mM NaCl, 2 mM MgCl₂, 0.005% Digitonin). The bait was incubated with the beads for 1 hour at 4°C on a slow rotor. The concentration to saturate the beads is 2µg protein /µl of beads. After this the unbound protein was washed off

and the bead were once washed with binding buffer (500µl). At each step the MoBi cols were centrifuged at 1000 rpm. Then a mix containing the prey was added. In the binding assays this mix contains - 1mM Nuclear Transport Receptor, 3µM RanGTP added to a MH1 lysate. The MH1 lysate was prepared as follows: MH1 bacterial cells were harvested at OD 0.6 and resuspended in 50 mM Tris, 2 mM MgCL₂, 50 mM NaCl, 1 DTT pH 7.5. Then they were lysed and the supernatant post centrifugation was used. This lysate is used as a non specificity control. This was then incubated at 4°C for 3 hours. After this, the MoBi Cols were centrifuged, the flowthrough was saved for SDS page analysis, washed twice with binding buffer and the elution was done using 50µl of SDS loading buffer without DTT. The flowthrough, wash and elution samples were analysed using SDS PAGE.

For GST purification the steps were the same except that the elution buffer was 50 mM Tris, 100 mM NaCl, 5 mM MgOAc, 10 mM DTT, 20 mM reduced Glutathione.

2.23 TRANSFECTIONS (courtesy Heinz-Jürgen Dehne):

HeLa 11.1 and HeLa Kyoto were used in this work. The cells were maintained in OPTI-MEM media and splitted at approximately 80% to 90% confluency with the media being replaced every two days.

The transfection was carried out as per the instructions of the **FuGENE 6 Transfection Reagent** manual. The cells were split a day before the transfection. At about 80% confluency the cells were transfected the following day with 1µgDNA and 3µl Fugene 6 (transfecting agent) and kept at 37°C for 24 hours. The ratio of the DNA and transfecting agent was optimised as per the instructions of the Fugene 6 manual. The experiments were carried out in 24 well plates with 1 ml media (OPTI-MEM) per coverslip. The preparation of the complex of Fugene6 and DNA followed by the transfection procedure was

performed in accordance with the guide lines of the **FuGENE 6 Transfection Reagent Instruction Manual Version 2, May 1999.**

The cells were allowed to grow for 24 hours at 37°C. The cells exposed to Leptomycin B were exposed 20nM LMB for 3 hours and then fixed. After the incubation period, the cover slips were fixed as per the protocol noted below and analysed using a confocal microscope. A brief protocol has been noted:

Day 1: Cell splitting

Day 2: Transfection (30% confluence)

Fugene 6 – 3 µl (optimised as per requirement) + Plasmid – 1µg in a 50µl scale – the mixture is made in media.

This is added to 1 ml total volume media per coverslip in a 24 well plate.

Day 3: 1 set of cells (+leptomycin B control) – add LMB (20 nano molar) directly to the coverslip with no change in media – incubate at 37°C for 3 hours.

Same day: Cells are fixed.

- suck off medium,
- per well:
- add 500µl PBS ,
- suck off PBS,
- add 500µl 4%PFA in PBS @room temperature for 5 minutes,
- wash once with PBS as before,
- add 400µl of the final Hoechst solution (works fast),
- take out coverslips,
- dip them shortly into Water
- suck off excess water with filter paper at the edge and mount with mounting medium on slides (2µl drop of VectaShield mounting medium - Vector Laboratories, Burlingame, CA, USA).

2.24 CONFOCAL MICROSCOPY: This is performed using Leica SP5 confocal laser scanning microscope equipped with a 63× oil immersion objective (Leica, Wetzlar, Germany).

3 Results

3.1 eIF2 β forms an export complex with CRM1

eIF2 β is a 38kDa translation factor and it was assumed that it can enter the nucleus passively since it is smaller than the size exclusion limit of the NPC permeability barrier (40kDa) or in other cases its accidental entry can also occur during nuclear membrane reformation after cell division.

Previously, nuclear exclusion of eIF2 β was shown to be Leptomycin B sensitive, thus, CRM1 dependent. (Bohnsack et al., 2002). When CRM1 action was blocked, it resulted in rapid nuclear accumulation of eIF2 β and this could not be supported solely by passive diffusion. This suggested that eIF2 β was actively imported. The dependence of nuclear export of eIF2 β on CRM1 was re-confirmed by expressing GFP tagged eIF2 β in HeLa cells by transfection. The sensitivity of eIF2 β to Leptomycin B (LMB) was tested Figure 1. This is an anti-fungal antibiotic that selectively inhibits CRM1 function by alkylation of a Cysteine residue in its binding cleft and therefore interrupting the interaction of CRM1 with its cargo.

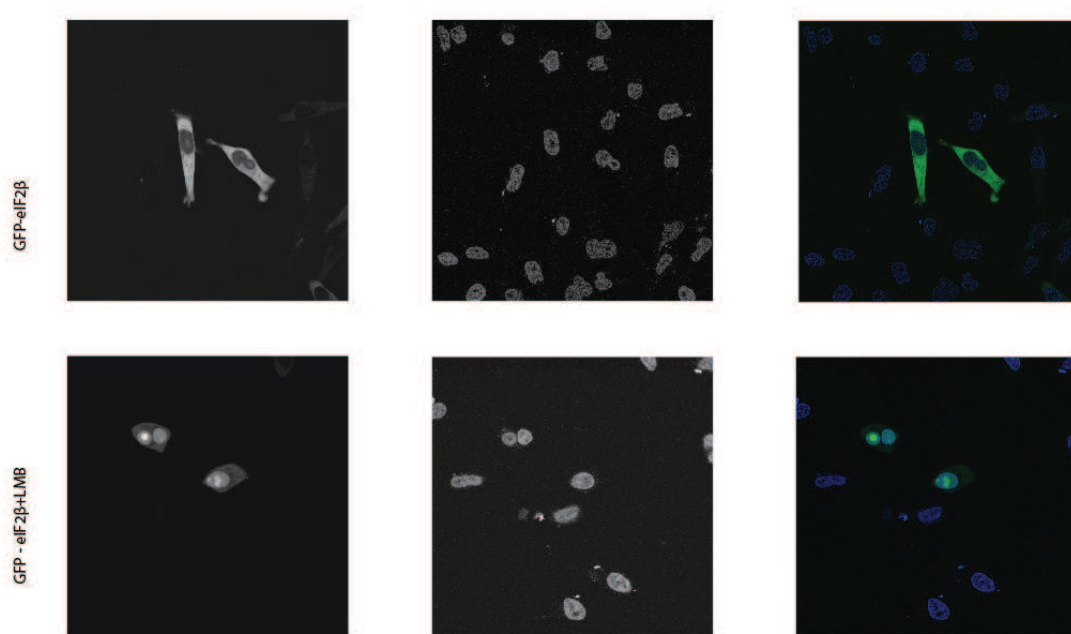


Figure 1

Figure 1: HeLa cells were transfected with a plasmid (1µg) expressing GFP tagged eIF2β. The transfecting agent used was Fugene 6 (1.8 µl per ml of media) and the cells were transfected for 24 hours. One set of the cells were fixed after 24 hours of transfection and the other set was exposed to Leptomycin B (20 nM) for 3 hours, then fixed and all these steps were executed at 37°C. The cells were fixed using PBS and 4%PFA. The nuclear DNA was stained using Hoechst solution.

In **figure 1** GFP tagged eIF2β was clearly shown to be cytosolic. When LMB was added to the transfected cells, eIF2β was localised in the nucleus proving that its export is solely CRM1 dependent.

eIF2β was expressed in BLR at 18°C overnight with 0.2 mM IPTG induction and it was purified via Nickel affinity chromatography. As seen in **figure 2**, eIF2β expresses well and is stable.

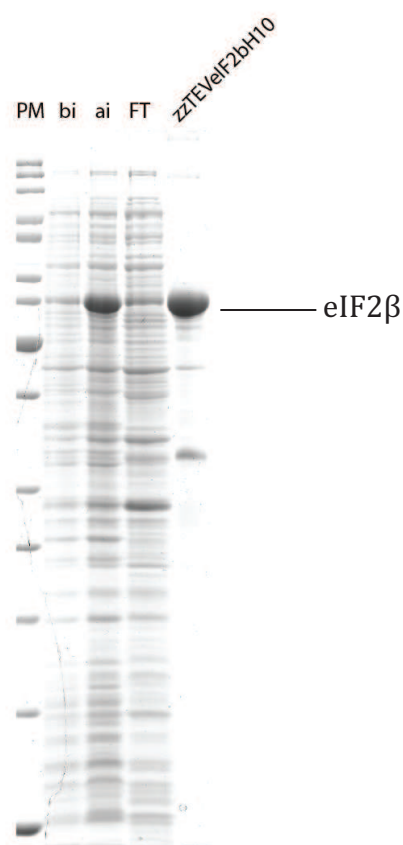


Figure 2

eIF2β was expressed in BLR at 18°C, overnight with 0.2 mM IPTG induction. bi – before induction and ai is after induction – these samples were taken directly from culture, pelleted and resuspended in SDS loading buffer. eIF2β has a Histidine tagged and was purified using nickel affinity chromatography. The flowthrough during purification consisting the unbound protein was also loaded – FT and the elution was done in elution buffer (50 mM Tris, 100 mM NaCl, 500 mM Imidazole, 5 mM DTT, 2 mM MgOAc2, pH 7.5) - labelled zzTEVeIF2bH10. 10 µl of the elution diluted to a 1:1 ratio with SDS loading buffer was analysed by SDS PAGE.

Following this, direct interaction between CRM1 and eIF2β along with RanGTP was tested in vitro by binding assays. This had been previously demonstrated by Thomas Güttler. In these assays, zz tagged eIF2β was immobilized on IgG

sepharose beads and exposed to bacterial lysate supplemented with CRM1 (1 μ M) in one set and CRM1 (1 μ M) + RanGTP (3 μ M) in the other set of samples. The immobilized cargo is exposed to a bacterial lysate, because it acts as an internal control to assess a genuine interaction by providing competition from soluble bacterial proteins. As shown in the **figure 3**, eIF2 β only interacts with CRM1 in the presence of RanGTP.

Since this interaction does not occur in the absence of RanGTP, it can be concluded that CRM1, RanGTP and eIF2 β directly interact to form an export complex.

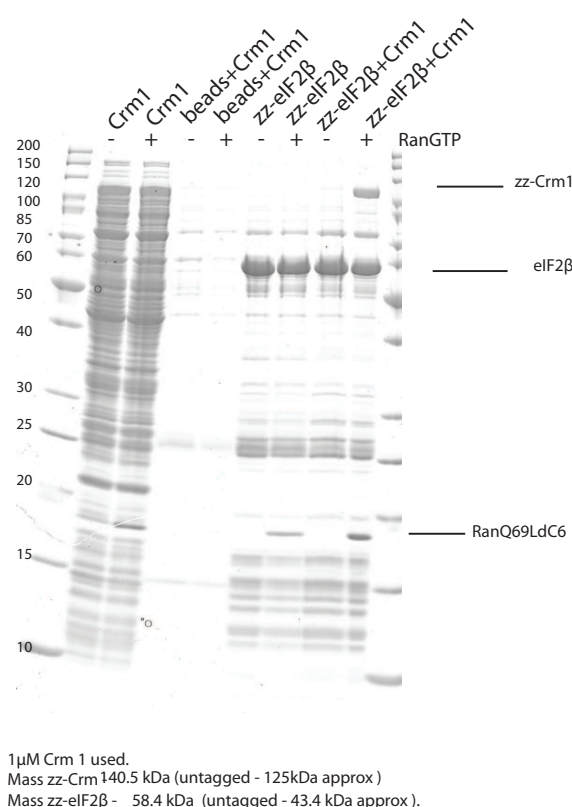


Figure 3:

eIF2 β was immobilized via the zz tag on IgG sepharose beads. It was then incubated in a bacterial lysate with different supplements as noted in the figure. The lysate is present in all samples and was used a non specific binding control in these pulldowns. From left to right the loading pattern is as follows: The input lysate with 1 μ M CRM1, lysate with 1 μ M CRM1+RanGTP (3 μ M), IgG beads incubated with: lysate+CRM1, with lysate + CRM1+ RanGTP, IgG beads immobilized with eIF2 β incubated with - lysate, lysate with RanGTP, lysate with CRM1, and lysate with CRM1+RanGTP. In all cases CRM1 concentration used is 1 μ M and RanGTP concentration used is 3 μ M. The proteins were incubated at 4°C for 3 hours, then washed and eluted with SDS loading buffer for analysis by SDS PAGE.

3.1.1 Location of eIF2 β export signal using bioinformatic servers

The next step was to elucidate the recognition mechanism of eIF2 β by CRM1. To do this, the immediate aim was to identify export signal of eIF2 β that is recognised by CRM1.

The eIF2 β sequence was submitted for analysis in bioinformatic servers for identification of a nuclear export signal that could be recognised by CRM1. There was no clear indication of the presence of an NES that fitted the classical Leucine rich consensus. This could suggest two possibilities. Either the NES prediction software was imprecise in its search and gave a false negative. The other reason suggests that eIF2 β may have a unique export signal which binds CRM1 in another site than the hydrophobic cleft. Thus, eIF2 β could employ a novel mechanism for its recognition and consequent export by CRM1.

3.2 Does eIF2 β interact with CRM1 in the binding cleft?

The hydrophobic cleft is a rigid region, which is responsible for interacting with the classical Leucine rich regions. After the crystal structure of CRM1 bound to SNF1 and RanGTP was solved, the mechanistic aspects of this interaction became clear. In the cytosolic unbound state CRM1 is in its 'relaxed conformation'. After interacting with RanGTP in the nucleus, it adopts a 'strained conformation', which suggests that there are long range effects exerted via the HEAT repeats of CRM1 (Monecke et al., 2009). This consequently opens the hydrophobic cleft and allows the NESs to gain access and fit into the rigid binding cleft of CRM1. As a consequence of this cooperation, an export complex is formed. After solving the crystal structure, mutants were generated to break the cooperative binding of RanGTP and cargo to CRM1 during formation of an export complex. One such mutant was considered and it has the mutation A541E in CRM1. A541E blocks the binding site of CRM1, because it is located in the binding region of Φ 3 hydrophobic residue and this interrupts the binding of the NES into the hydrophobic cleft of CRM1. Hence, CRM1 cannot bind to the NES whether RanGTP is present or absent (**figure 4**) (Monecke et al., 2009).

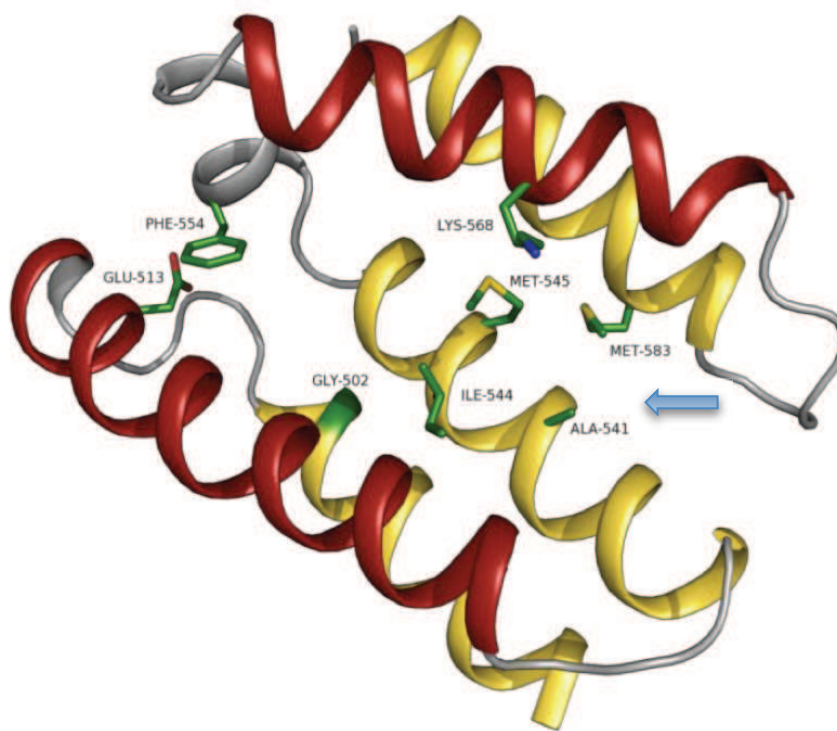


Figure 4 Hydrophobic cleft of CRM1 in the open conformation.

The residue mutated to disenable the cooperative binding of cargo and RanGTP to CRM1 has been showed by the blue arrow (figure adapted from Kevser Gencalp Lab Report 2).

In this experiment a comparative analysis was performed (experiment not shown). A classical Leucine rich NES is the PKI type NES. This is a good example to see how a classical Leucine rich NES interacts with CRM1 and the hydrophobic cleft mutant of CRM1. Binding assays of eIF2 β and PKI NES were performed with CRM1 wildtype and the A541E CRM1 mutant. The goal of this experiment is to see, if the uncoupling of binding of PKI (classical NES) and RanGTP in the CRM1 mutant A541E is also seen in eIF2 β . If this is the case, then the NES of eIF2 β does in fact bind in the hydrophobic cleft of CRM1 in a similar manner to PKI-NES.

When eIF2 β interaction with the A541E CRM1 mutant was tested, no interaction was observed irrespective of the presence or absence of RanGTP.

This comparative analysis of binding of PKI-NES and eIF2 β NES to the CRM1 mutants showed that the eIF2 β NES does in fact interact with CRM1 in the

hydrophobic binding cleft. This also suggests that the eIF2 β NES may contain a sequence similar to the PKI NES.

The following experiments were dedicated to identifying the region of eIF2 β that binds the CRM1 hydrophobic cleft.

3.3 Fragment mapping of eIF2 β to locate the residues that contain the export signal for recognition by CRM1.

Truncated versions of eIF2 β were expressed and these constructs had a Histidine tag for purification and a zz tag, which can be used to immobilize these on IgG beads for binding assays.

The binding assays were performed in the presence of a bacterial lysate supplemented with CRM1 in one case and CRM1+RanGTP in the other case. If an export complex is formed, then a trimeric complex of the cargo, RanGTP and CRM1 will be formed. Initially, five truncated constructs were tested for binding to CRM1 in the presence and absence of RanGTP and it was observed that the N terminal unstructured region bound with highest affinity. As the construct progressed with increasing deletions from the N terminal region and including the structured eIF2 β domain, the binding affinity kept decreasing (**figure 5**). Following this more truncated constructs were tested. The interactions of the different eIF2 β constructs with CRM1 in the presence and absence of RanGTP have been shown in Figures 5A, 5B and 5C.

Figure 5-A

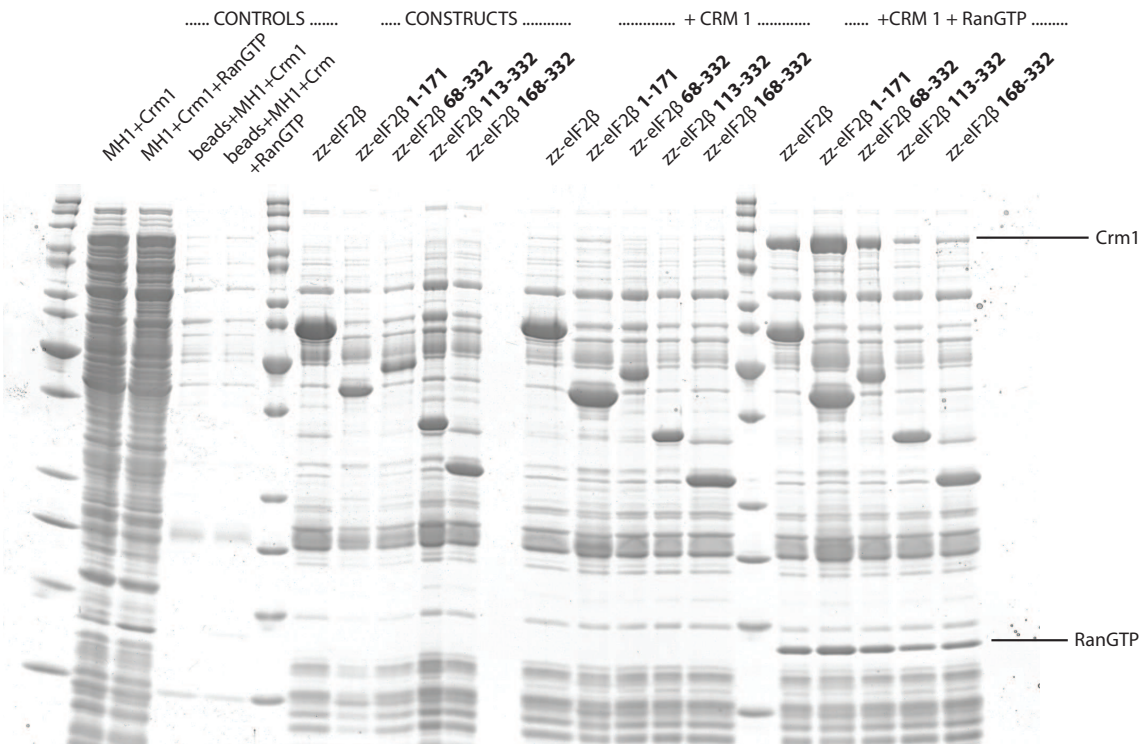


Figure 5A Interaction of eIF2β truncated constructs with CRM1 and RanGTP

Figure 5-B

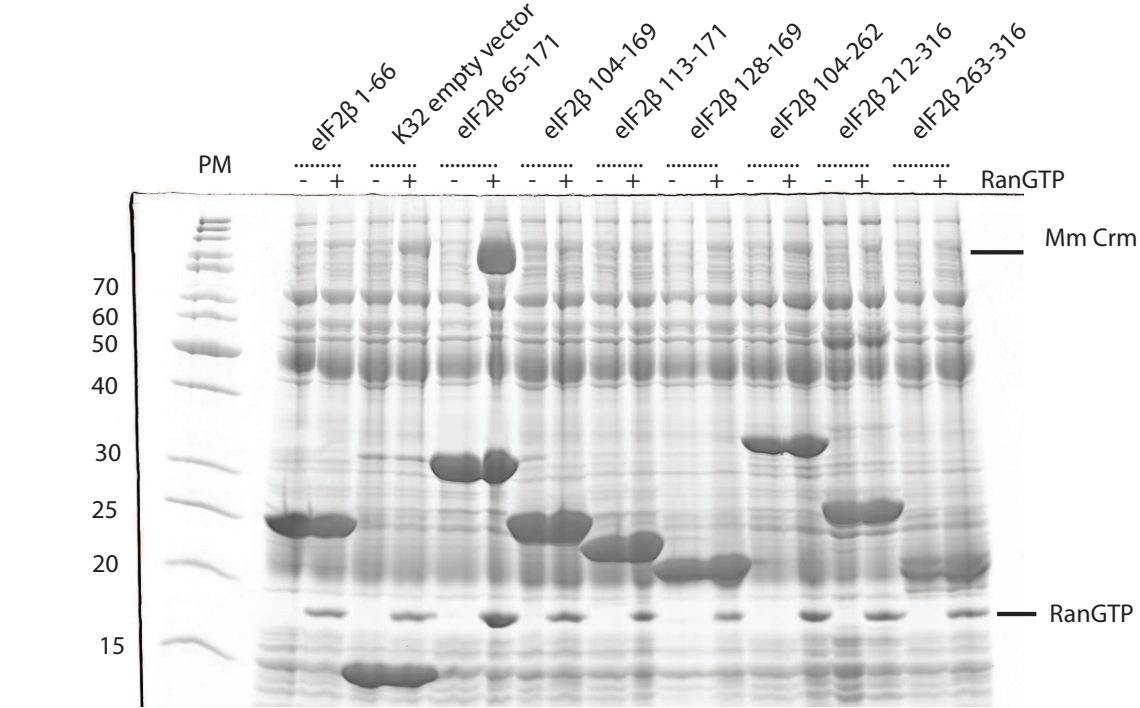


Figure 5-C

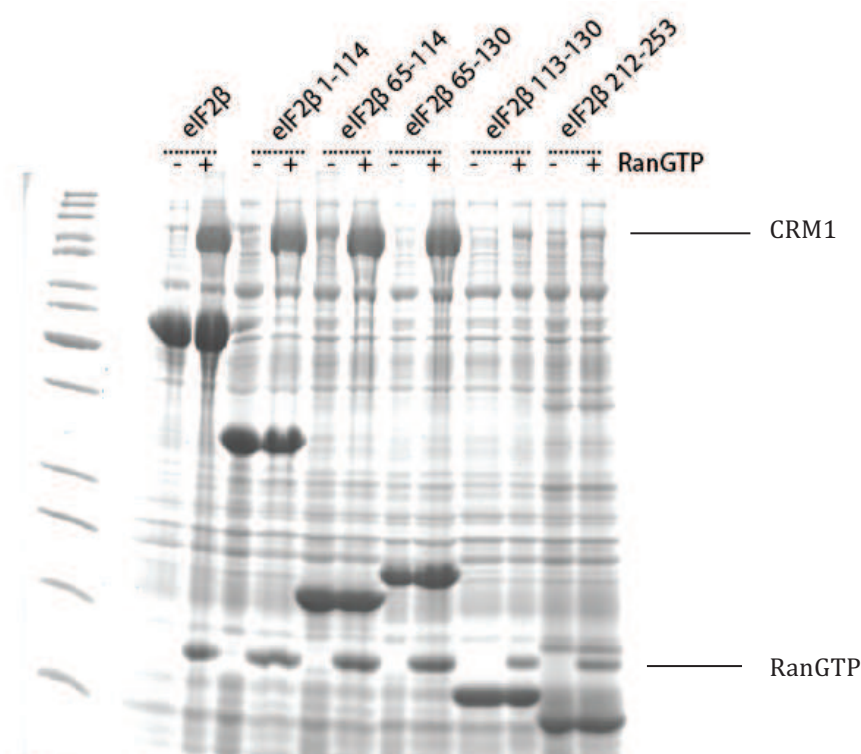


Figure 5A,5B and 5C:

The truncated constructs of eIF2 β were expressed in NEB Turbo and purified via Nickel affinity chromatography. The purified proteins also had a zz tag and were used for pulldowns. The pulldown involved immobilization of the construct via the zz tag on IgG beads (10 μ l beads, 2mg/ml amount of construct to saturate the beads). One set of these immobilized constructs were incubated with a bacterial lysate (MH1 bacterial strain) containing 1 μ M CRM1 and the other set was incubated with 1 μ M CRM1 + 3 μ M RanGTP for three hours at 4°C (500 μ l of lysate). After this, the unbound material was washed off the beads and the final elution was done in 50 μ l with SDS loading buffer without DTT. 10 μ l of this was loaded on SDS PAGE for analysis.

As seen in **figure 5A, 5B and 5C**, after performing pulldowns, residues spanning 65-114 amino acids proved to contain the minimal fragment, which binds CRM1 with the highest affinity. This region of interest lies in the unstructured part of eIF2 β and includes the second K box. The sequence of the residues spanning 65 to 114 has been shown below:

(65) ASDDLDDLNFFNQKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTPED (114)

The summary of the strength of interactions of the different constructs with CRM1 (in the presence of RanGTP) has been depicted in the figure 6.

YTYEELLNRVFNIMREKNPDMVAGEKRKFVMKPPQVVRVGTKKTSFVNFTDICKLLH
RQPKHLLAFLLAELGTSGSIDGNNQLVIKGRFQQKQIENVLRRYIKEYVTCHTCRSPDTI
LQKDTRLYFLQCETCHSRCSVASIKTGFQAVTGKRAQLRAKAN

The deleted fragment has been highlighted in **red font**.

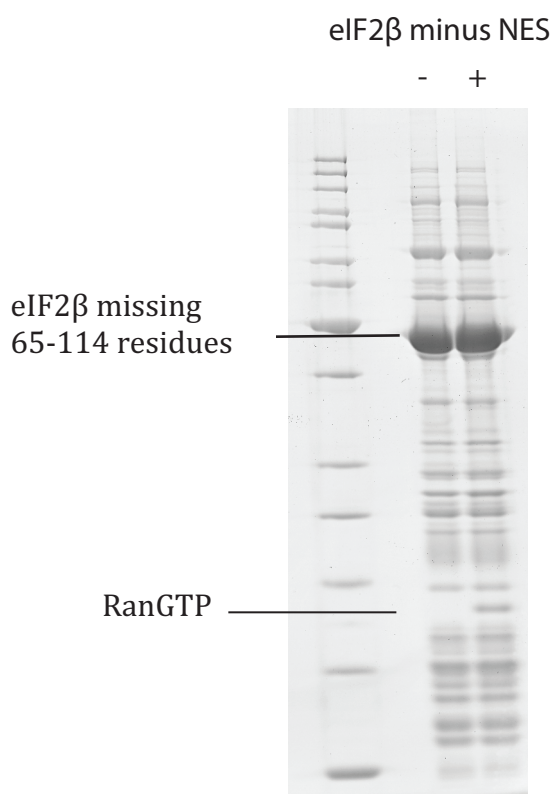


Figure 7

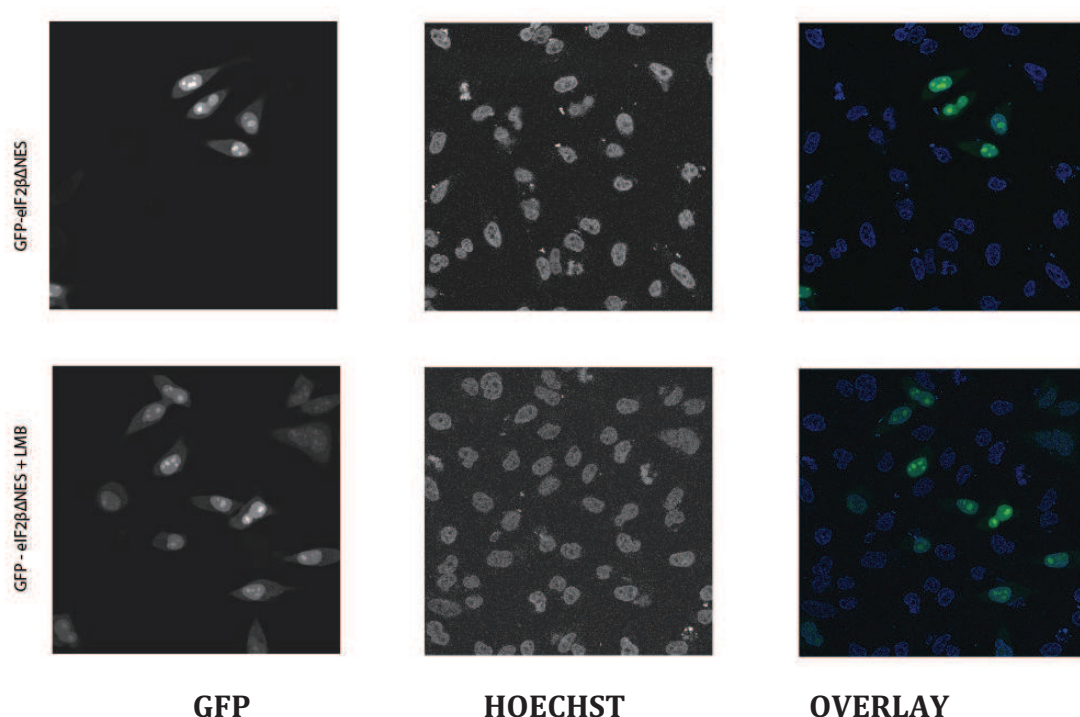
Pulldowns with CRM1 in the presence and absence of RanGTP have been depicted. The zz tagged constructs were immobilized on IgG beads and exposed a bacterial lysate with 1μM CRM1 and 1μM CRM1 + 3μM RanGTP at 4°C for 3 hours. The samples were eluted with 50μl of SDS PAGE loading buffer and 10μl was loaded for analysis.

An *in vivo* validation was done using transfection to prove that the nuclear export of eIF2β depends on the fragment spanning residues 65-115 (**figure 8**). GFP tagged eIF2β constructs were expressed in HeLa cells and their location of these GFP- eIF2β constructs was analysed. The full length eIF2β (shown in Figure 1) was shown to be cytosolic. Here, the identified fragment suggested to contain the NES (65-114 residues) and full length eIF2β missing the 65-114 amino acid stretch were transfected. The GFP tagged eIF2β domain at the C terminus was also tested to confirm that this does not contribute to the nuclear export of eIF2β.

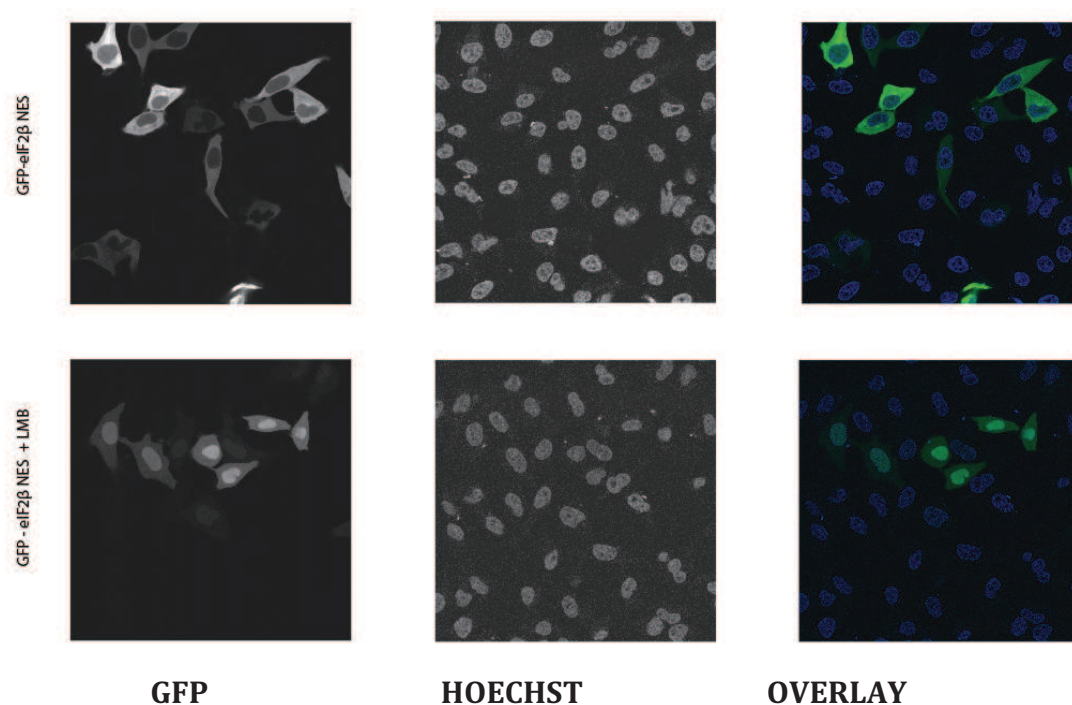
The full length eIF2 β and the identified NES were localised in the cytosol. On the other hand, the full length eIF2 β missing the 65-114 amino acid stretch and the eIF2 β domain at the C terminus were found localised in the nucleus. This validates that the nuclear export of eIF2 β is dependent on the region spanning amino acids 65-114.

The dependence of nuclear export of eIF2 β export on CRM1 was also confirmed by performing the transfection in the presence of Leptomycin B. It was observed that all the constructs remained localised in the nucleus when exposed to Leptomycin B highlighting the fact that this export solely depends on CRM1.

A) GFP tagged eIF2 β missing the 65 to 114 amino acid residues are located in the nucleus.



B) GFP tagged 65 to 114 amino acids of eIF2 β are located in the cytosol.



C) GFP tagged eIF2 β domain at the C terminus is located in the nucleus.



Figure 8A, 8B and 8C:

HeLa cells were transfected with a plasmid (1 μ g) expressing GFP tagged eIF2 β constructs. The transfecting agent used was Eugene 6 (1.8 μ l per ml of media) and the cells were transfected for 24 hours. One set of the cells were fixed after 24 hours of transfection and the other set was exposed to Leptomycin B (20 nM) for 3 hours before fixing and all these steps were executed at 37°C. The cells were fixed using PBS and 4%PFA and the nuclear DNA was stained using Hoechst solution.

The next step was to identify the exact residues that contribute to the export signal. To narrow down the region that interacts with CRM1 to form an export complex two more deletion constructs were made within the 65 to 114 amino acid stretch (**figure 9**).

Further deletions within the small 65 to 114 eIF2 β construct:

(65) ASDDLDDLNFFNQKKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTPED (114)

Construct A:

ASDDLDDLNFFNQKKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTPED
deleted

Construct B:

ASDDLDDLNFFNQKKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTPED
 deleted

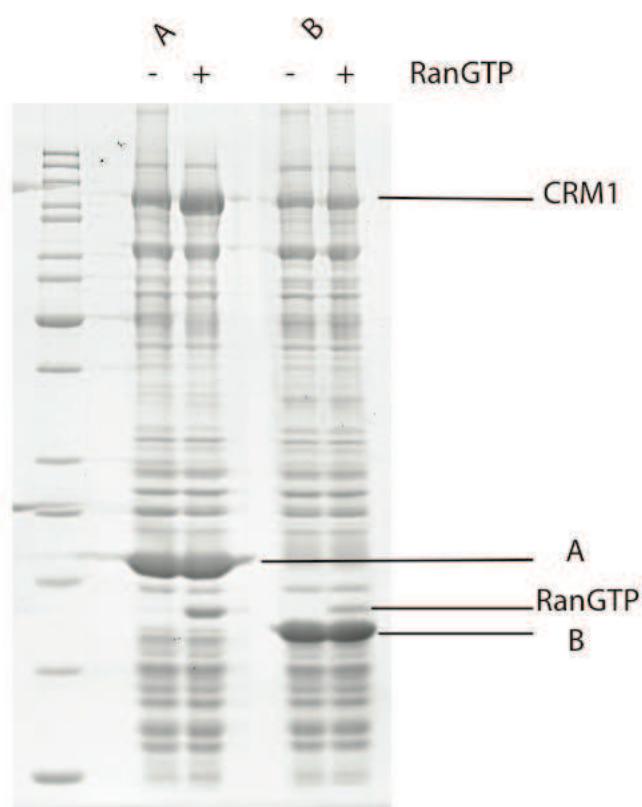


Figure 9

Pulldowns with CRM1 in the presence and absence of RanGTP have been depicted. The immobilized constructs on the IgG beads are the zz tagged ones. The zz tagged construct saturated the IgG beads and the constructs were exposed to a bacterial lysate with 1 μ M CRM1 and 1 μ M CRM1 + 3 μ M RanGTP at 4°C for 3 hours. The elution was performed using 50 μ l of SDS loading buffer and 10 μ l was analysed by SDS PAGE.

Based on the pulldowns, it was clear that nuclear export signal was contained in the residues spanning 87 to 114 or IFDIDEAEEGVKDLKIESDVQEPTPED, since the construct missing this region did not interact with CRM1 irrespective of the presence or absence of RanGTP. It is also clear that the stretch 65 to 75 residues or ASDDLDDLNFF does not contribute to the export signal, since its absence did not prohibit interaction of the eIF2 β with CRM1 + RanGTP.

3.4 Deciphering the NES binding sequence in eIF2 β :

From the previous experiments it was clear that the fragment corresponding to the amino acids 87 to 114 in the N terminal unstructured part of eIF2 β contains the NES and this interaction occurs in the binding cleft of CRM1. This is a strong indication that the export signal should correspond to the classical Leucine rich signal. The NES consensus we referred to was established in the Güttler et al., 2010 and has been shown below:

$x\Phi^0xx\Phi^1xxx\Phi^2xx$ or $xxx\Phi^3x\Phi^4$

The hydrophobic residues in this stretch were designated the name ' Φ residues' (numbered from 0 to 4). The 'x' denotes the amino acids surrounding the hydrophobic residues or also called the 'spacer residues'

The sequence of this fragment was re-analysed to locate residues, which may fit the consensus defined for the classical Leucine rich export. Here, in addition to the defined requirements of the NES e.g. favoured hydrophobic residues, appropriate spacing etc. another aspect was considered.

According to the defined consensus, Leucine, Isoleucine, Valine, Phenylalanine etc. are residues preferred at the hydrophobic Φ positions of the NES. However, if each Φ position was occupied by a 'perfectly matching' residue for maximum binding, then the strength of the NES would be unnaturally strong. The NES should have enough affinity for the cargo to be successfully translocated to the appropriate destination. However, following transport, the cargo should also be

easily disassembled from the transport receptor. If the NES is extremely strong, its disassembly may prove to be difficult and this would result in a futile cycle with the cargo remaining bound to the transport receptor. Thus, sometimes, the hydrophobic positions are occupied by residues that are less favoured and do not contribute substantially to the binding to the transport receptor. In fact their actual role is to modulate the affinity of the NES to the receptor as required.

Keeping in mind that a best binding hydrophobic residue is not needed at every Φ position, the sequence was re analysed to locate a sequence that resembles the defined consensus. Some NES like sequences had been predicted based on the NES prediction algorithm (courtesy of Koray Kirli). Based on the previously described biochemistry analysis, it was clear that the sequence highlighted in yellow contained the NES. Further analysis helped us to narrow down and suggest the marked residues as the hydrophobic amino acids that may occupy the Φ positions of the NES. An amino acid stretch (sequence highlighted in yellow- figure 10) was speculated to form the NES. Here an Alanine occupies the Φ_1 position of the NES.

Hs eIF2 β

SGDEMIFDPTMSKKKKKKKPFMLDEEGDTQTEETQPSETKEVEPEPTEDKDLEADE
EDTRKKDASDDLDDLNFFNQKKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTEPEDD
LDIMLGNNKKKKKKNVKFPDEDEILEKDEALEDEDNKKDDGISFSNQTPAWAGSERD
YTYEELLNRVFNIMREKNPDMVAGEKRKFVMKPPQVVRVGTKKTSFVNFTDICKLLH
RQPKHLLAFLLAELGTSGSIDGNNQLVIKGRFQQKQIENVLRRYIKEYVTCHTCRSPDTI
LQKDTRLYFLQCETCHSRCSVASIKTGFQAVTGKRAQLRAKAN

DIDEAEEGVKDLKI

0 1 2 3 4

Figure 10 - predicted NES in Hs eIF2 β

To validate our approach, the amino acids occupying the Φ positions were mutated successively to Alanine, Glycines, Serines and Leucines. These NES mutants were then analysed for their binding capacity to CRM1 in the presence and absence of RanGTP (**figure 11**).

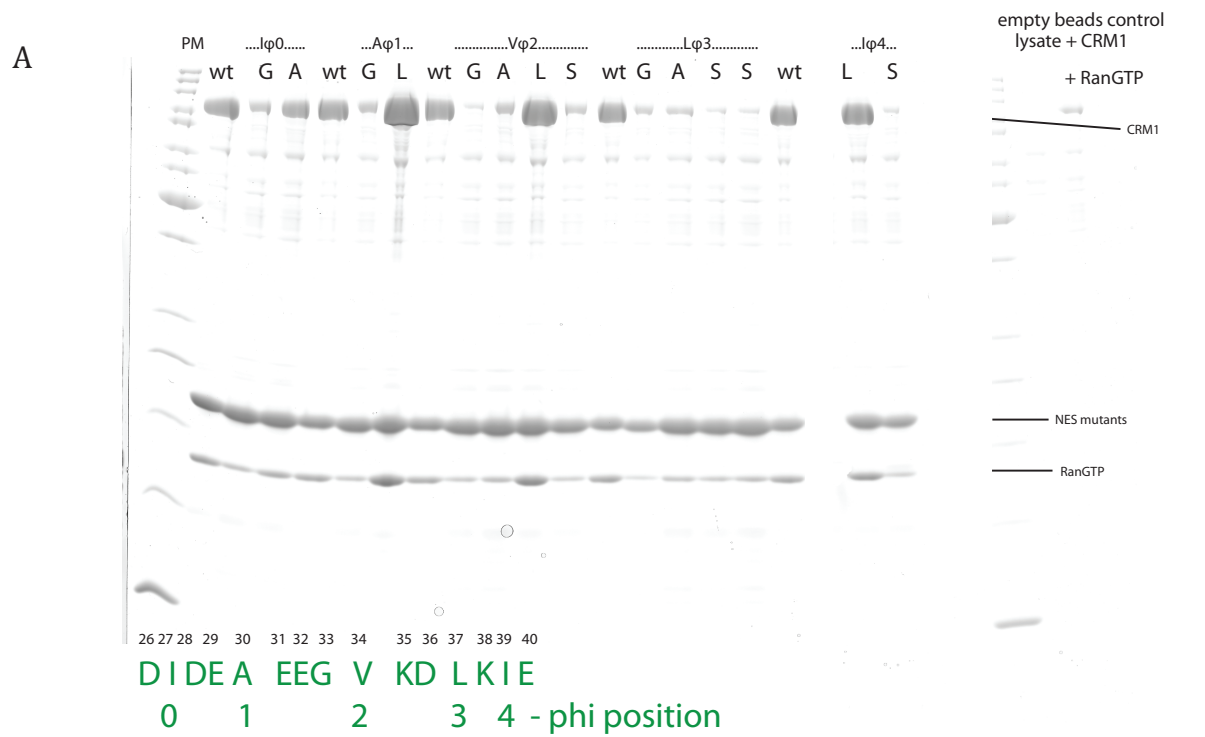
Pulldowns with CRM1 in the presence and absence of RanGTP are depicted in figure 11. The immobilized constructs on the IgG beads are the zz tagged. The zz tagged construct saturated the IgG beads and the constructs were exposed to a bacterial lysate with 1 μ M CRM1 and 1 μ M CRM1 + 3 μ M RanGTP at 4°C for 3 hours. The results of the binding assays were analysed on SDS PAGE.

Following the binding assay, it was clear that the mutation of these designated Φ residues had a dramatic influence on the interaction of the NES with CRM1. When the Φ 0 residue was mutated to Glycine the binding went down dramatically and when changed to Alanine, there was a slight decrease in binding strength compared to the wildtype NES.

The Φ 1 position is occupied by an Alanine and this is an atypical residue for a Leucine rich classical NES. When this position was changed to a Glycine, there was a dramatic decrease in the binding to CRM1. On the other hand, when this Φ 1 Alanine was mutated to a Leucine, there was a dramatic increase in binding to CRM1. Thus, this seems to be a sensitive point in the NES and we expected that the type of residue present at this location in the NES has a great influence on the affinity of the NES for CRM1. This point has been discussed in later experiments.

The Valine residue (Φ 2) was mutated to Glycine and this abolished binding to CRM1. Mutation to Alanine and Serine showed a dramatic decrease in affinity to CRM1. When Valine was mutated to a Leucine, there was an evident increase in the binding affinity. Therefore this is a sensitive position as well in the NES.

When the Leucine residue or Φ 3 was mutated to a Serine, Glycine and Alanine, the binding to CRM1 was abolished. Hence a Leucine seems to be crucial at this specific position of the NES.



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

DASDDLDDLNFFNQKKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTEPE

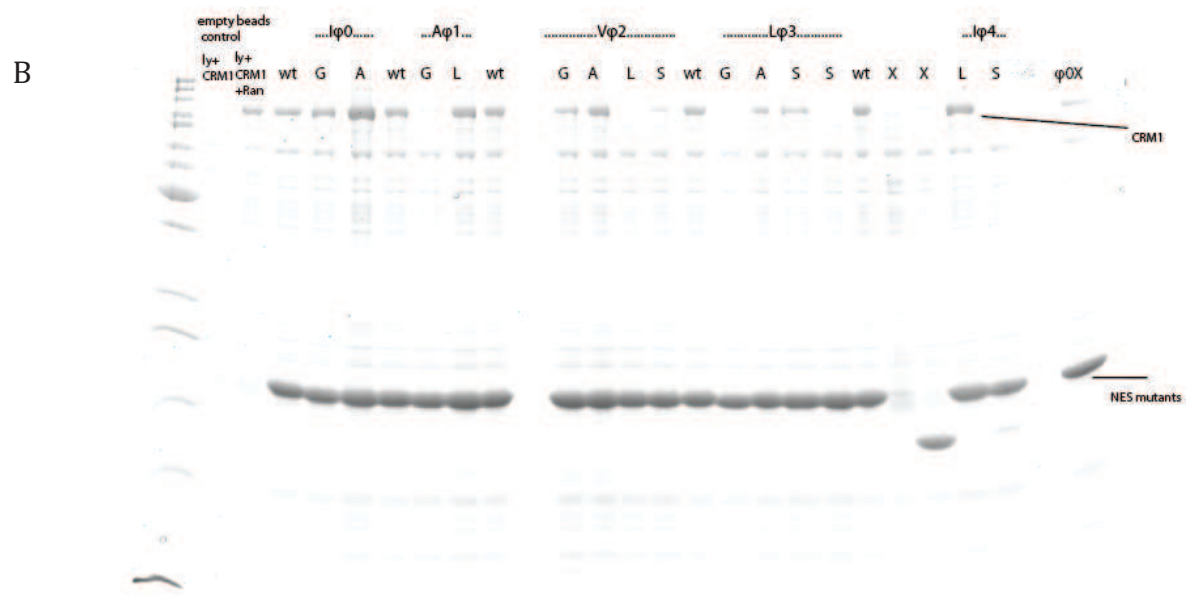


Figure 11

Figure 11: The mutated constructs of eIF2 β -NES were immobilized via the zz tag on IgG beads (10 μ l beads, 2mg/ml amount of construct to saturate the beads). One set of these immobilized constructs were incubated with a bacterial lysate (MH1 strain) containing 1 μ M CRM1 and the other set was incubated with 1 μ M CRM1 + 3 μ M RanGTP for three hours at 4°C (500 μ l of lysate). After this, the unbound material was washed off the beads and the final elution was done in 50 μ l with SDS loading buffer without DTT. 10 μ l of this was loaded on SDS PAGE for analysis. The CRM1 used is wt. The variation in the binding affinity between CRM1 wt and eIF2 β -NES mutants in the presence of RanGTP has been depicted (A) and in the absence of RanGTP (B). Various hydrophobic residues contributing to the Φ positions of the NES have been mutated.

Lastly, when the Isoleucine at Φ 4 position was mutated to a Serine, the binding to CRM1 was abolished and a change to Leucine did not result in a dramatic change in binding affinity (**figure 11**).

In figure 11B, the binding of the mutants in the absence of RanGTP was tested. It is important to consider that there is some background binding of CRM1 to the IgG beads. But it is evident that in some cases, a weak interaction of the mutants with CRM1 can be observed in the absence of RanGTP too.

This analysis explains why the conventional bioinformatic portals could not detect the NES in eIF2 β because of the Alanine at the Φ 1 position. Alanine is an atypical residue for a hydrophobic position. As seen before, this residue actively contributes to the NES interaction with CRM1 and when mutated to Arginine abolishes binding to CRM1. When the other Φ residues were mutated to Alanines to check for the effect on binding to CRM1 and RanGTP, it was clear that this residue dramatically decreases interaction with CRM1 at all Φ positions other than Φ 1.

A glycine scan was done for the hydrophobic Φ residues. These Φ residues of the predicted NES were systematically mutated into Glycines and expressed in HeLa cells with a GFP tag by transfection. In figure 12, the mutations were performed on the full length protein and analysed to check the location of the mutated protein. In all cases, eIF2 β was shown to be localised in the nucleus proving that mutation of any Φ residue to Glycine renders the NES inactive for nuclear export.

eIF2 β Glycine mutants of the hydrophobic Φ residues

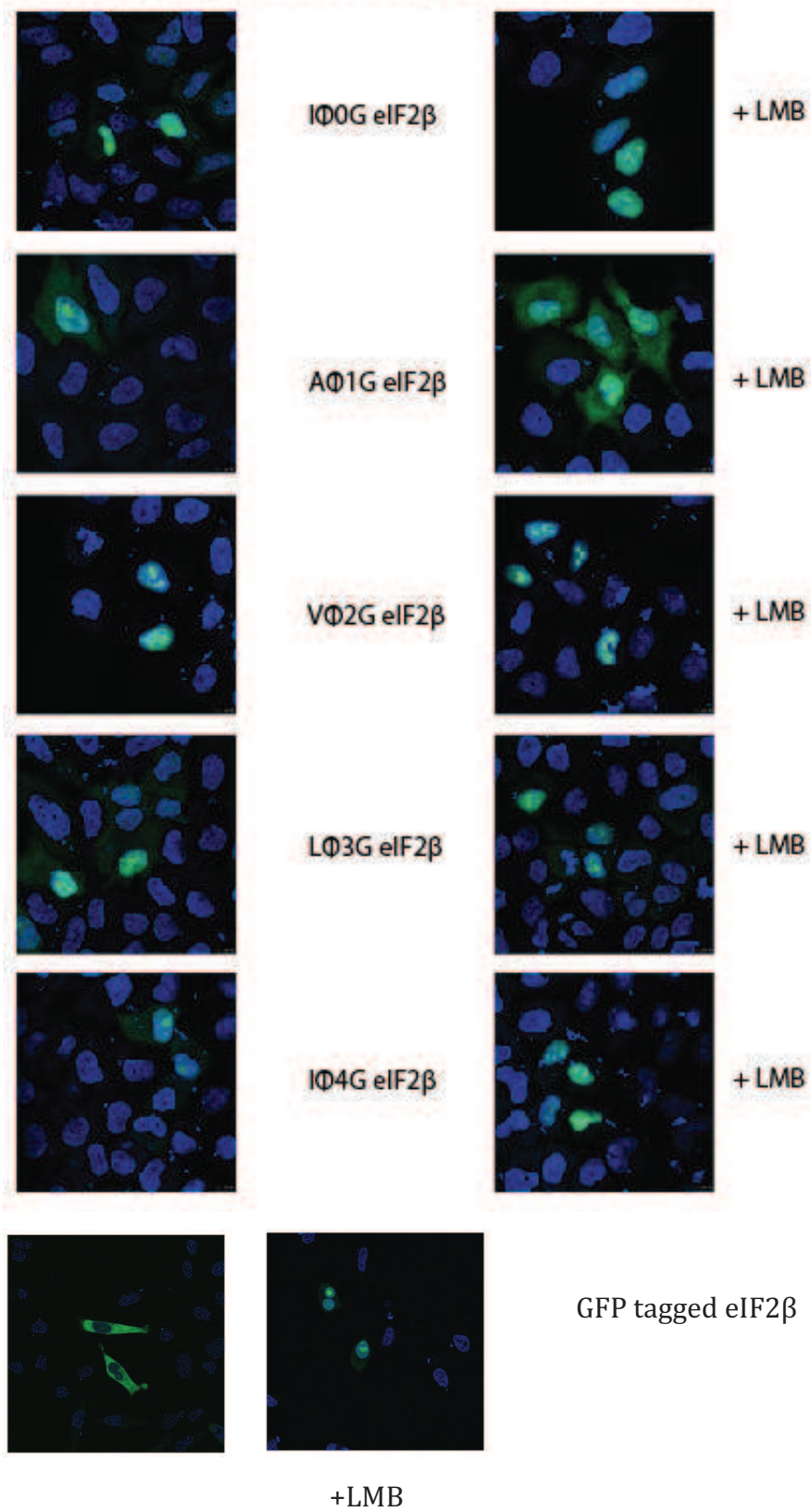


Figure 12

Figure 12: HeLa cells were transfected with a plasmid (1µg) expressing GFP tagged eIF2β constructs. The transfecting agent used was Fugene 6 (3 µl per ml of media) and the cells were transfected for 24 hours. One set of the cells were fixed after 24 hours of transfection and the other set was exposed to Leptomycin B (20 nM) for 3 hours before fixing and all these steps were executed at 37°C. The cells were fixed using PBS and 4%PFA and the nuclear DNA was stained using Hoechst solution.

As a control, each sample was also tested in the presence of LMB to validate CRM1 dependent export. Since eIF2β has poly Lysine tracks spanning its unstructured and charged N terminal region, it has a tendency to accumulate in the nucleoli - this is the reason for non uniform distribution of the GFP tagged mutants in the nucleus.

Another round of validation was done by creating Glycine mutants of the eIF2β Φ residues on the NES template alone. Here as well, there were two sets of samples - with and without LMB. As seen in **Figure 13**, in the NES constructs as well, when the Φ residues are mutated successively to Glycines, the NES of eIF2β is not recognised by CRM1 for nuclear export.

There were two residues, Isoleucine and Phenylalanine, which were immediately adjacent in location to the identified NES. The nature of the amino acids suggested that they could participate in the NES binding. As shown by previous research, in the classical NES there is a prevalence of residues like Isoleucine, Phenylalanine and Valine. Hence, the Phenylalanine and Isoleucine were mutated one at a time to Arginine. These were tested for interaction with CRM1 in the presence of RanGTP to check for formation of an export complex.

Regarding residue numbering, please refer to the numbering shown in the sequence below. The hydrophobic residues suggested to contribute to the NES are highlighted in **yellow**. The I and F residues being tested are shown in **bold print**. Suggested eIF2β NES:

I	F	D	I	D	E	A	E	E	G	V	K	D	L	K	I	E
87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103

eIF2 β NES Glycine mutants of the hydrophobic Φ residues

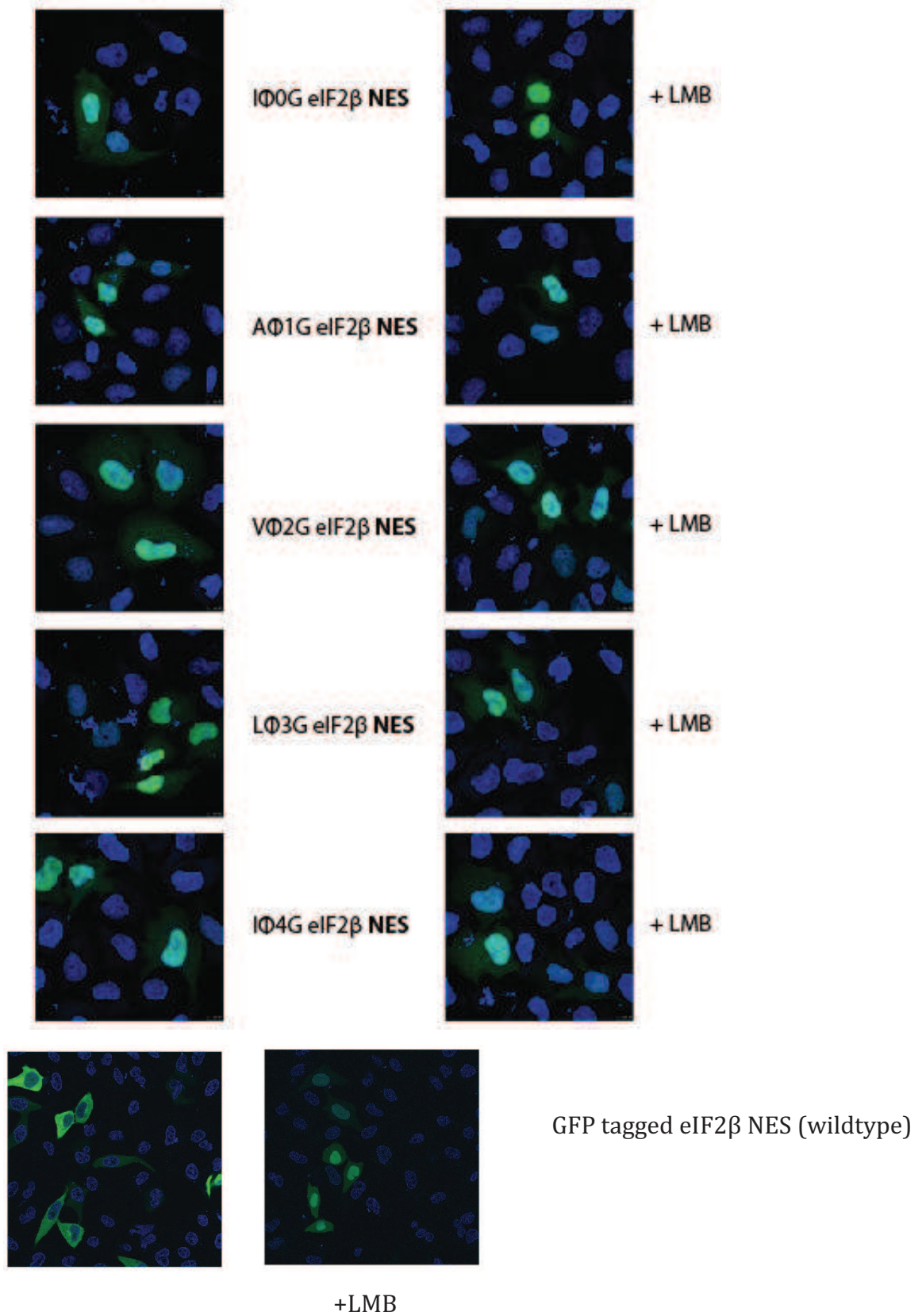


Figure 13

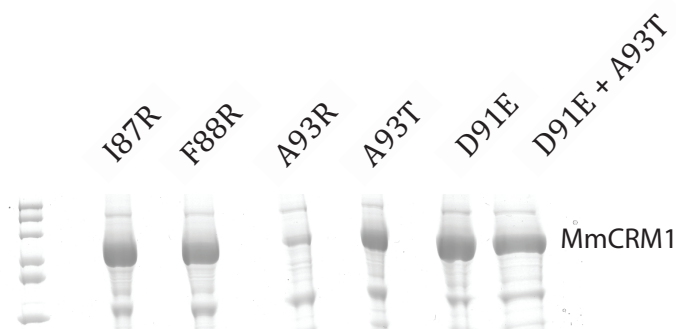
The procedure for transfection is the same as described in Figure 12.

When binding assays were performed with these mutants, it was clear that there was no change in the binding to CRM1 compare to the natural NES sequence. This ensured that these residues do not contribute to the NES (**figure 14**).

In figure 14, The Alanine at the Φ 1 position was mutated to Arginine to see if this residue contributes affinity of the NES for CRM1 for export. A Φ 1R abolished export complex formation. After mutagenesis, it was clear that Alanine does infact play an active role in modulating the affinity of the eIF2 β NES to CRM1 for nuclear export.

3.5 Comparison of eIF2 β NES in different species: Modifying the *Homo sapien* NES to match the eIF2 β NES from *Xenopus tropicalis*

The sequence of eIF2 β was compared to other homologs and it highlighted the importance of the identified NES sequence, because it was highly conserved. In *Xenopus tropicalis* there were subtle changes in one spacer residue and Φ 1 position in the NES when compared to the *Homo sapien* NES. The *Homo sapien* eIF2 β NES was mutated to match the *Xenopus tropicalis* NES. A slight decrease was observed in the binding to CRM1+RanGTP by mutating Alanine Φ 1 (residue 93 in the sequence below) residue to Threonine, (spacer) Aspartate residue (no 91) in the NES to Glutamate and when both the mutations were combined in one construct. This proved to be a cross species validation of the NES sequence (figure 14).



Mutating I87 to R and F88 to R to confirm if they also contribute to the NES.

DASDDLDDLNFFNQKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTPE

Mutating A93 to R and see if this atypical hydrophobic residue can be replaced by a polar residue and be tolerated in the NES.

DASDDLDDLNFFNQKKKKKTKKIFDIDEAEEGVKDLKIESDVQEPTPE

Adapting the Hs eIF2 β NES to Xt eIF2 β by successively mutating D91 to E, A93 to T and then combining both mutations.

	LEE	T	EEG	V	KN	L	KI	- Xt
	D	I	E	A	EEG	V	KD	LKI - 2 β Hs
	0	1		2		3	4	
		E						
			T					
		E	T					

Figure 14

The eIF2 β -NES containing fragment has been mutated to analyse, if the hydrophobic residues adjacent to the identified Φ residues also contribute to the binding with CRM1 in the presence of RanGTP. A cross species validation was performed by analysis of mutants modified to match *Homo sapien* NES to the *Xenopus tropicalis* NES. A Φ 1R was also tested. The mutants were immobilized on IgG sepharose, exposed to a bacterial lysate, supplemented with 1 μ M CRM1 and 3 μ M RanGTP for 3 hours at 4°C. The complex was eluted by SDS loading buffer and analysed by SDS page. The sequence of the construct is shown to clarify the position of the mutations.

3.6 Alanine Φ 1 plays an active role in modulating the eIF2 β NES affinity for CRM1

Alanine Φ 1 plays an active role in influencing the affinity of the NES to CRM1, as seen in the pulldowns. To further validate this, the GFP tagged NESs with the Alanine Φ 1 mutants - A Φ 1R and A Φ 1T were expressed in HeLa cells by transfection (figure 15). As compared to the wildtype NES, it is clear that A Φ 1R renders the NES inactive and A Φ 1T decreases the affinity of the NES for CRM1 - hence in this case the nuclear exclusion is not as strong as the wildtype NES.

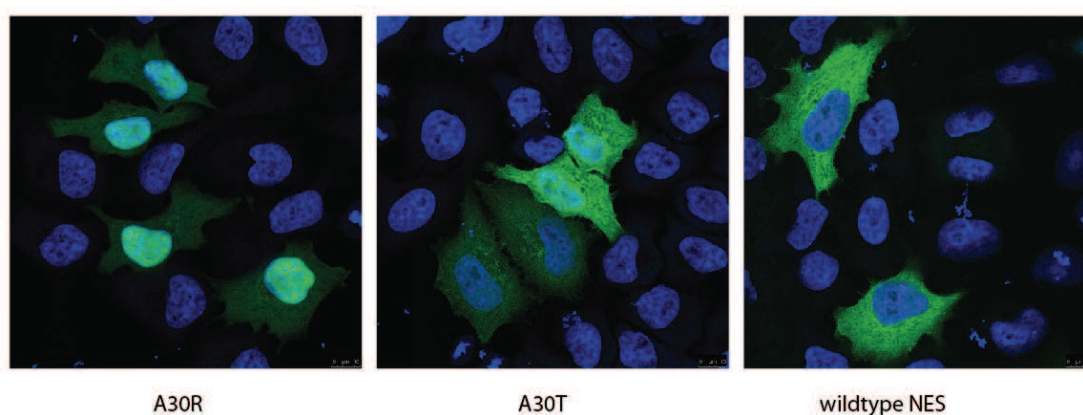


Figure 15:

HeLa cells were transfected with a plasmid (1 μ g) expressing GFP tagged eIF2 β constructs. The transfecting agent used was Fugene 6 (3 μ l per ml of media) and the cells were transfected for 24 hours. One set of the cells were fixed after 24 hours of transfection and the other set was exposed to Leptomycin B (20 nM) for 3 hours and all these steps were executed at 37°C. The cells were fixed using PBS and 4%PFA and the nuclear DNA was stained using Hoechst solution.

3.7 Analysis of the PKI-NES Φ residue mutants: introducing atypical residues at the Φ positions in the NES

The work done so far has crucially depended on the NES consensus derived from the classical PKI NES. In the eIF2 β NES, surprisingly, an Alanine was found at the Φ 1 position. In the Xenopus eIF2 β NES, a Threonine was located at the Φ 1 position. These are both atypical Φ residues, and we wanted to address the issue

of the presence of atypical Φ residues in the context of a classical NES - the PKI NES.

Here we analysed a variant of the PKI NES with a high affinity for CRM1 - called super-PKI or S-PKI. In this, the Φ positions were modified successively to residues which are not usually found in NES hydrophobic residues e.g. Cysteines, Serines, Alanines and Threonines. The binding of these constructs to CRM1 was analysed by binding assays. A high affinity variant was used for the binding assays with CRM1 in the presence and absence of RanGTP to detect subtle changes in affinity (**figure 16**).

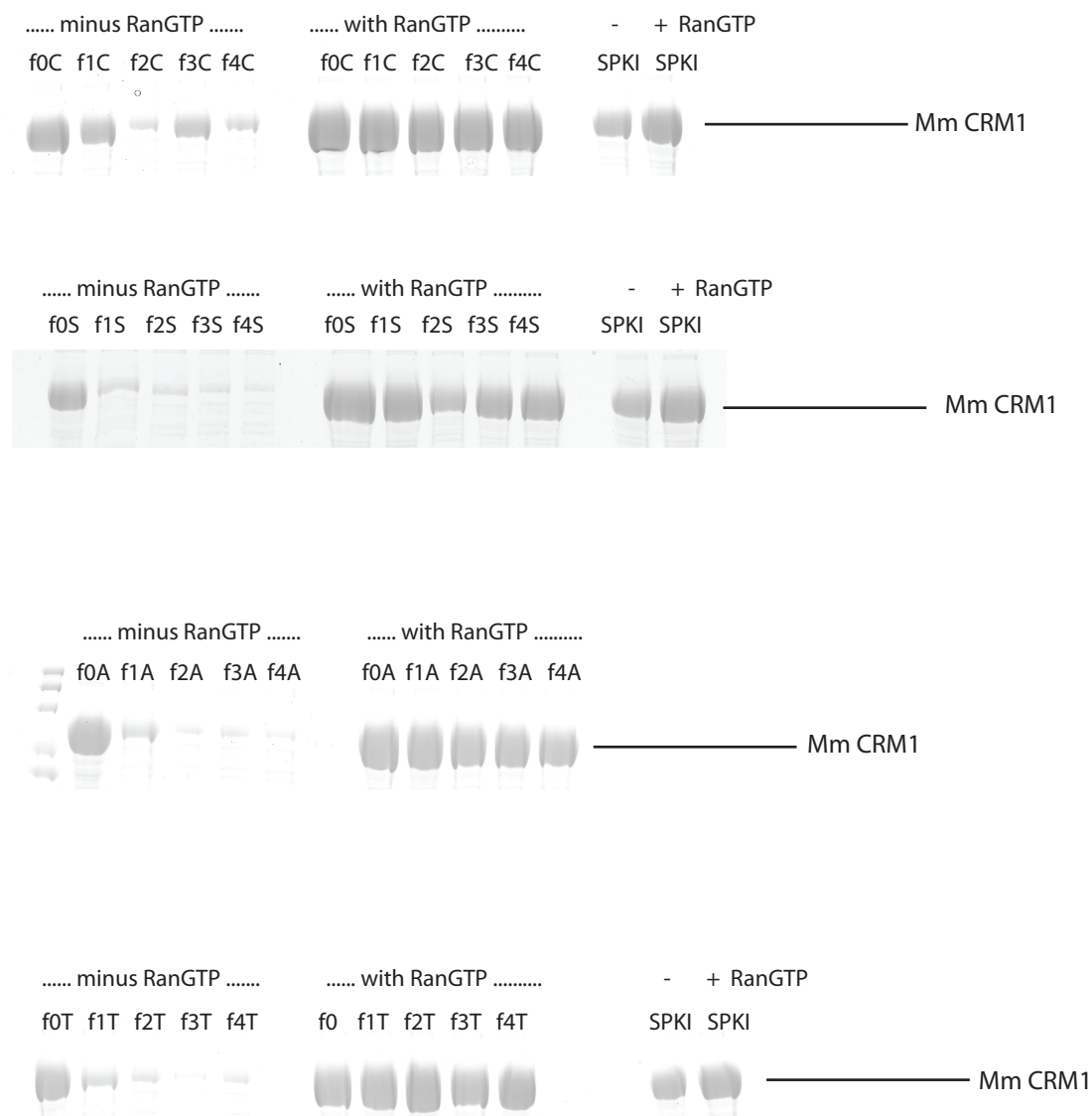
These mutants bind very strongly to CRM1 in the presence of RanGTP as they are high affinity variants. An interaction with CRM1 was evident even in the absence of RanGTP, but it was considerably weaker when compared to the condition with RanGTP. Based on the pulldowns, in the minus RanGTP binding condition, the changes in affinity of the different mutants for CRM1 was apparent.

The Φ 0 position was most tolerant to the mutations and showed little variation in binding to CRM1 in the presence of RanGTP. Even in the absence of RanGTP, the different Φ 0 mutants showed least sensitivity to the introduction of atypical Φ residues. This suggests that this Φ residue may even be skipped. The Φ 0 position of eIF2 β was also not as sensitive as the other Φ positions to the introduction of Glycines and Alanines suggesting that NESs that follow the classical consensus can tolerate a greater variety of residues at the Φ 0 position. One reason for this may be the structure of the binding cleft. It may allow greater variety in the residues that can be allowed at this position.

All the Φ 1 and Φ 3 mutants showed strong interaction with CRM1 in the presence of RanGTP. In the minus RanGTP condition, Φ 1 and Φ 3 mutants showed maximum variation in binding to CRM1 depending on the residue introduced. In the absence of RanGTP, these two positions tolerated Cysteins and showed an interaction with CRM1, but when Serines, Alanines or Threonines were introduced, they showed no interaction with CRM1.

Thus, Cysteine can be tolerated at $\Phi 1$ and $\Phi 3$ positions in the PKI NES. $\Phi 1$ position in both PKI NES and eIF2 β NES can tolerate atypical residues.

The $\Phi 2$ and $\Phi 4$ positions were sensitive to the mutations introduced and interacted with CRM1 with high affinity only in the presence of RanGTP and in the absence of RanGTP, there was an abolishment of interaction with CRM1.



f denotes the phi position in the NES binding site. S PKI is the control.

Figure 16

Super PKI NES mutants were analysed to test their affinity to CRM1 in the presence and absence of RanGTP. f denotes Φ position and the number denotes the position in the classical NES consensus. The constructs were zz tagged and immobilized on IgG sepharose beads and incubated 1 μ M CRM1 and 1 μ M CRM1 + 3 μ M RanGTP in binding buffer (50mM Tris, 50mM NaCl, 2mM MgOAc₂, 0.05% Digitonin, pH 7.5). The eluted samples were analysed on SDS page.

3.8 Analysis of the atypical Φ 1 Alanine residue of the eIF2 β NES: Understanding which residues can be permitted in this position and allow NES activity

Due to previous experiments, it was clear that the type of amino acid present at Φ 1 position of the eIF2 β NES has a dramatic effect on the formation of an export complex with CRM1. This encouraged us to examine Φ 1 position in more detail. When an Arginine or Glycine was introduced, the interaction was abolished, but upon introduction of a Leucine, there was a dramatic increase in the interaction with CRM1 and RanGTP. The Alanine to Leucine mutant where the affinity for CRM1+RanGTP increases dramatically, was also tested by transfection. GFP tagged eIF2 β Alanine Φ 1 to Leucine mutant was expressed in HeLa cells and was found to be strictly excluded into the cytosol - Figure 17. This further validates that Alanine Φ 1 is a sensitive point and can modulate the affinity of the NES to CRM1+RanGTP.

Nuclear exclusion of GFP tagged eIF2 β Alanine Φ 1 to Leucine mutant.

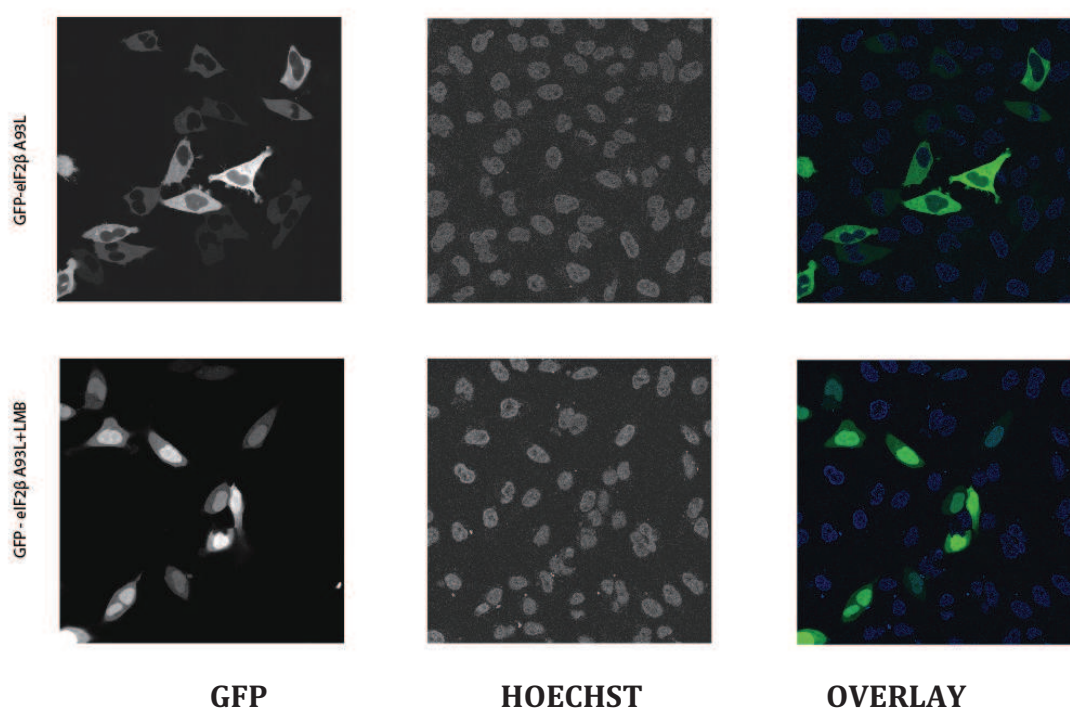


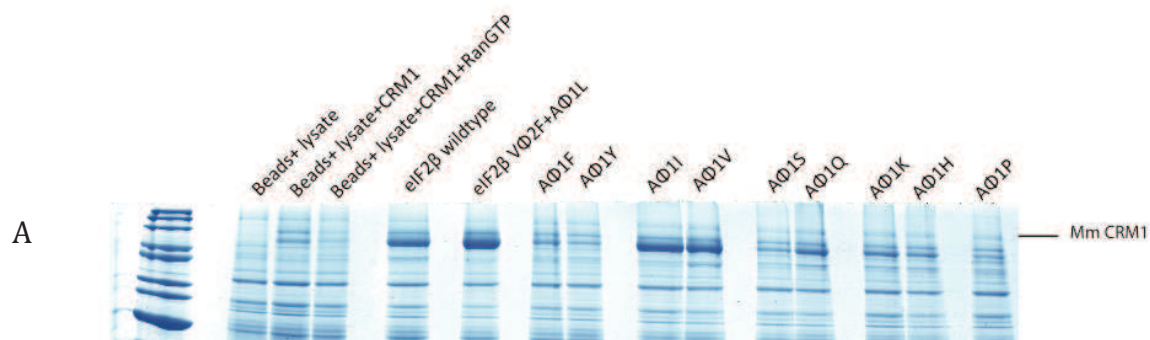
Figure 17

HeLa cells were transfected with a plasmid (1 μ g) expressing GFP tagged eIF2 β constructs. The transfecting agent used was Fugene 6 (1.8 μ l per ml of media) and the cells were transfected for

24 hours. One set of the cells were fixed after 24 hours of transfection and the other set was exposed to Leptomycin B (20 nM) for 3 hours and all these steps were executed at 37°C. The cells were fixed using PBS and 4%PFA and the nuclear DNA was stained using Hoechst solution.

We wanted to analyse which residues are allowed at this Φ 1 position from the biological perspective for successful export by CRM1. We only focused on issues pertaining to the export signal in these experiments and not on the influence on the translation function of eIF2 β . As shown in **figure 18**, this position was replaced by bulky, hydrophobic, polar, charged and rigid amino acid and tested for interaction with CRM1. From this experiment it is clear that bulky residues like Tyrosine, Phenylalanine and conformationally restricted residue like Proline, cannot be tolerated, presumably because the steric hindrance does not allow proper entry or fitting of the NES in the rigid CRM1 binding cleft. Polar residues like Glutamine, Serine and charged residues like Lysine and Histidine are also not tolerated, presumably because they interrupt a hydrophobic interaction with a CRM1 residue.

We validated this in vivo by expressing a GFP tagged eIF2 β mutants in HeLa cells to see if the mutated NES can get recognised by CRM1 and get exported (**figure 18**). The transfections further supported the result of the binding assays, suggesting that the Φ 1 position in the eIF2 β plays a crucial role in modulating the affinity of the NES for CRM1 and can only occupy amino acids of a hydrophobic and aliphatic side chain. As a control, we also tested each sample with LMB to ensure that the export is solely CRM1 dependent (figure 18 C).



Φ		0			1				2			3		4	
	D	I	D	E	A	E	E	G	V	K	D	L	K	I	E
					F										
					Y										
					I										
					V										
					S										
					Q										
					K										
					H										
					P										
					R										
					L										

null
Very weak
weaker
wildtype
stronger
Very strong

Fig 18 A) Ala Φ 1 hydrophobic residue was mutated to different amino acids belonging to the hydrophobic, polar, charged and bulky groups. The zz tagged mutants were immobilized and pulldowns were performed using 1 μ M CRM1 and 1 μ M CRM1 + 3 μ M RanGTP in bacterial lysate at 4°C for 3 hours (internal specificity control). The eluent was loaded per sample and analysed by SDS page. The affinity of the different NES mutants for CRM1 binding is also depicted in the table.

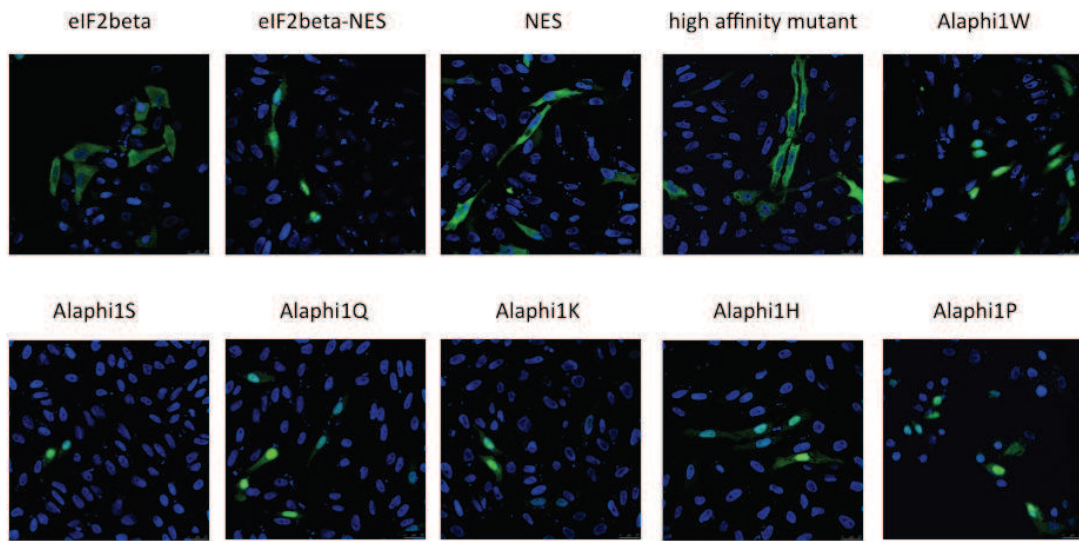
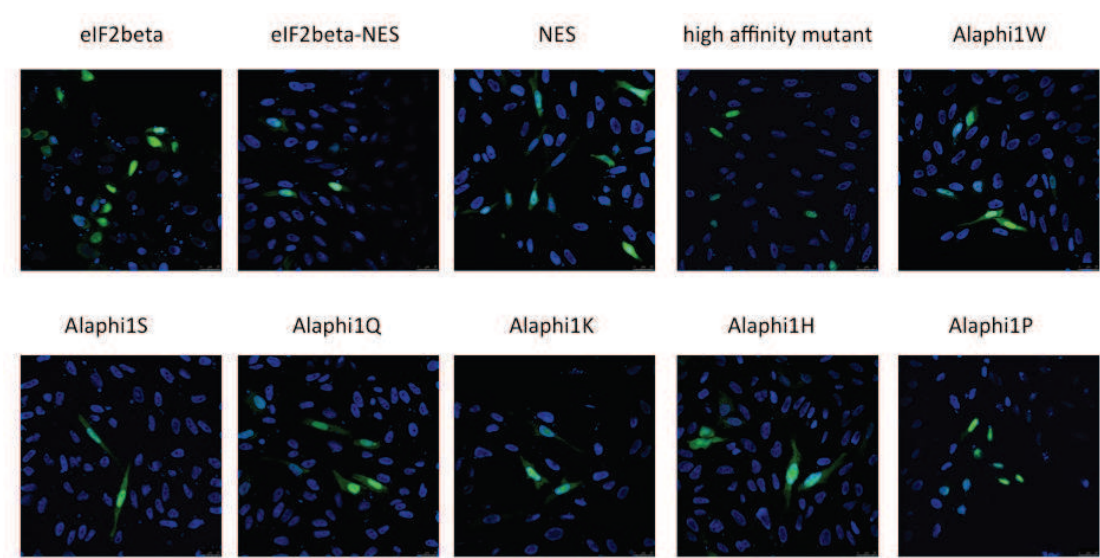


Fig 18 B) HeLa cells were transfected with a plasmid (1 μ g) expressing GFP tagged eIF2 β constructs. The transfecting agent used was Fugene 6 (1.8 μ l per ml of media) and the cells were transfected for 24 hours. The cells were fixed after 24 hours of transfection and all these steps were executed at 37°C. The cells were fixed using PBS and 4%PFA and the nuclear DNA was stained using Hoechst solution. High affinity mutant - A Φ 1 L, eIF2 β -NES means this construct does not consist of the NES, NES is the Nuclear Export Signal. **Fig 18 C)** HeLa cells were transfected with the same procedure as described in 18B. Here, the samples were treated with Leptomycin B for 3 hours at 37°C before fixation.

Figure 18 C - Transfection tested with Leptomycin B.



3.9 Analysis of the spacer residues in the eIF2 β consensus

So far the emphasis has been on understanding more about the hydrophobic Φ residues, when it comes to defining the NES recognised by CRM1 to build the NES prediction algorithms. Based on literature, it is evident that the spacer residues surrounding the Φ positions are also crucial in the NES binding and influencing the affinity to CRM 1. In the NES of the proteins An3 and PKI, when the spacer residues were made more acidic, the affinity to CRM1 to form an export complex increased (Güttler et al., 2010). In our opinion, the NES definition is incomplete and we want to highlight the role of the spacer residues in more detail. Thus, we want to explore the contribution of the spacer residues in interaction of the classical NES with CRM1 and RanGTP by using the eIF2 β NES as a model. This information will be used to refine the NES consensus.

According to our results, the eIF2 β follows the classical NES consensus:

$x\Phi^0xx \Phi^1 xxx\Phi^2xx$ or $xxx\Phi^3x\Phi^4$

In the forthcoming experiments, we wanted to develop a consensus of the residues, which are permitted at these spacer positions. We started by analysing the NES from a structural perspective. Here we also included the residue immediately adjacent to the last Φ position because it may have an effect on the NES owing to its close proximity. The NES consensus we took into account was:

$x\Phi^0xx \Phi^1 xxx\Phi^2xx$ or $xxx\Phi^3x\Phi^4x$

We mutated the spacer residues successively to Prolines and Glycines and tested the export forming capacity of these mutants by performing binding assays. We chose Prolines and Glycines, because they have poor helix forming propensities. Prolines can break a helix, because of steric hindrance and inability to form a hydrogen bond. Glycine tends to break helices as well, because of the high

flexibility it contributes in the protein structure. This makes it entropically expensive to adopt the relatively constrained alpha helical structure.

The binding assays performed will give us a fairly good idea about the differences in binding capacity of the various eIF2 β mutants to CRM1 during formation of an export complex. However, an in vivo validation is essential to correlate our finding in a biological system. To do this, GFP tagged eIF2 β spacer mutants were expressed in HeLa cells by transfection and we wanted to see the intracellular distribution of the mutants.

This analysis has been depicted in **figure 19 A**. In the binding assays it was seen that introduction of Prolines in many spacer positions decreases the interaction with CRM1 in the presence of RanGTP. This was also expected since it is a rigid residue, which would not allow the NES to snake its way through the rigid and rather narrow CRM 1 binding cleft. The binding affinities have been shown in a colour code, where light purple depicts the affinity corresponding to the wild type constructs. As the colour gets a deeper shade of green, the affinity goes up and shades of red depict a decrease in the affinity of the mutant to CRM1. Black shows a complete abolishment in binding. Maximum decrease in binding was seen in positions G96, K98, D99 and K101 - residues surrounding hydrophobic Φ residue 2 and 3. E92, E94 and E95 were also extremely sensitive to the mutation and showed negligible binding to CRM1 (+RanGTP). The dramatic decrease in affinity for CRM1 in these spacer mutants, shows that a secondary structure is being formed. By introducing Prolines, these structures are being interrupted and this results in loss of binding. For instance, in SNP1 there is a fold between hydrophobic Φ residues 0 and 1, 1 and 2, 2 and 3 - any of these could be interrupted by the presence of Prolines. It is surprising that Prolines are also not tolerated at non helical regions like spacer positions adjacent to Φ 4 - K101P and E103P also show negligible interaction with CRM1+RanGTP. On the other hand, when Proline is introduced at spacer positions surrounding Φ 0 (D89P and D91P) the mutants showed a weaker interaction with CRM1 and RanGTP compared to the wildtype NES. Thus positions surrounding Φ 0 have a greater tolerance for Prolines compared to the other spacer positions - this could happen

because around this position the NES exits the binding cleft and can make room to accommodate Proline. But, deep inside the narrow hydrophobic cleft, Proline can not be placed easily. Moreover, at the edge of the NES near $\Phi 0$, it is probable that no major secondary structures are being formed to allow correct fitting of the NES in the CRM1 cleft. Thus, Proline at this spacer position may not have the possibility to disrupt any important structural conformation required for the NES binding to CRM1.

Figure 19A

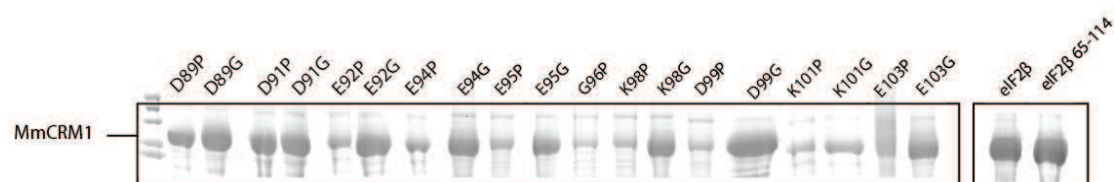
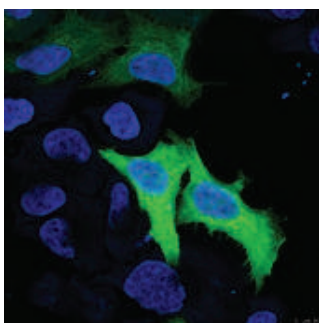


Figure 19-A

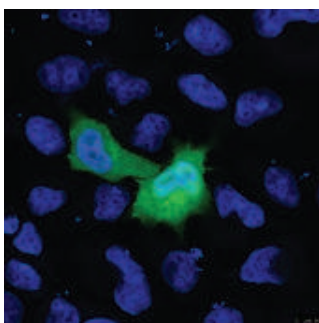
The eIF2 β -NES spacer residues were analyzed by mutagenesis. The zz tagged mutants were immobilized and pulldowns were performed using 1 μ M CRM1 and 1 μ M CRM1 + 3 μ M RanGTP in bacterial lysate at 4°C for 3 hours (lysate is the internal specificity control). The eluent was loaded per sample and analysed by SDS.

The next step was to analyse the eIF2 β mutants where the GFP tagged Proline mutants were expressed in HeLa cells by transfection. We can immediately see nuclear localisation of most of the mutants. This supports the results from the binding assays and is also plausible since Proline is restricted in conformation and it is capable of causing hindrance when the NES tries to fit in the rigid CRM1 binding cleft.

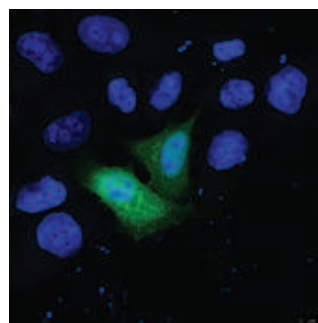
CONTROL



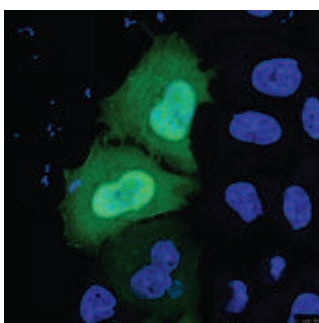
D89P



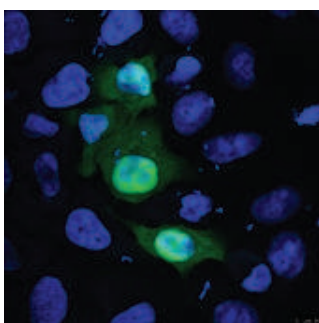
D91P



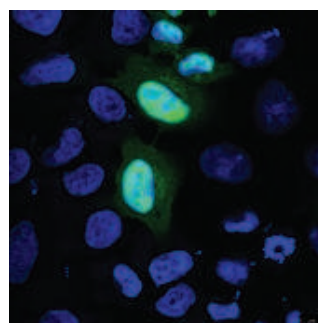
E92P



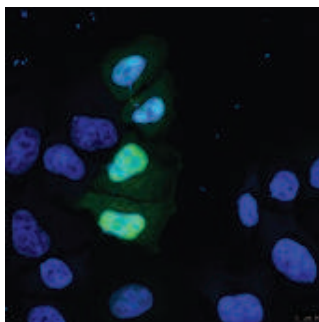
E94P



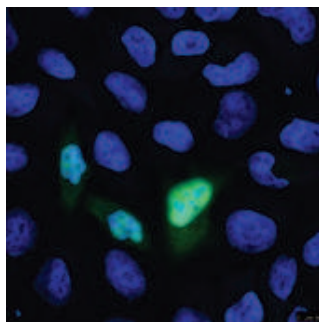
E95P



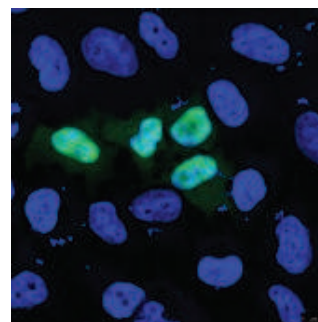
G96P



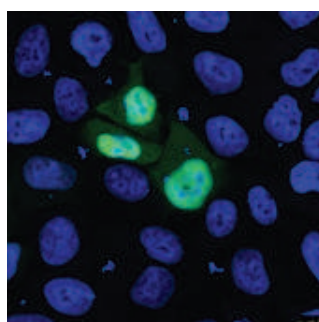
K98P



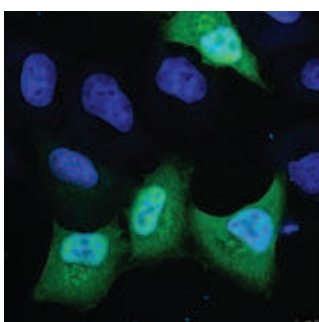
D99P



K101P



E103P



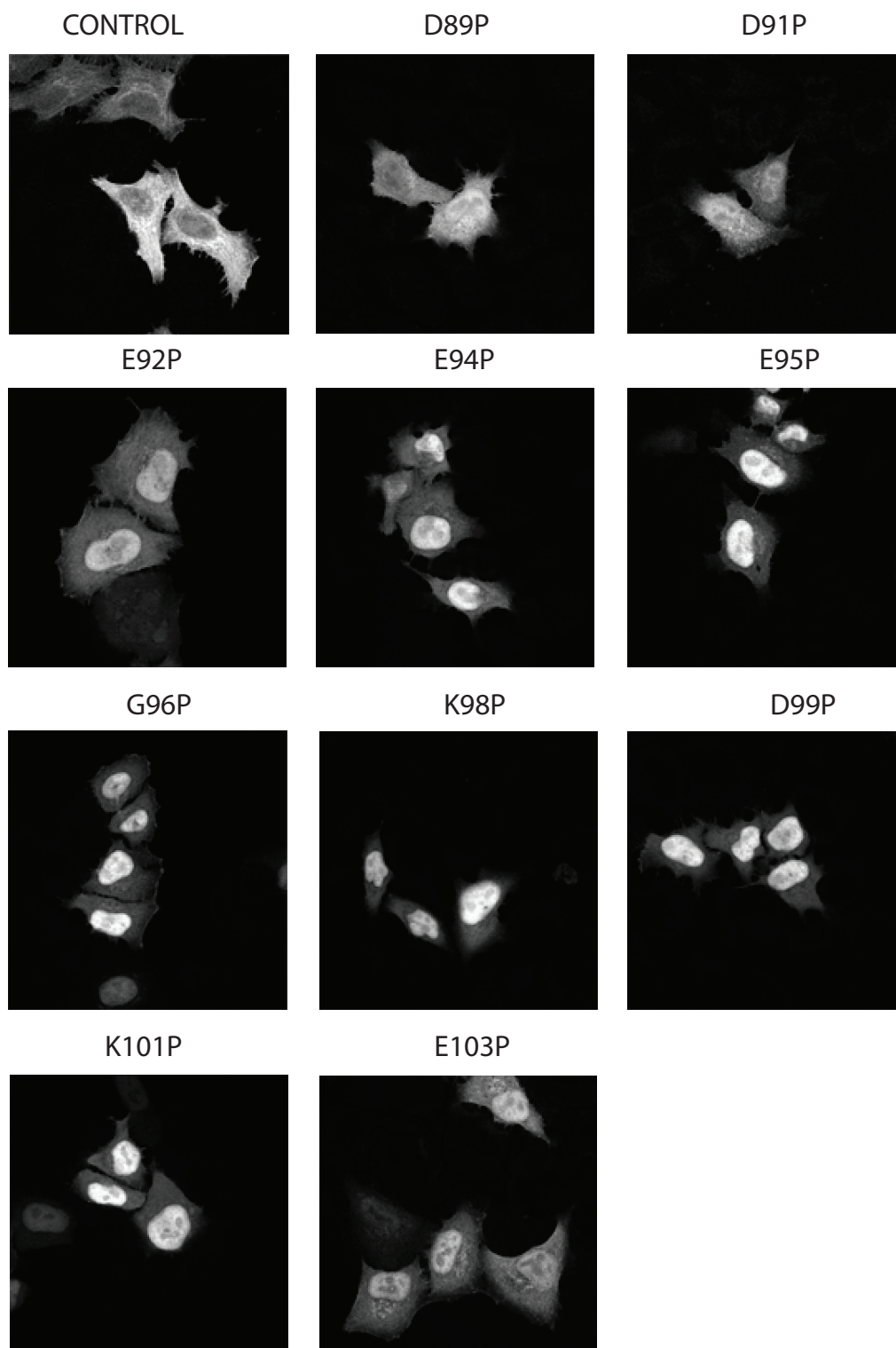


Figure 19 B- Proline mutants of the NES spacer residues.

GFP tagged Proline mutants of the eIF2 β NES spacer residues were transfected into HeLa cells. HeLa cells were transfected with a plasmid (1 μ g) expressing GFP tagged eIF2 β constructs. The transfecting agent used was Fugene 6 (1.8 μ l per ml of media) and the cells were transfected for 24 hours. The cells were fixed after 24 hours of transfection and all these steps were executed at

37°C. The cells were fixed using PBS and 4%PFA and the nuclear DNA was stained using Hoechst solution. Below the overlay, the same transfections have been shown in a grey scale. The GFP tagged protein is shown in grey colour to show the intracellular distribution clearly.

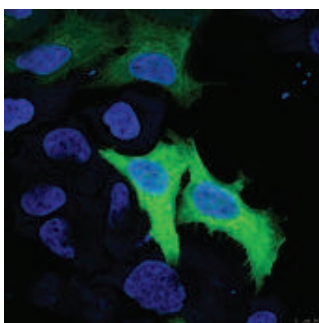
The Proline mutants of the spacer residues D89 and D91, that are adjacent to Φ 0, show very weak export as compared to the wildtype NES. A weak interaction of these mutants is also seen in the binding assays. Proline mutants of E92, E94, E95, G96, K98, D99 and K101 - spacer residues surrounding Φ 1, Φ 2, Φ 3 and Φ 4 show strong nuclear localisation and abolishment of nuclear export. The mutant of spacer residue adjacent to Φ 4 - E103 also shows strong nuclear localisation, which also supports the negligible interaction as seen in the binding assays. This emphasises that this residue also plays an important role in the interaction and fitting of the NES in the CRM1 binding cleft (**Fig 19 B**).

The introduction of Glycines at the spacer positions also decreased the binding to CRM1, but this was evident on specific positions like E95, K98 and K101-residues surrounding the second and third Φ position (**figure 19A**). The Glycines are better tolerated than the Prolines, because they are more flexible and pose less steric hindrance in the NES fitting in the binding cleft. As seen in other NES structures, the spacer region between Φ 2 and Φ 3 positions forms an alpha helix to bring the hydrophobic residues in the correct position in the hydrophobic cleft. In the process of positioning the Φ residues, these spacer residues form an alpha helix like structure. Glycine probably disfavours these structures, because of its highly flexible nature and hence decreases the ability of the NES to fit into the binding cleft, consequently damaging the binding capacity to CRM1. After the spacer position of K98 (following Φ 2 position) a secondary structure is often seen e.g. HIV-Rev NES, SNP1 -NES and this can also get interrupted by Glycines.

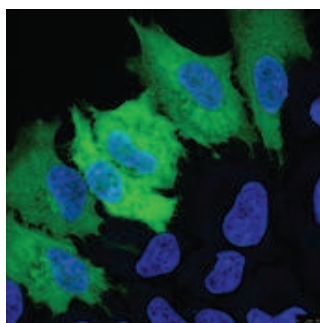
Following this, the Glycine spacer residues were transfected into HeLa cells and their localisation varied depending on the location of the mutation. Glycine mutants of D89 and D91, surrounding Φ 0, showed cytosolic localisation that was almost as effective as the wildtype NES. The Glycine mutants of spacer residues E92, E94 and E95 showed distribution in nucleus and cytosol, hence a definite weakening of the export signal. The mutant of K98G, adjacent to Φ 2 also

showed nuclear and cytosolic localisation. The spacer residues surrounding L Φ 3 showed interesting behaviour. D99G showed cytosolic localisation (weaker than wildtype NES) but K101G showed nuclear localisation. The mutant of spacer residue E103, adjacent to I Φ 4, also tolerated a Glycine and showed cytosolic localisation similar to the wildtype NES. Thus, spacer residues surrounding A Φ 1, V Φ 2 and L Φ 3 are sensitive to the introduction of Glycines. This also validates the results from the binding assays that indicate less tolerance of Glycine at spacer positions E95G, K98G and K101G. Possibly, Glycine mutants are disrupting a secondary structure being formed between Φ 1 and Φ 2; and Φ 2 and Φ 3. This mutant also influences the fitting of the I Φ 4. In general, Glycine mutants of spacer residues are better tolerated than Proline mutants and show position specific effects (**Fig 20A**).

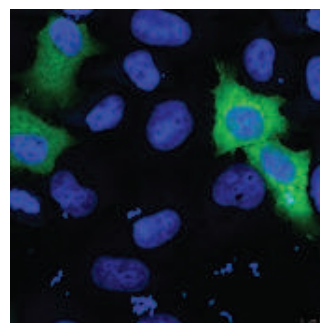
CONTROL



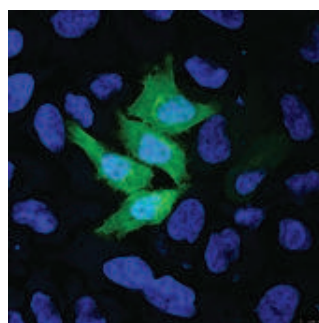
D89G



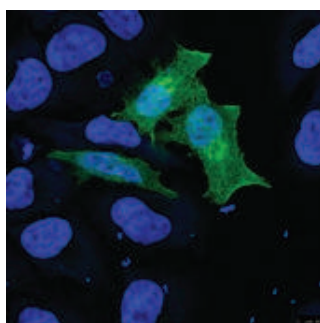
D91G



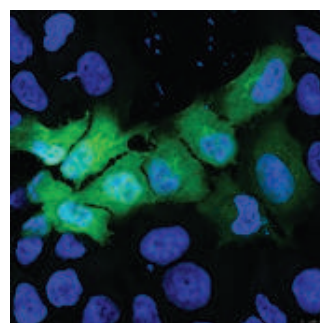
E92G



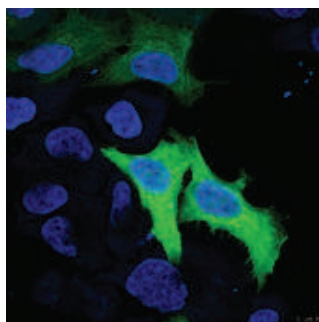
E94G



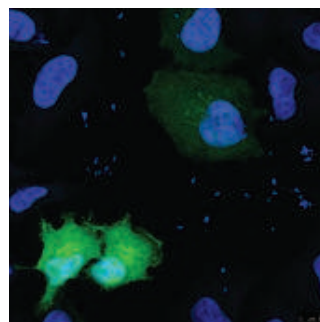
E95G



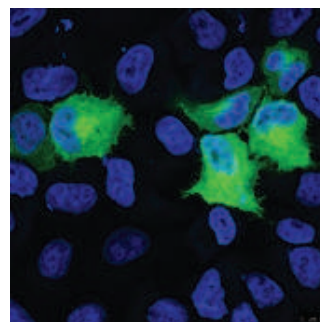
G96



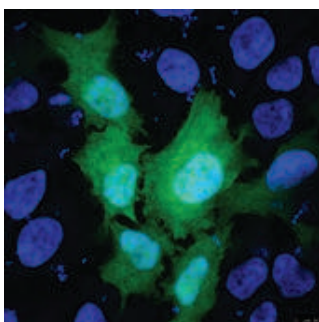
K98G



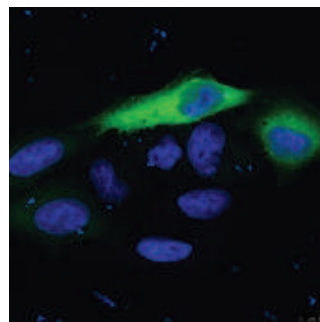
D99G



K101G



E103G



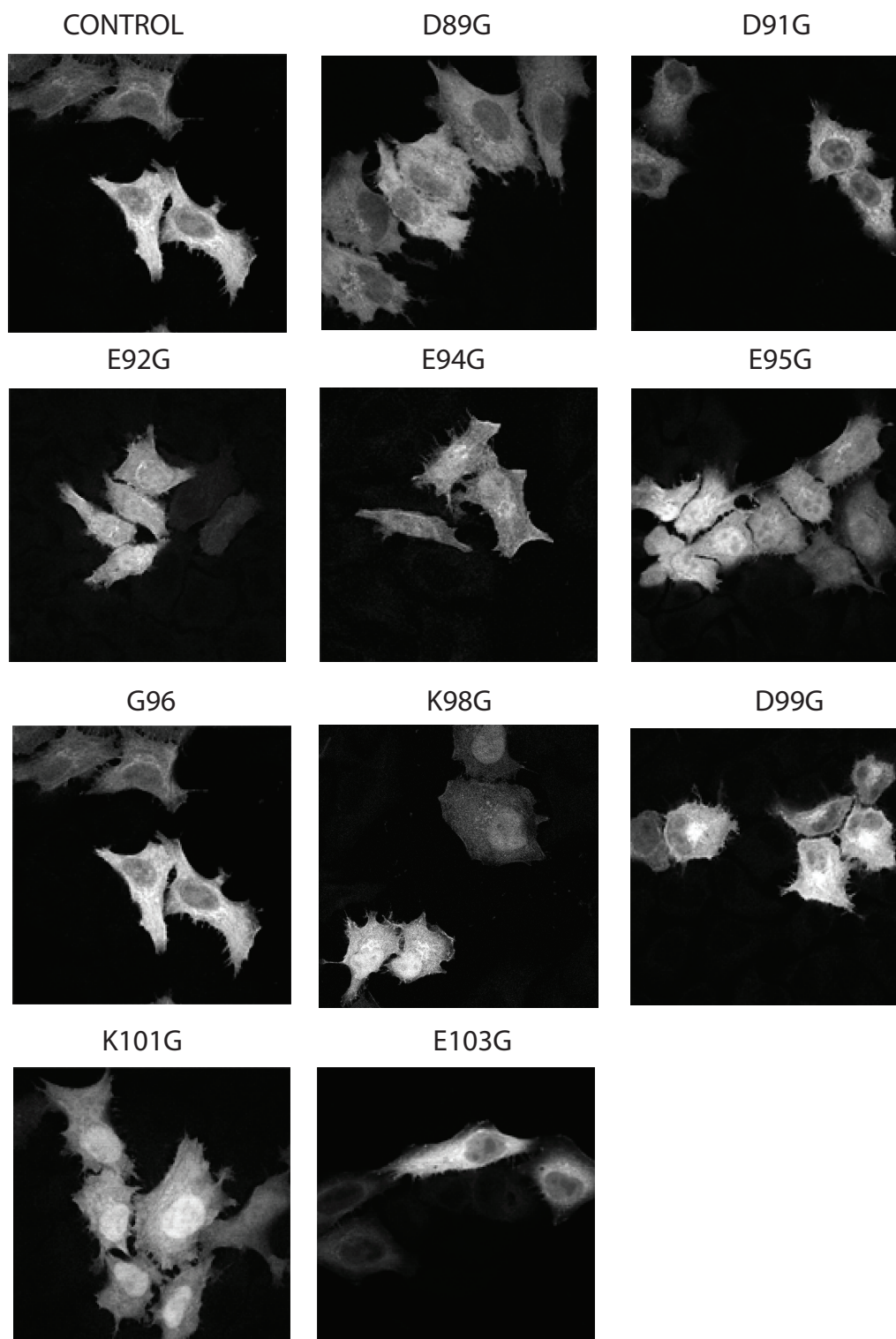


Figure 20A - Glycine mutants of the NES spacer residues.

GFP tagged mutants of the eIF2 β NES were transfected into HeLa cells. The spacer residues of the NES were successively mutated to Glycines. The procedure is the same as in Figure 19B. Below the overlay, the same transfections have been shown in a grey scale. The GFP tagged protein is shown in grey colour to show the intracellular distribution clearly.

In addition to the spacer residues defined by the NES consensus, the residue immediately adjacent to the last Φ position was also included in the analysis. It was clear from the binding assays, that this position is very sensitive to both Proline and Glycine and plays a crucial role in the interaction of the eIF2 β NES to CRM1 and its successful nuclear export. Hence, this residue should be considered in the existing NES consensus as well.

In the next step, the aim was to understand the role of the amino acid side chain in the spacer position. These residues were successively mutated to bulky, hydrophobic and polar residues. These mutants were then analysed for interaction with CRM1 by binding assays and in vivo by transfections.

When Tyrosine was introduced at the spacer positions, the binding to CRM1 went down dramatically. This was not a surprise, since the presence of a bulky residue probably blocks the entry of the NES in the CRM 1 binding cleft. This effect was more pronounced around Φ positions 1 and 2. Regions in the spacer positions occupied by E92, E94, E95, G96 and K98 often form secondary structures, which are easily disfavoured by bulky residues **Figure 21A**. Weak interactions were seen when spacer residues D99, K101 and E103 were mutated to Tyrosine. Maybe the hydrophobic cleft is not as narrow at these positions and has greater capacity to tolerate bulky residues compared to spacer positions surrounding Φ 1 and Φ 2. D91Y, next to Φ 0 also shows a weak binding to CRM1+RanGTP.

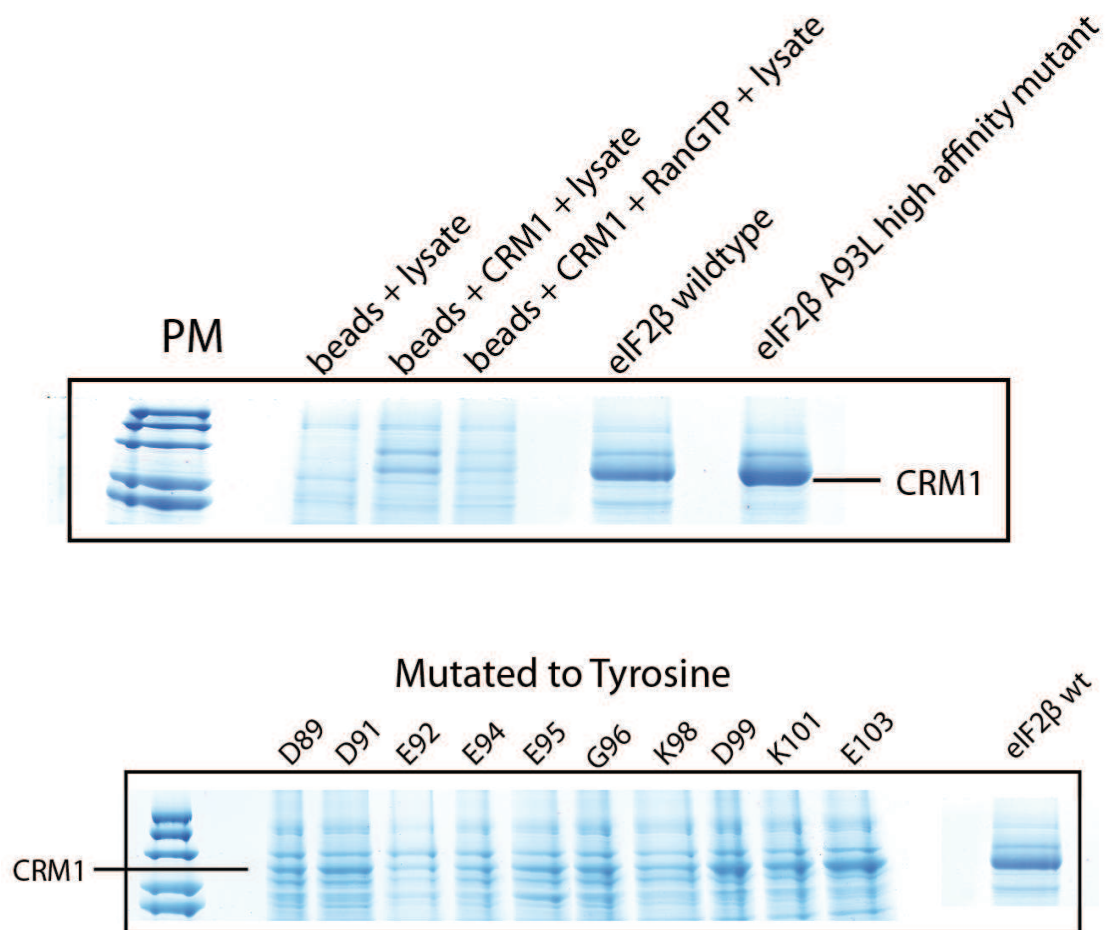
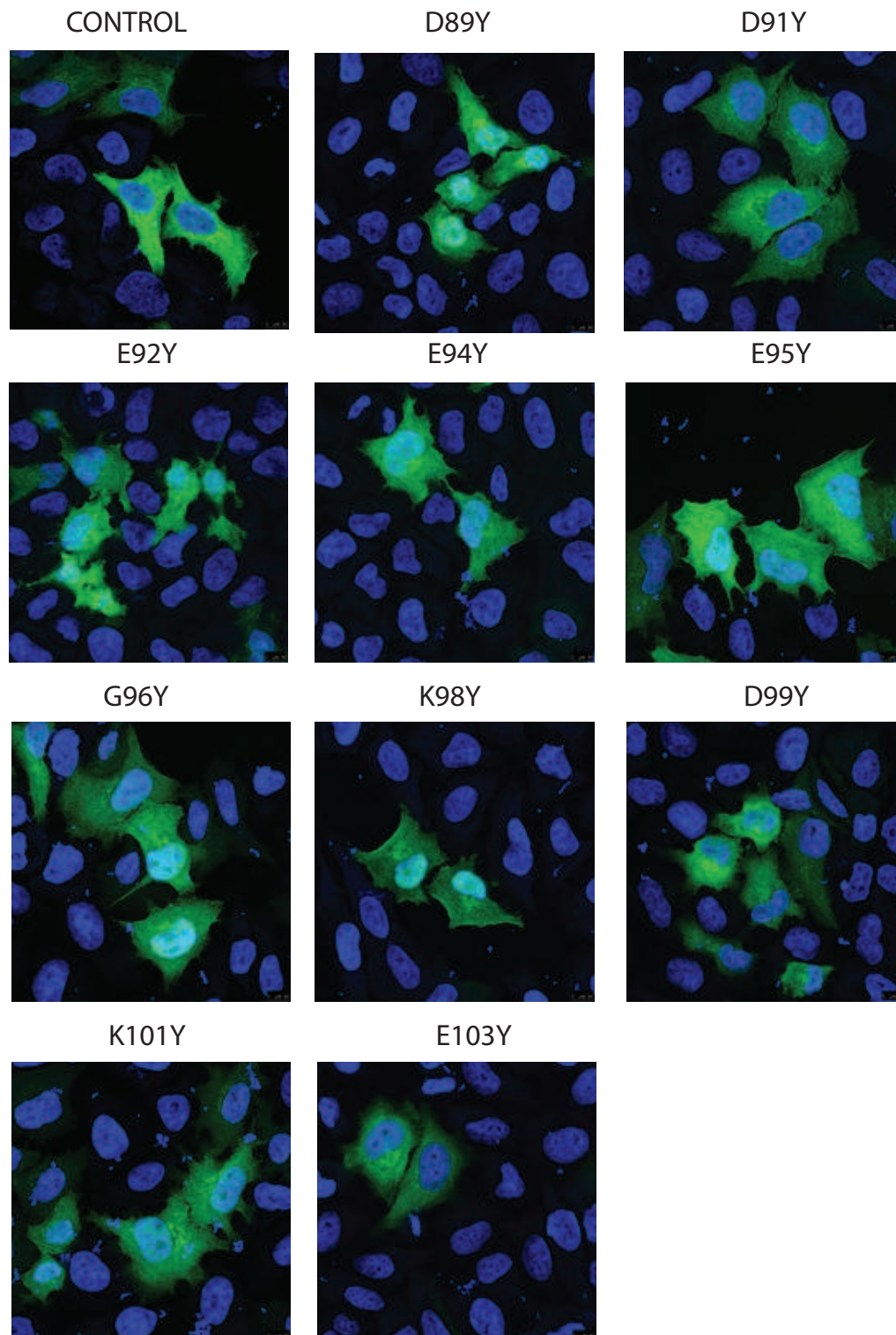


Figure 21A

The procedure is same as in 19A.

The spacer mutants changed to Tyrosine were expressed in HeLa cells with a GFP tag (Fig 21B). Tyrosine mutant of D89, spacer residue adjacent to IΦ0 shows nuclear localisation. D91Y shows a weakening of the export signal and the mutant is distributed in the nucleus and cytosol. E92Y, E94Y and E95Y also show distribution throughout the cell nucleus and cytosol. G96Y and K98Y show more nuclear localisation and emphasises the intolerance of bulky residues around Φ positions 1 and 2. K101Y also shows nuclear and cytosolic localisation. D99Y and E103Y show weak cytoplasmic distribution. Thus, spacer residues surrounding Φ3 and Φ4 are not as sensitive to the introduction of Tyrosines as Φ1 and Φ2. In summary, the bulky residues are not tolerated at spacer positions surrounding Φ0, Φ 1 and Φ 2. This maybe because a bulky residue is difficult to

accommodate in the narrow binding cleft, even at the edge of the NES at spacer position D89. Tyrosines may also disrupt important interactions with the CRM1 backbone in the binding cleft because of the difference in ionic properties compared to the original spacer residues.



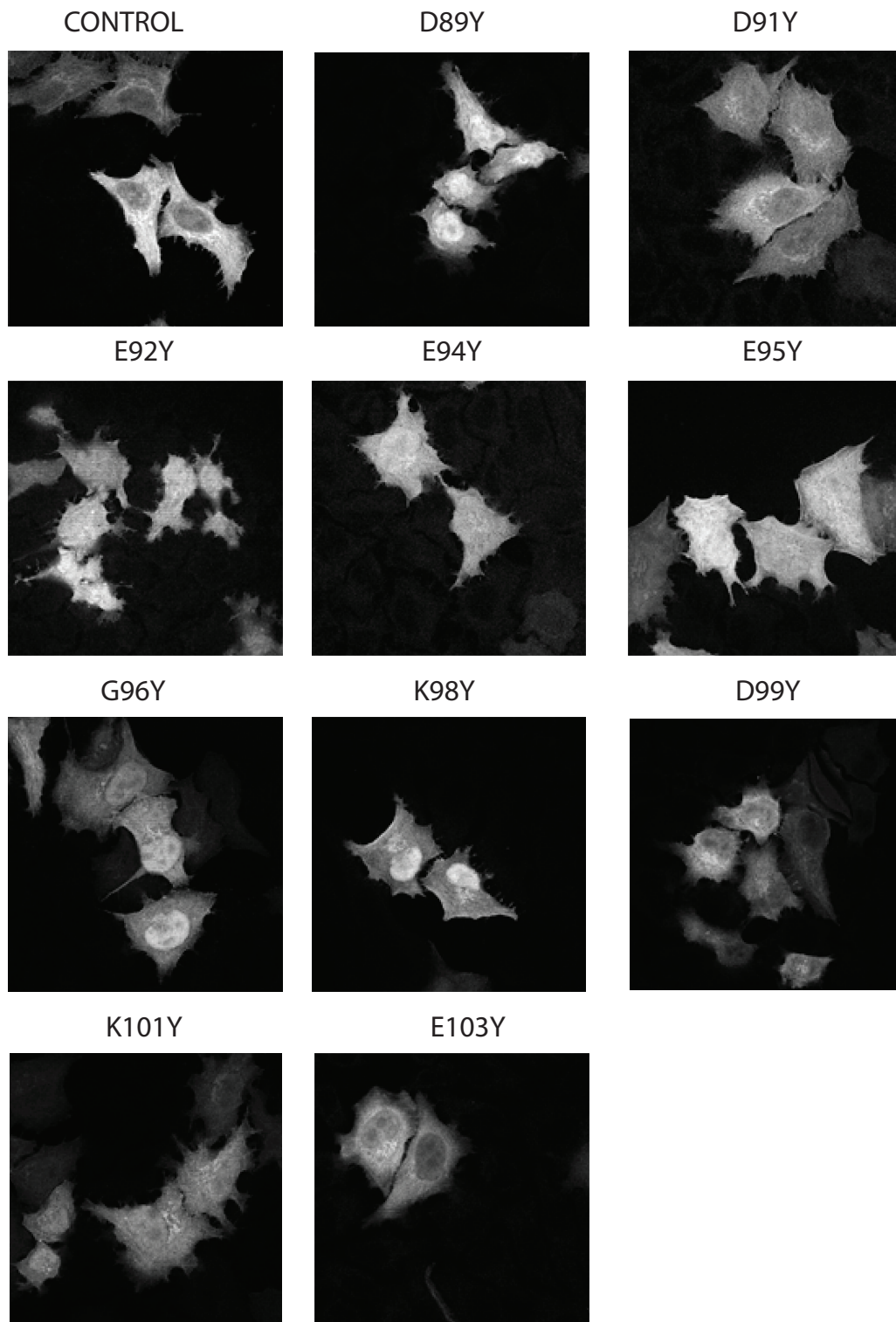


Figure 21B - Tyrosine mutants of the NES spacer residues.

GFP tagged mutants of the eIF2 β NES were transfected into HeLa cells as done in the previous assays. The spacer residues of the NES were successively mutated to Tyrosines in this case. The

procedure is the same as in Figure 19B. Below the overlay, the same transfections have been shown in a grey scale. The GFP tagged protein is shown in grey colour to show the intracellular distribution clearly.

Introduction of hydrophobic Alanine also decreased the binding to CRM1 at specific positions. G96A in particular is sensitive to the introduction of Alanine resulting in a decrease in binding to CRM1+RanGTP. This could be because of the hydrophobic nature of Alanine due to which it did not interact with the CRM1 backbone and may even tend to pull away from a charged/polar surface in the CRM1 binding cleft. This causes a decrease in the capacity of the NES to fit in the binding cleft when Alanine is introduced in the spacer position (Fig 22A). (some Alanine mutants are in preparation and have not been depicted).

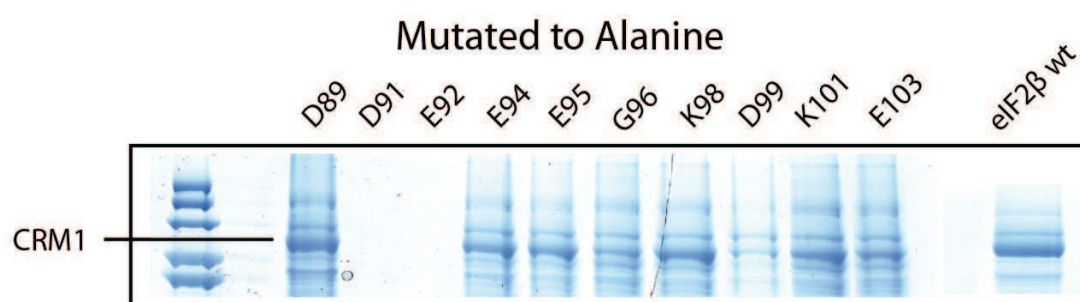
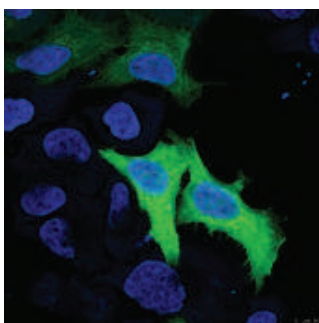


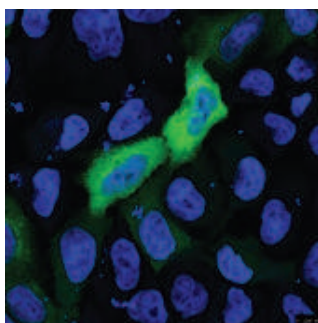
Fig 22A (same procedure as 19A)

In **figure 22B**, the spacer mutants have been mutated to Alanines and their GFP fusions have been transfected in HeLa cells. Based on the results of the binding assays, no dramatic effect was seen upon introduction of Alanines at most of the spacer positions in comparison to the wildtype CRM1 control. In the transfections, the GFP tagged mutants behaved similar to the GFP tagged wildtype NES regarding their cytoplasmic localisation in most spacer positions. However, at spacer positions G96, K101 and K98, the mutants showed a weakening in the nuclear export. Hence, spacer residues adjacent to VΦ2 and to a certain extent, LΦ3 are sensitive to the introduction of a hydrophobic residue. Maybe this mutation to Alanine interrupts an electrostatic interaction with the CRM1 backbone and results in consequently weakening of the export signal.

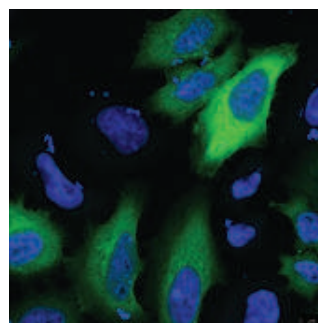
CONTROL



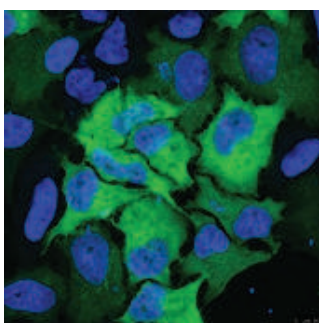
D89A



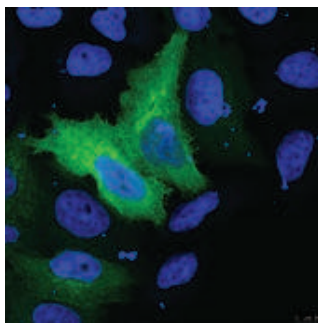
E92A



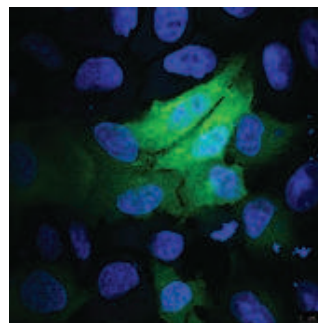
E94A



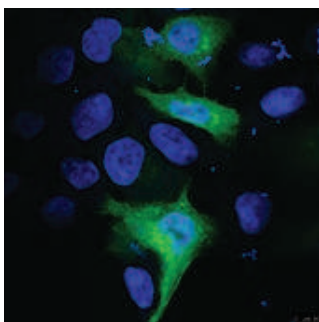
E95A



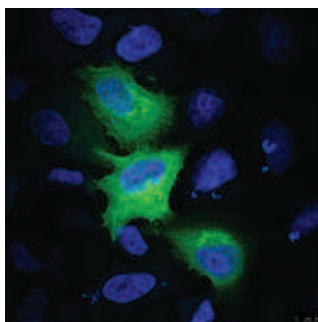
G96A



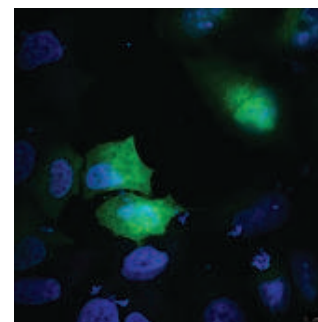
K98A



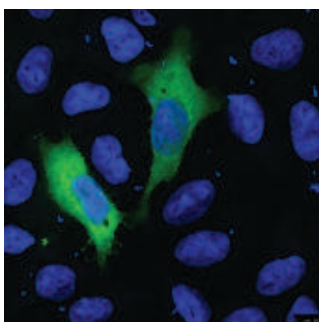
D99A



K101A



E103A



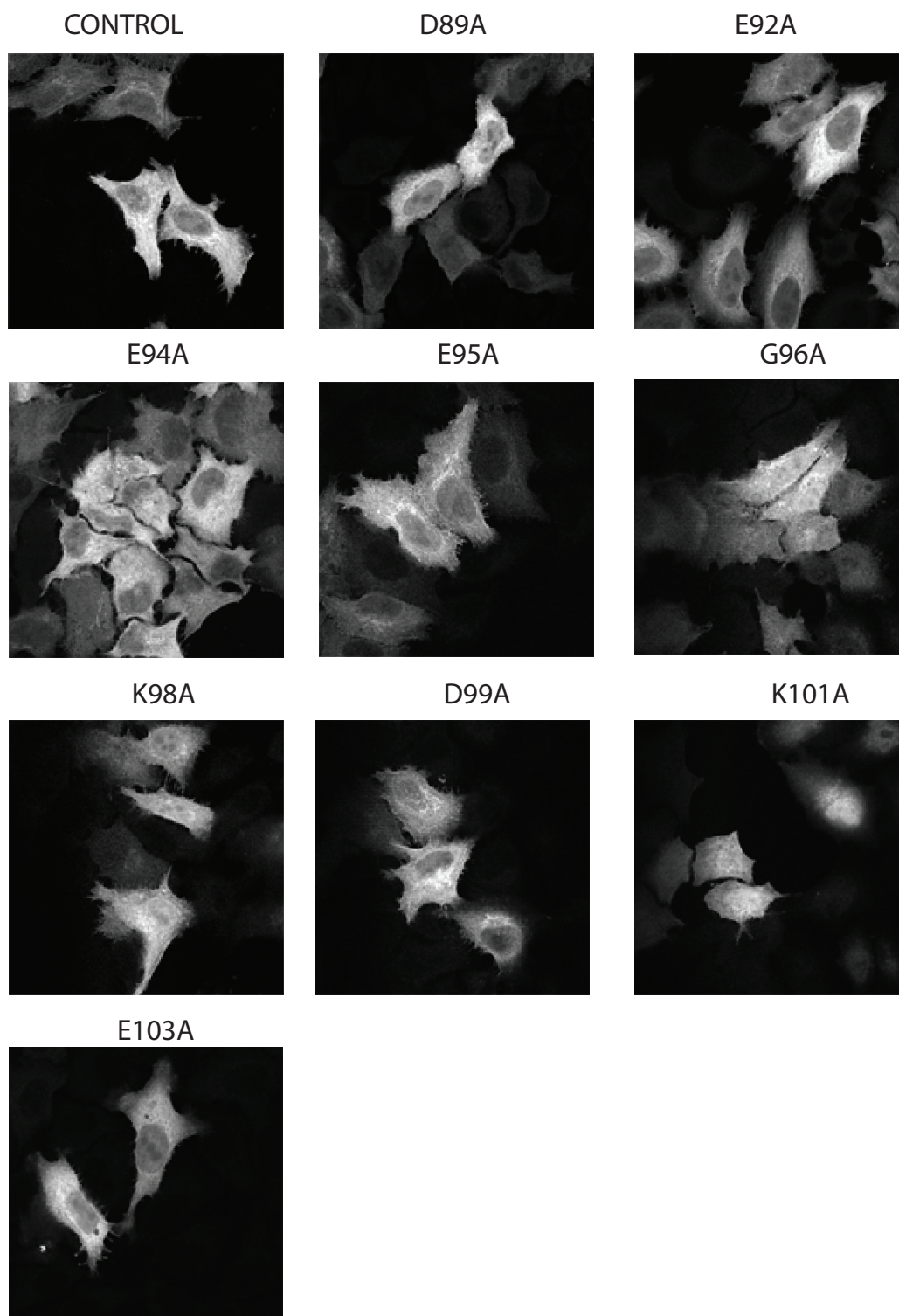


Figure 22B - Alanine mutants of the NES spacer residues.

GFP tagged mutants of the eIF2 β NES were transfected into HeLa cells as done in the previous assays. The spacer residues of the NES were successively mutated to Alanine in this case. The procedure is the same as in Figure 19B. Below the overlay, the same transfections have been

shown in a grey scale. The GFP tagged protein is shown in grey colour to show the intracellular distribution clearly.

When the polar residue Serine was present in the different spacer positions, there was a slight decrease in the binding to CRM1 at all the positions, but the effect was uniform. There was no dramatic effect at a particular position upon introduction of Serine. This is a polar residue, which can interact with the charged CRM1 backbone residues unlike hydrophobic residues. Thus, Serines are better tolerated in the spacer positions compare to Alanine (fig 23). (note: some mutants are missing in the analysis and are presently under preparation).

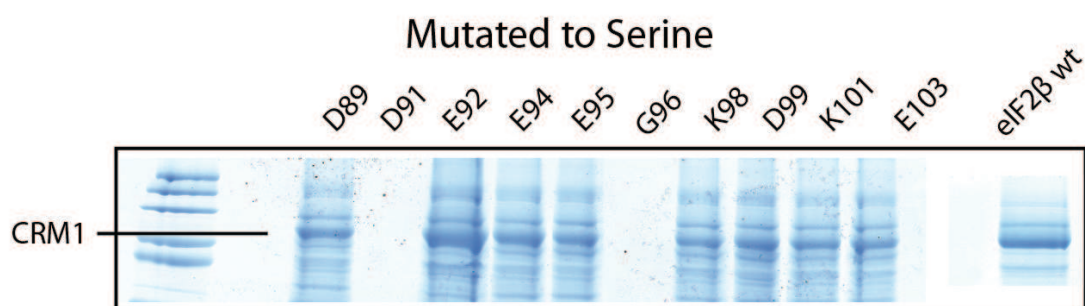


Figure 23 (same procedure as 19A)

Upon introduction of Glutamines at positions G96, K98 and K101, there was a dramatic decrease in binding to CRM1. Maybe, these mutations disturb the NES in adopting the correct three-dimensional structure and making the correct contacts with the CRM1 backbone to allow best fitting of the NES in the hydrophobic cleft (fig 24A).

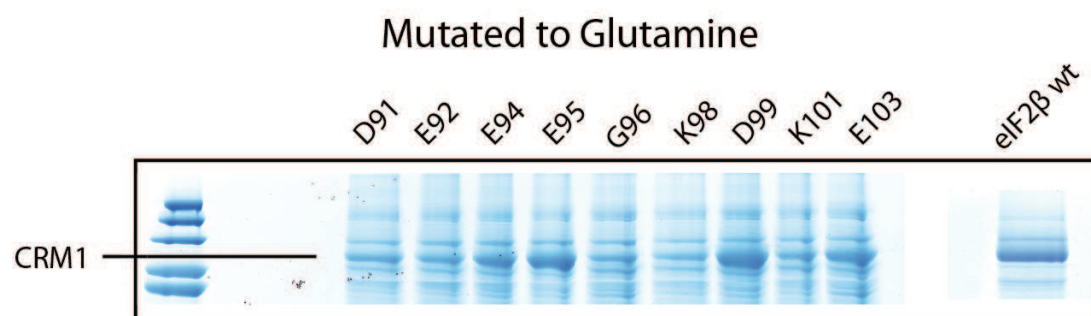


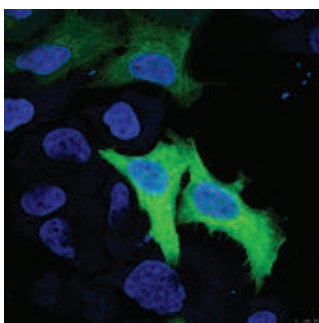
Figure 24A (same procedure as 19A)

Based on the binding assays, it was seen that the affinity of the Glutamine mutant to CRM1+RanGTP changed depending on the location of the mutation. The

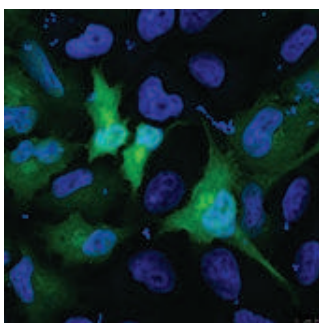
different GFP tagged Glutamine mutants of the spacer residues were transfected and the location of the mutants was analysed. The different Glutamine mutants have been shown in **figure 24B**. Glutamine mutant at spacer position D89 showed a weakening of nuclear export suggesting less tolerance of Glutamine close to $\Phi 0$. Maybe a decrease in charge at this position hinders an interaction with the CRM1 backbone. Mutants at positions D91, E92, E94 and E95 showed localisation which was more cytosolic than nuclear, but the export efficiency was lesser in comparison to the wildtype NES. Hence reduction in the charge at spacer positions surrounding A $\Phi 1$ and V $\Phi 2$ is tolerated and can allow nuclear export. Glutamine mutants at G96, K98 and K101 showed a strong nuclear localisation compare to the wildtype and hence are sensitive points for the Glutamine mutations. Thus spacer positions around $\Phi 2$ and $\Phi 3$ do not tolerate Glutamines. This suggests that an electrostatic interaction with the CRM1 binding cleft or a secondary structure is being disrupted by these mutations. In general, $\Phi 2$ is a sensitive point in the NES with low tolerance when other residues were introduced as well.

D99 and E103 can easily tolerate Glutamines and show a cytoplasmic distribution. These 2 spacer positions showed successful nuclear export when Glutamine, Glycine, Tyrosine and Alanine were introduced as well. Only Prolines were not tolerated. Thus these two spacer positions offer more flexibility to the residue being introduced.

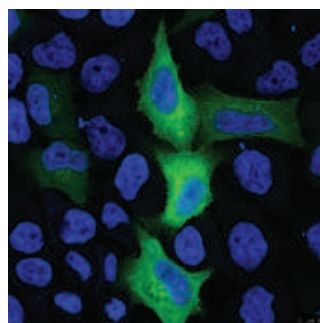
CONTROL



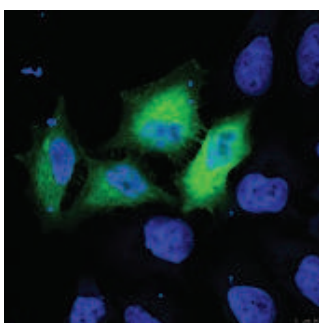
D89Q



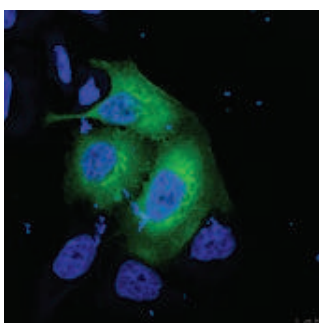
D91Q



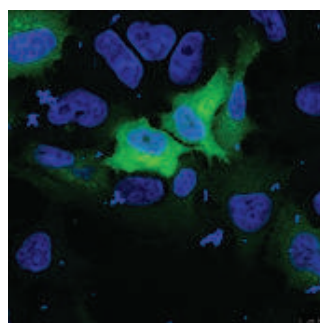
E92Q



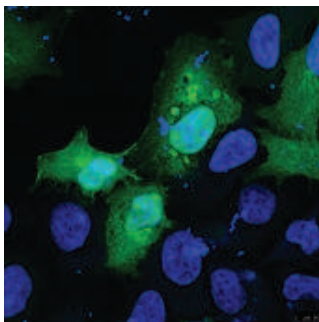
E94Q



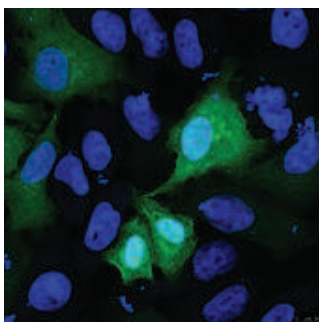
E95Q



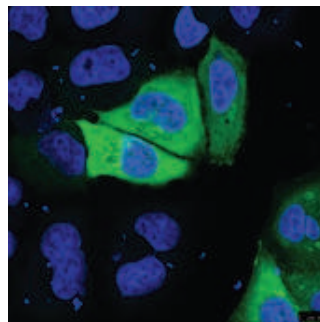
G96Q



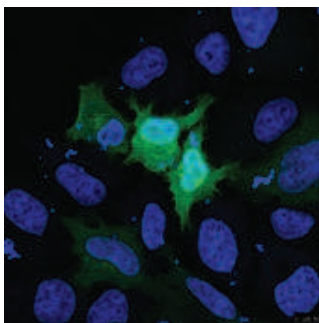
K98Q



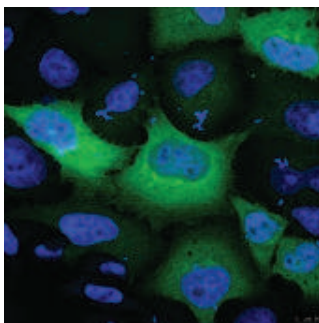
D99Q



K101Q



E103Q



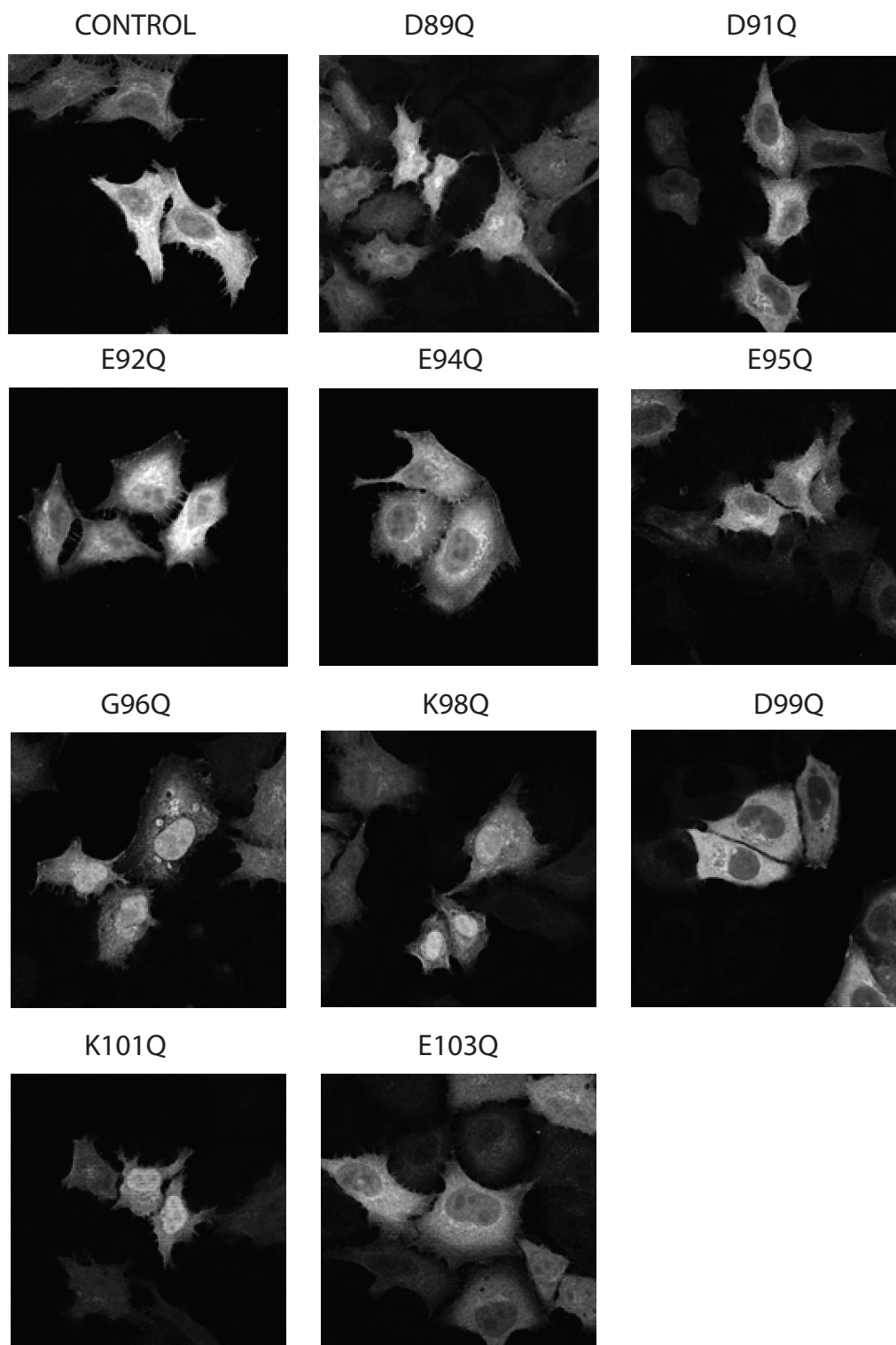


Figure 24B- Glutamine mutants of the NES spacer residues.

GFP tagged mutants of the eIF2 β NES were transfected into HeLa cells as done in the previous assays. The spacer residues of the NES were successively mutated to Glutamine in this case. The procedure is the same as in Figure 19B. Below the overlay, the same transfections have been

shown in a grey scale. The GFP tagged protein is shown in grey colour to show the intracellular distribution clearly.

Based on the binding assays, we can see that there is a fluctuation in the binding of the NES to CRM1 depending on the position and type of residue introduced. The binding assays give a rough idea about the variation in the affinity of the NES for CRM1 and RanGTP in the different mutants compared to the wildtype. In general, the spacers around Φ 1,2 and 3 cannot tolerate bulky residues. Around Φ 2 and 3 acidic residues are crucial and a decrease in charge causes a dramatic decrease in binding. When the small hydrophobic Alanine is introduced at positions around Φ 2 and Φ 3, the binding to CRM1 is abolished (summary in **figure 25**). Hence, bulky and rigid (Proline) residues are not tolerated at any position and acidic residues are preferred at the spacer positions, especially around Φ positions 2 and 3.

D99 and E103 can Glutamine, Glycine, Tyrosine and Alanine. The spacer positions around Φ 0 are also generous in their capacity to tolerate variable residues. Thus these spacer positions offer more flexibility to the types of residue that can be allowed.

Mut to	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103
	D	I	D	E	A	E	E	G	V	K	D	L	K	I	E
Y															
Q															
S															
A															
P															
G															

null
Very weak
weaker
wildtype
stronger
Very strong

Figure 25 - Summary of the interaction of the different spacer mutants with CRM1 for nuclear export. The strength is represented in a colour code shown on the left.

To really see if the spacer residue after Φ 4 position is crucial for a complete classical NES, a truncation of the 65-114 amino acids stretch can also be studied. This construct ends at Φ 4 residue and can be tested for binding to CRM1 to form an export complex. This can be used for many purposes eg. This is important especially when designing NESs which can be artificially added to a protein.

In summary, the contribution of the spacer residues from a structural point of view and questions regarding the charge, size and hydrophobicity of the spacer residues was addressed. To do this, the spacer residues were successively mutated to Tyrosine (bulky), Glutamine (polar) and Alanine (hydrophobic) amino acids.

The transfections performed have validated the in vitro binding assays to judge which residues are permitted at the spacer positions. Referring to figure 25, (summary of the affinities of the different mutants for CRM1+RanGTP) it is clear that the spacer positions have specific residues that can be permitted. These interact with the CRM1 backbone and help the NES adopt secondary structures to position the hydrophobic Φ residues at the correct places in the binding cleft.

Information regarding the residues that can be incorporated in the spacer positions of the classical Leucine rich signal will be used to develop more accurate NES prediction tools. Based on these experiments, it is also clear that the spacer residue immediately adjacent to the Φ 4 position also plays an important role in regulating the NES binding to CRM1 for export (marked in red).

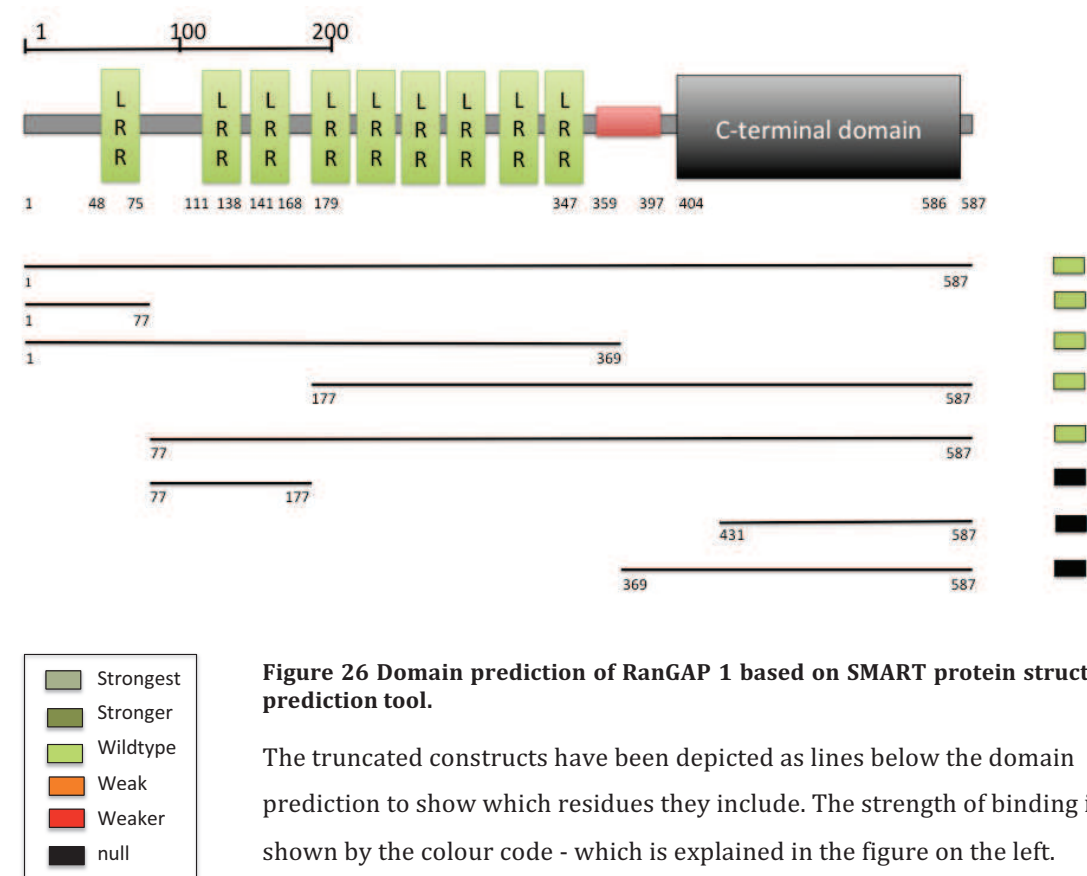
$x\Phi^0_{xx} \Phi^1_{xxx}\Phi^2_{xx}$ or $xxx\Phi^3_x\Phi^4_x$

To really see if the spacer residue after Φ 4 position is crucial for a complete classical NES, a truncation of the NES amino acids stretch can also be studied. This construct can be designed to end immediately after Φ 4 residue and can be tested for binding to CRM1 to form an export complex. An accurate NES stretch of residues can be used for many purposes e.g. This is important when designing an NES to fuse to another protein.

3.10 Hs RanGAP 1

RanGAP 1 is a crucial protein for nuclear export complex disassembly. In addition, RanGAP 1 along with RanBP1, plays a crucial role in maintaining directionality of transport across the nuclear envelope. This gradient is also maintained by RanGEF - RCC1, which is present in the nucleus.

The domain prediction of RanGAP 1 is shown in **figure 26**. As portrayed, it is a Leucine rich protein with Leucine Rich Repeats or LRR spanning the N terminal region with a flexible acidic linker that connects the C terminal domain also known as the RanGAP 1 domain. It should be noted that the RanGAP 1 domain does not constitute the catalytic activity of the protein.



As mentioned before, RanGAP 1 can enter the nucleus accidentally e.g. during cell division and it needs to be exported out of the nucleus. From previous research, it has been shown that CRM1 is responsible for its nuclear export.

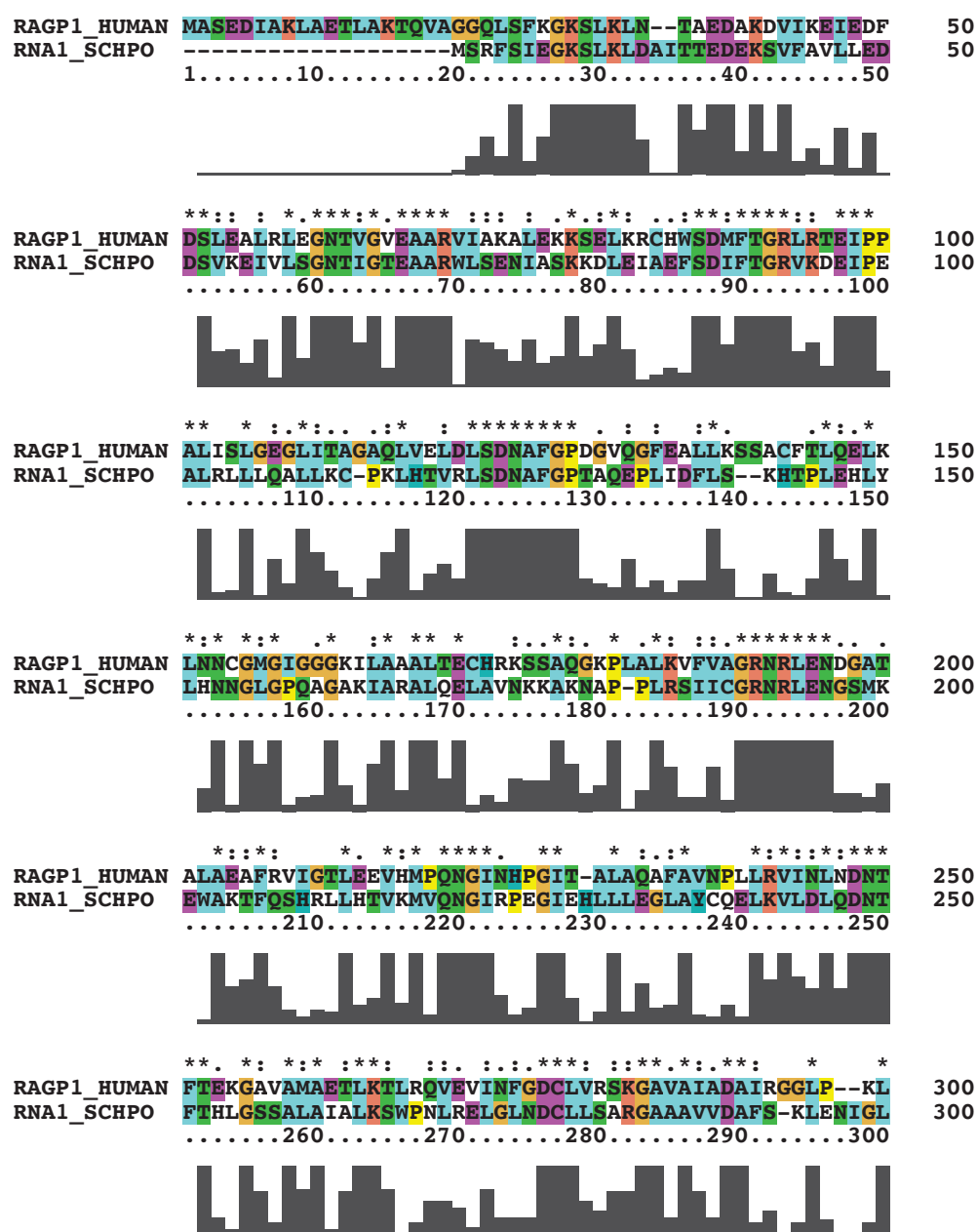
The goal of this project is to see, which residues form an export signal in RanGAP 1 for its successful recognition and nuclear export by CRM1. Following this, we will take this export complex for a structural analysis by X-ray crystallography.

3.11 NES prediction of the RanGAP 1 sequence using bioinformatic tools

When the sequence was analysed (using various bioinformatic tools) to find the classical CRM1 binding sequence there were many non specific hits. In some cases, NESs were identified in the structured hidden folds of the protein. This is a point of speculation. On one hand, it is an unlikely position for an export signal, which is usually exposed and easily accessible. The energy penalty to break a structured fold - so that the region can fit into the binding cleft of CRM1 is high and therefore this seems to be an improbable scenario. On the other hand, upon interaction with CRM1 and RanGTP, RanGAP 1 could undergo a conformational change to expose a hidden region that acts as an export signal.

The yeast homolog of RanGAP 1, Rna1p had been analysed and it was suggested that functional NLS and NES motifs were present in the protein sequence. It was suggested that Rna1p is exported by virtue of the identified NES (Feng et al., 1999). However, the crystal structure of Rna1p showed this region to be present in an alpha helical region which, in our opinion, seems to be an unlikely position for an NES that is ideally solvent exposed and flexible. Based on the crystal structure of Rna1p we wanted to make truncated constructs and test them for interaction with CRM1+RanGTP. In this way, an unbiased approach would be adopted to identifying the NES for CRM1 binding in Rna1p. Then the identified NES would be validated by mutagenesis and the mutants would be tested for Rna1p activity and interactions to ensure that the protein folding is not affected. Based on sequence conservation between Rna1p and RanGAP 1, we wanted to extrapolate our finding from Rna1p NES to identify the NES in RanGAP 1. For this strategy to succeed a high degree of conservation in the homologs is important.

When the sequence of Rna1p and Hs RanGAP 1 were compared, it was clear that the sequence conservation was very low. Rna1p is also missing the C terminal domain of RanGAP 1. Hence deducing the recognition mechanism used by CRM1 to recognise RanGAP 1, based on Rna1p and CRM1 interaction proved to be difficult. The alignment has been shown in Figure 27.



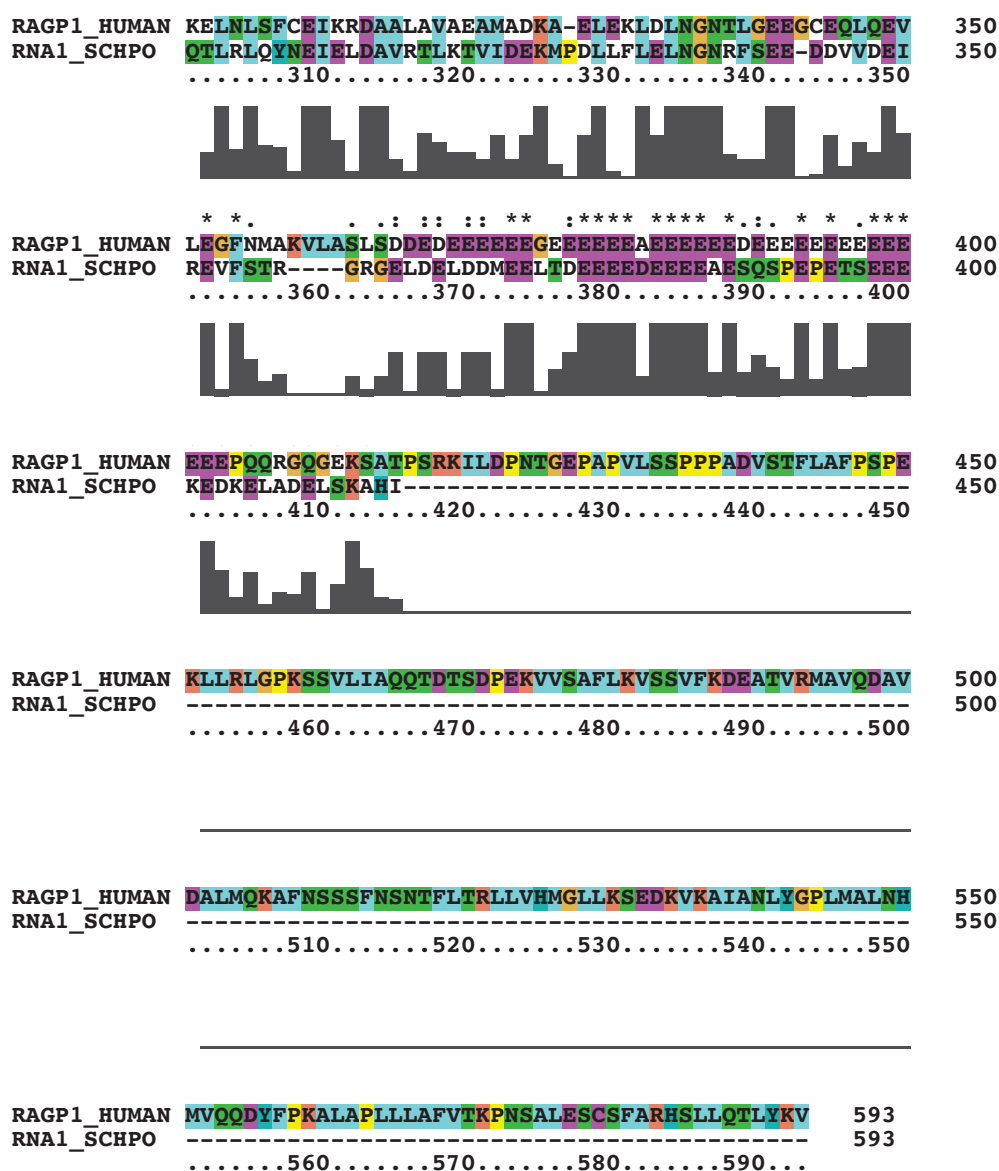


Figure 27: Alignment of Hs RanGAP 1 and Rna1p using Clustal W.

The height of the bars above the sequence is proportional to the degree of similarity of the sequences. The different colours simply depict the different amino acids.

3.12 Locating the NES by generated truncations of RanGAP 1 and testing for interaction with CRM1 and RanGTP

Since the bioinformatic approach seemed imprecise, an unbiased approach was taken to locate the export signal in RanGAP 1. Based on protein structure prediction and analysing conservation sequence patterns, truncated constructs were designed for RanGAP 1. These constructs did not interrupt conserved

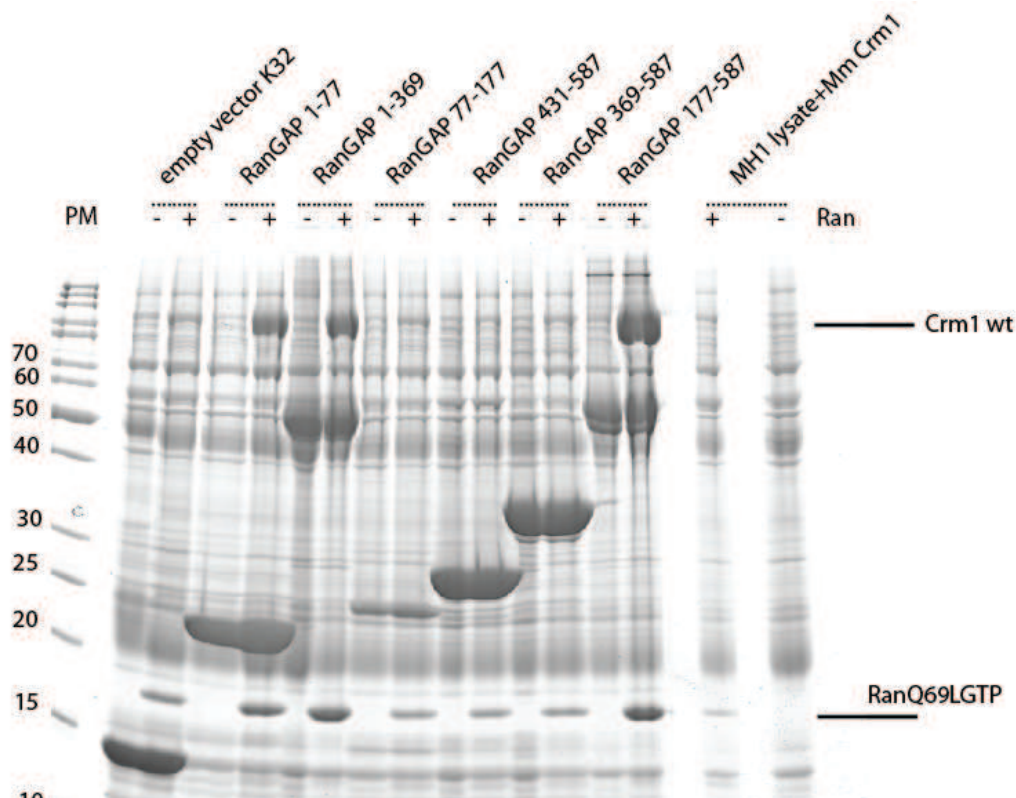
regions or structured folds and were expressed with a Histidine and a zz tag. By virtue of the Histidine tag they were purified using Nickel affinity chromatography and the zz tag was used to immobilize the constructs on IgG sepharose to perform binding assays with CRM1 and RanGTP.

The binding assays were performed in a similar manner as the eIF2 β assays. The immobilized constructs were exposed to a bacterial lysate with CRM1 (1 μ M) in one sample and in the other CRM1 (1 μ M) + RanGTP (3 μ M). After elution with SDS loading buffer, they were analysed on SDS PAGE.

The constructs designed have been shown in the figure 26 and based on the binding assays the strength of interaction of the different constructs with CRM1 in presence of RanGTP has also been analysed.

Based on the binding assays it was clear the RanGAP 1 interacts with CRM1 only in the presence of RanGTP, which is a characteristic feature of export complexes (**figure 28**).

A



B

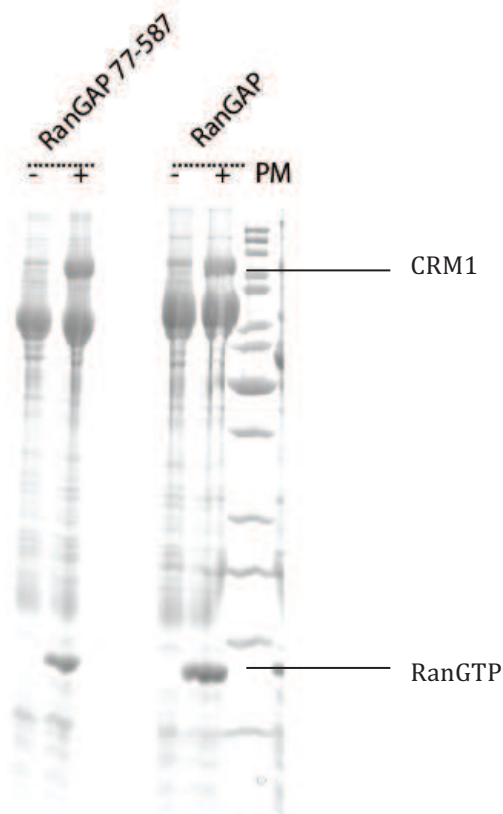


Figure 28 (A and B)

The truncated constructs of RanGAP 1 were expressed in NEB Turbo and purified via Nickel affinity chromatography. The purified proteins also had a zz tag and were used for pulldowns. The pulldown involved immobilization of the construct via the zz tag on IgG beads (10 μ l beads, 2mg/ml amount of construct to saturate the beads). One set of these immobilized constructs were incubated with a bacterial lysate (MH1 lysate) containing 1 μ M CRM1 and the other set was incubated with 1 μ M CRM1 + 3 μ M RanGTP for three hours at 4°C (500 μ l of lysate). After this, the unbound material was washed off the beads and the final elution was done in 50 μ l with SDS loading buffer without DTT. 10 μ l of this was loaded on SDS PAGE for analysis.

The binding assays of RanGAP 1 with CRM1 have been depicted in **figure 28**. The constructs spanning the LRR closest to the N terminal region spanning residues 1 to 77 aa interacts with CRM1 and RanGTP with high affinity. The same behaviour was noticed for the LRR closest to the RanGAP 1 domain. The LRR in the middle (residues 77-177 aa) did not interact with CRM1 in the absence or presence of RanGTP. The entire LRR stretch (1-369 aa) and the RanGAP 1 domain with almost the entire LRR stretch (177-587 aa) also bound CRM1 with high affinity in the presence of RanGTP.

Both versions of the RanGAP 1 domain, excluding and including the acidic linker did not show any affinity for CRM1 irrespective of the presence of RanGTP. However, when the LRR were included as shown in constructs 369-587 aa and 77-587 aa there was a strong binding with CRM1 and RanGTP.

The alignment of the binding has been shown in **figure 26** and it is clear that the LRR plays a crucial role in the interaction with CRM1 and contains the NES.

By aligning the sequences there are two probable regions that consist of the export signal, which are residues spanning 1-77 amino acids and 177-369 residues. This opens up room for questions like does RanGAP 1 have two export signals? Is there a single correct NES - if so, which one is it? Is a hidden sequence exposed upon truncating a continuous fold acting as a false positive?

Prediction of NES by re-analysis of the RanGAP 1 sequence and using previous information from the binding assays

The sequence of RanGAP 1 was re analysed (courtesy to Koray Kirli) to locate an NES that follows the PKI-NES consensus as defined in Güttler et al., 2010. The sequence of RanGAP 1 has been noted below and the highlighted regions were suggested to have an NES. The sequence in red is the C terminal structured domain and it is improbable that it has an NES because this region came with evolution (absent in yeast) and it is unlikely that the protein would use an entirely new domain for nuclear export compare to its homolog in yeast. The hit indicated in green could be a possible positive hit since it is located at the end of the LRR repeats.

Hs RanGAP 1

MASEDIAKLAETLAKTQVAGGQLSFKGKSLKLNTAEDAKDVIKEIEDFDSLEALRLEGN
 TVGVEAARVIAKALEKKSELKRCHWSDMFTGRLRTEIPPALISLGEGLITAGAQLVELD
 LSDNAFGPDGVQGFALLKSSACFTLQELKLNNCGMGIGGGKILAAALTECHRKSSAQ
 GKPLALKVVFVAGRNRLENDGATALAEAFRVIGTLEEVHMPQNGINHPGITALAQAFV
 NPLLRVINLNDNTFTEKGAVAMAETLKTTLRQVEVINFGDCLVRSKGAVAIADAIRGGL
 PKLKELNLSFCEIKRDAALAVAEAMADKAELKLDLNGNTLGEEGCEQLQEVLEGFN
 MAKVLASLSDDEDEEEEEEGEEEEEEAEDEEEEEDEEEEEEEEEEEEEEPQQRGQGEKS
 ATPSRKILDPNTGEPAPVLSSPPPADVSTFLAFPSPEKLLRLGPKSSVLIAQQTDTSDPE
 KVVSAFLKVSSVFKDEATVRMAVQDAVDALMQAFNSSFSNSNTFLTRLLVHMGLLK
 SEDKVKAIANLYGPLMALNHMVQQDYFPKALAPLLLAFVTKPNSALESCSFARHSLQ
 TLYKV

The suggested NES of the RanGAP 1 sequence have been listed below:

'FTGRLRTEIPPALI'	'AALAVAEAMADKAEL'	'QQDYFPKALAPLLL'
T L I A I	A V M A L	Q F L L L
Φ1 Φ2 Φ3 Φ4 Φ5	Φ1 Φ2 Φ3 Φ4 Φ5	Φ1 Φ2 Φ3 Φ4 Φ5
'GLITAGAQLVELDL'	'GCEQLQEVLEGFNM'	
L A L L L	C L L F M	
Φ1 Φ2 Φ3 Φ4 Φ5	Φ1 Φ2 Φ3 Φ4 Φ5	
	(most interesting hit)	
'GVQGFALLKSSACF'	'AFLKVSSVFKDEATV'	
V F L A F	F V F A V	
Φ1 Φ2 Φ3 Φ4 Φ5	Φ1 Φ2 Φ3 Φ4 Φ5	
'EAFRVIGTLEEVHM'	'NSNTFLTRLLVHMGL'	
A V L V M	S F L M L	
Φ1 Φ2 Φ3 Φ4 Φ5	Φ1 Φ2 Φ3 Φ4 Φ5	

The 1-77 region which was shown to interact with CRM1 in the presence of RanGTP in the binding assays, but in the bioinformatic search could not located it. When this region was manually inspected, the residues indicated below could possibly form an NES.

' IAKLAETLAKTQV '

I L L T V

Φ1 Φ2 Φ3 Φ4 Φ5

This is located at the beginning of the LRR and is easily accessible. So we will include this in our further analysis as well.

To validate that these are genuine NES signals, the sequence in these stretches needs to be analysed further. The predicted hydrophobic residues should be by mutated to study interaction with CRM1 and RanGTP. To ensure that the mutations have not caused misfolding, the catalytic activity of RanGAP 1 can also be tested. Another way to confirm that the conformation of the mutants is not affected, their interaction with RanBP1 and RanGTP can be tested. If the structure of the RanGAP 1 mutant is conserved, then the interaction with these proteins will not be affected in binding assays. A final validation can be performed by transfections.

It will be very interesting to see how RanGAP 1, along with RanGTP is exported by CRM1 without hydrolysing its own export complex. To allow export of RanGAP 1, there is a possibility that CRM1 binds RanGAP 1 in a way that it disables its catalytic activity or CRM1 interacts with RanGAP 1 and RanGTP in the export complex in a manner that spatially segregates the two proteins without allowing their interaction. These questions will be answered clearly by the crystal structure of this export complex.

3.13 EXPORTIN 4 and its cargo

There are some remarkable nuclear cytoplasmic receptors that can export and import cargo - hence they have a dual function of an importin and an exportin. It is interesting to understand how the presence of RanGTP can have a prominent effect on the structure of the nuclear transport receptor (NTR), such that the NTR can recognise a different set of cargo and attain a different function altogether (Görlich and Kutay, 1999).

Exportin 4 - for instance - was shown to export eIF5A and Smad 3. Later in 2009, it was shown to also import Sox2 and SRY into the nucleus (Lipowsky et al., 2000; Kurisaki et al., 2006; Gontan et al., 2009).

Solving the crystal structure of Exportin 4 bound to its export and import cargo will give deep insight into the function and mechanism of transport of this remarkable NTR. This would explain the conformational changes attained by Exportin 4 that allow it to recognise such different import and export cargo. It will also clearly explain how RanGTP influences the function of Exportin 4. By analysing the structures, we will look for a motif or a sequence that acts as a 'signal' for transport by Exportin 4 as well.

Exportin 4 was expressed in BLR bacterial strain at 0.1 mM IPTG at 18°C. It was expressed with a histidine tag and purified via Nickel affinity chromatography. It had a sumo cleavage site for the histidine tag removal. Different amounts of the purified Exportin 4 have been loaded in **figure 29**. There are minor contaminants but the protein expresses well and is purified and stable. After the tag cleavage and reverse Nickel purification, there is a decrease in the contamination and the protein continues to be stable. For high purity, it must be purified over a gel filtration and ion exchange chromatography to remove the fine contaminants.

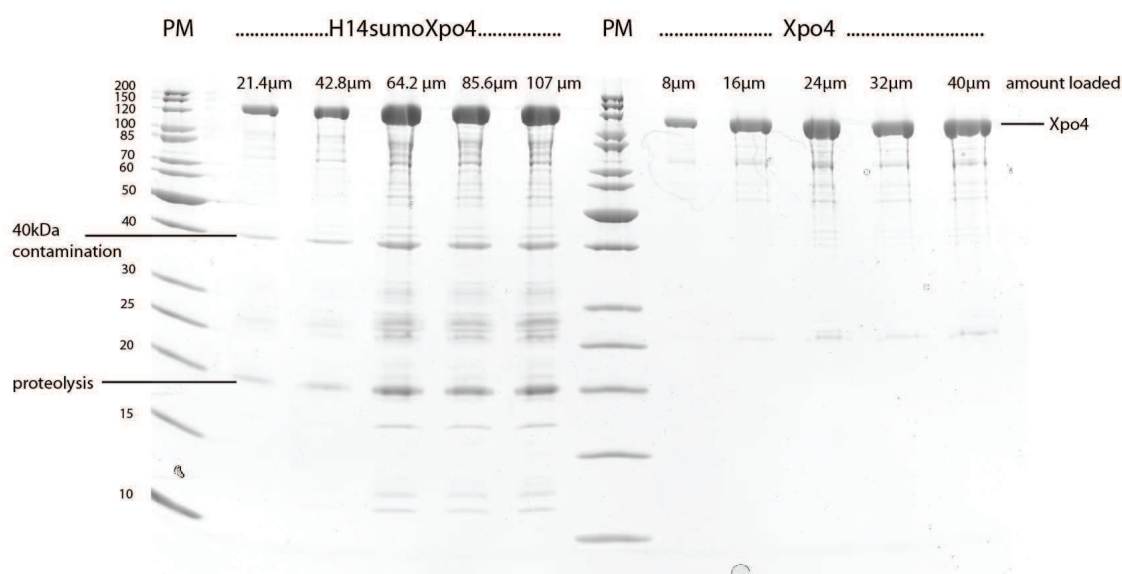


Figure 29: Purified Exportin 4

H14 sumo Xpo4 was transformed into an E.Coli strain (BLR) and this pre culture was grown over night at 37°C in 2YT media supplemented with Kanamycin. The next day this was diluted to an OD of 0.6 with media at 4°C. This diluted culture was grown at 18°C for half an hour and induced with 0.1 mM IPTG over night at 18°C. The following day, this culture was harvested, lysed and purified by Nickle affinity chromatography. A small amount of the eluted protein was diluted in a 1:1 ratio with SDS loading buffer and a 20µl aliquot was run on SDS PAGE. It was loaded in different amounts and analysis showed a 40kDa contamination (by MS identified to be LacI repressor) and a minor degradation product. On the right side the Xpo4 with a cleaved tag was loaded. The sumo tag was cleaved by keeping Xpo4: SUMO protease in a 100:1 ratio overnight at 4°C and performing a reverse nickel purification to remove the his tagged protease and cleaved off tags. By following this strategy the background could be reduced significantly.

Exportin 4 is expressed in a stable manner and can be purified in large amounts.

The challenge of this project was to get stable, homogenous and large amounts of cargo proteins.

3.14 eIF5A1

eIF5A 1 was shown to be exported by Exportin 4 (Lipowsky et al., 2000). eIF5A1 is the only known protein to have a post translational modification called hypusine. This modification is found in certain archaea and eukaryotes, but not in bacteria.

The modification is introduced in eIF5A after the protein has been folded. In Hs eIF5A, the Lysine 50 is modified and acted upon by Deoxyhypusine hydroxylase and Deoxyhypusine synthase for hypusination.

To crystallize a protein complex, there are several crucial requirements. Firstly, eIF5A 1 has to be fully modified and homogenous. Secondly, it should be expressed in large amounts and lastly, it needs to be extremely pure.

Our goal was to generate large and homogenous amounts of eIF5A1.

We chose an in vivo strategy to introduce the modification on eIF5A. The system for protein expression is usually bacteria, because it is convenient, economical and can be used to generate large amounts of protein. We used bacterial strain BLR and co-expressed eIF5A along with the modifying enzymes. eIF5A1 was cloned into one plasmid with a cleavable His tag and the hypusinating enzymes were cloned together on the other plasmid with a non cleavable His tag. These plasmids were compatible for co expression in bacteria and had different antibiotic resistances for selection. The strategy was to co-express the hypusinating enzymes with eIF5A in bacteria and allow hypusination to occur in vivo. After this, the bacterial culture was harvested, resuspended (50 mM Tris, 500 mM NaCl, 2 mM MgCl₂, 20 mM Imidazole, pH 7.5) and lysed (sonication). The soluble and non soluble components was separated by centrifugation. The soluble proteins in the supernatant (post centrifugation) were incubated with Nickel beads to allow the His tag containing proteins to bind (1 hour at 4°C). Following this, the beads were washed with a high salt buffer (50 mM Tris, 500 mM NaCl, 2 mM MgCl₂, 20 mM Imidazole, 5 mM DTT, pH 7.5) to remove the non

specific binding proteins from the beads. The immobilized proteins were eluted with an elution buffer containing high Imidazole concentration (50 mM Tris, 100m M NaCl, 500 mM Imidazole, 2 mM MgCl₂, 5 mM DTT, pH 7.5).

When soluble proteins would be obtained in the final elution, the next step of separating the enzymes from the eIF5A was to be performed. This would involve a buffer exchange to remove the imidazole followed by a TEV cleavage to remove the Histidine tag from eIF5A (not the enzymes) and then it run over a Nickel column again (reverse Nickel step). The tagged enzymes, the tagged protease and the cleaved off His tags would be left behind and the hypusinated eIF5A would elute out.

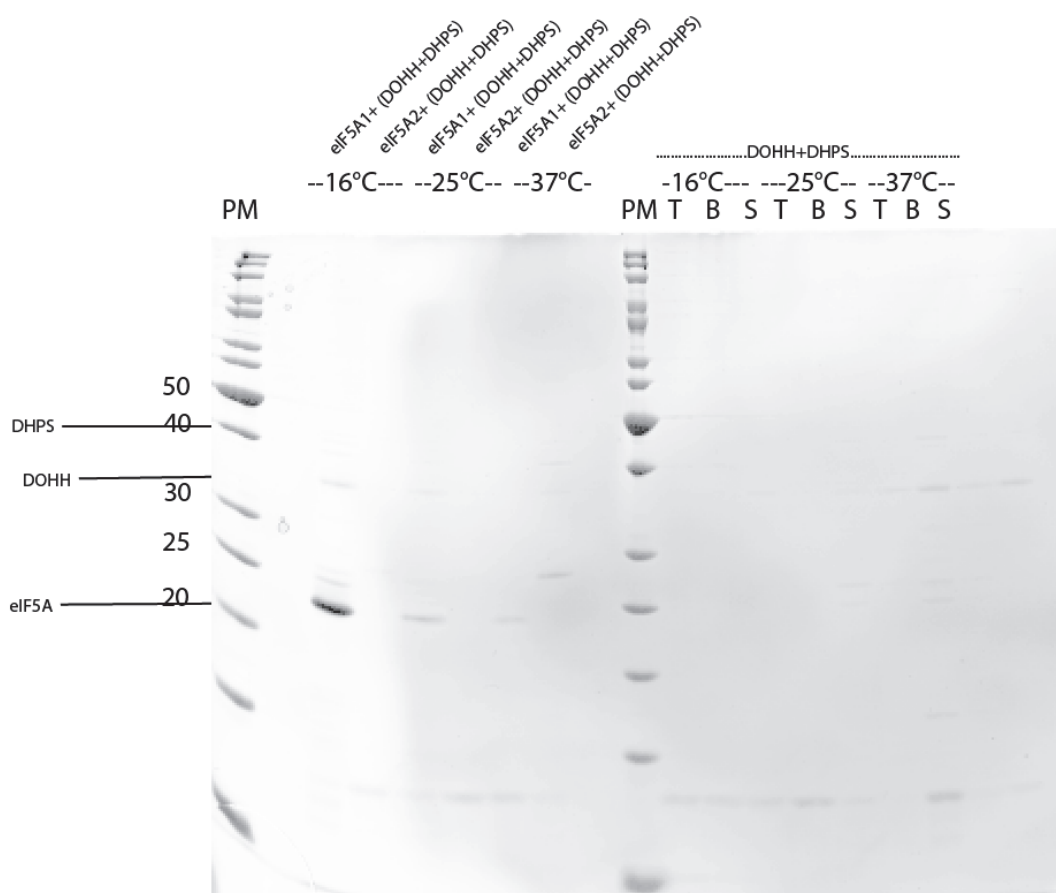
Initially there were some challenges in expression with Deoxyhypusine hydroxylase (DOHH) and Deoxyhypusine Synthase (DHPS). The expression of these enzymes was almost negligible. In order to solve this, expression was tested in BLR, TOP10F' and DE3Star at 37°C, 25°C and 16°C. In addition, various IPTG induction strengths were also tested (0.05, 0.1, 0.5, 1 mM IPTG). eIF5A 1 co-expression with the hypusinating enzymes was also tested under different growth conditions. The bacterial cultures were grown at different conditions and a small sample of cells were harvested and resuspended in SDS loading buffer for analysis by SDS PAGE.

When the hypusinating enzymes were expressed from the single bicistronic plasmid, there is no apparent difference in the expression levels of the two enzymes when comparing (bi) before induction and the (ai) after induction samples. They were expressed with different IPTG strengths, temperatures and in different bacterial strains. When eIF5A was expressed with the enzymes, expression at 16°C in DE3Star (BLR was also good) seemed to be the best temperature and strain for eIF5A expression and here DHPS expression was also detected. These are samples have a high background because of the presence of other bacterial proteins. It was important to do a purification step to pull out the His tagged proteins to clearly confirm how the enzymes are expressed.

A brief summary of the elution is shown in **figure 30**. Here the DOHH+DHPS were co expressed with eIF5A1 and eIF5A2 at different temperatures (16°C, 25°C and 37°C). On the right side of the gel, the enzymes have been co-expressed in different bacterial strains (T = TOP10F', B= BLR and S = DE3Star) at three temperatures - 16°C, 25°C and 37°C. It was evident when the enzymes were co expressed there was no soluble protein purified upon elution.

Viewing the co-expression of eIF5A 1 and eIF5A2 with the hypusinating enzymes (left side of the gel), not only was there no evidence of expression of the enzymes, but the eIF5A 1 and eIF5A 2 expression levels fell dramatically. eIF5A 1 and eIF5A 2 have been expressed in bacteria without the hypusinating enzymes and these are robust in their expression or purification. In this work, the major isoform eIF5A 1 is of importance, thus for further analysis only eIF5A1 was tested.

Based on this observation, it is clear that the enzymes are impaired in expression.



DHPS -47.5, DOHH -37, eIF5A - 22kDA. T-TOP10F', B-BLR, S-DE3Star.

Figure 30 Purified eIF5A and the hypusinating enzymes (DHPS and DOHH) after co-expression

eIF5A1 and eIF5A2 were co expressed with the hypusinating enzymes Deoxyhypusine synthase (DHPS) and Deoxyhypusine hydroxylase (DOHH) in BLR bacterial strain. The different temperatures used for IPTG induction (0.2 mM IPTG) were 16°C, 25°C and 37°C. On the right is the co-expression of the enzymes in 3 different bacterial strains – TOP10F' (T), BLR (B) and DE3Star (S). The proteins were purified by Ni²⁺ affinity chromatography and the elution was loaded (10µl of eluent from a 700ml scale preparation).

Seeing this, there was a clear indication that expression of the enzymes needed to be optimised. After reviewing literature, it was established that DOHH is an Iron containing enzyme and it would be possible that Iron may not only play a role in the catalysis but also in folding of the enzyme (Cano et al., 2010).

In the next step, Iron was added in the media in the form of FeSO₄ along with NTA as a chelating agent. This was done to compensate for any shortage in Iron

during expression of the DOHH in E.Coli. Based on previous optimisations, DE3 Star was chosen for expression at 0.5 mM IPTG at 16°C over night induction.

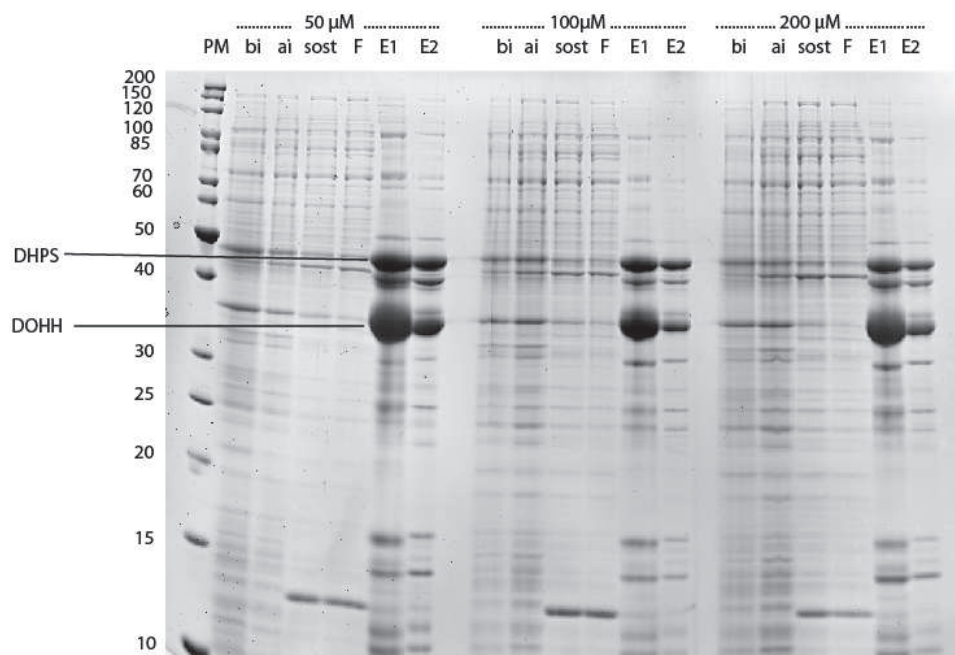
When the expression of the enzymes was tested, Iron was added in amounts of 50mM, 100mM and 200mM along with the chelating agent NTA. It was clear that there was an obvious increase in the expression levels of the enzymes and in amount of the soluble proteins which were eluted as shown in **fig 31**. When the enzymes were co-expressed with eIF5A1, there was a dramatic increase in the amount of the protein expressed and eluted after purification, as evident in **fig 31**. Hence, there was a dramatic increase in the expression and solubility of the enzymes and eIF5A when Iron was added in the media. This experiment highlights Iron being crucial for both expression and folding of DOHH. Iron concentration in all tested amounts seem to enhance expression in positively correlating amounts.

The next step was to test eIF5A for the presence of the hypusine modification. After co expression with the hypusinating enzymes, eIF5A was purified via its His tag, using Nickel affinity chromatography. The eluted protein was sent for a Mass Spectrometry analysis. However, Matrix Assisted Laser Desorption/ (MALDI) did not prove to be sufficient to detect the distinct modification. The next step will be to analyse the modification using Electrospray Ionization or ESI. With ESI an accurate mass analysis will be performed. A comparison will be done between the unmodified and hypusinated eIF5A1 to see the success of the in vivo strategy for hypusination.

In 2011, a publication was released where a co-expression strategy was also used for eIF5A hypusination. Here, as in this work, eIF5A was co-expressed with the hypusinating enzymes and the strategy for modification was successful and therefore, validates our approach (Park et al., 2011).

After obtaining homogenous and hypusinated eIF5A, trials for complex formation with Exportin 4 and RanGTP will be optimised. This stable complex will then be taken further for crystal trials.

A



B

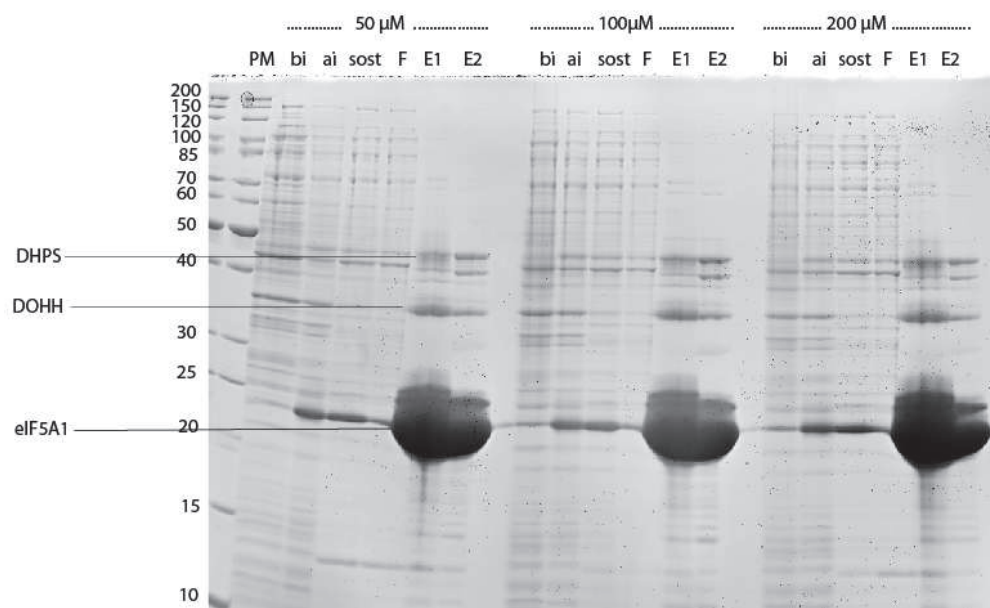


Figure 31 Expression of DHPS+DOHH with Iron in media and co-expression of eIF5A1 along with hypusine enzymes in Iron containing media.

Key: bi - before induction, ai - after induction, sost - supernatant after lysis and centrifugation, F - flowthrough after binding to Nickel beads, E1,E2 - elution of His tagged proteins with Imidazole containing buffer.

The expression was done in DE3Star at 16°C over night with 0.5 mM IPTG induction. 40mOD loaded per well except for the elution - both E1 and E2 10 μ l. E1, E2 - elution with the buffer (50

mM Tris, 500 mM Imidazole, 100m MNaCl, 5 mM DTT, 2 mM MgOAc₂, pH 7.5). In **figure 31A** - the construct expressed is a bicistronic vector with Hs DOHH (37kDa) and Hs DHPS (45.7 kDa). In **figure 31 B** - eIF5A is co-expressed with DOHH+DHPS. 2YT media with the supplements (0.2%glycerol and 20mM K₂HPO₄) was used with additional Iron to ensure better folding of DOHH. Iron was supplied as FeSO₄ with NTA (8mM FeSO₄ and 20mM NTA stock solution). 3 different concentrations of FeSO₄ were used - 50µM, 100µM and 200µM.

3.15 Sox 2 and SRY

Exportin 4 was shown to import Sox 2 and SRY. Sox 2 and SRY are transcription factors, which contain the HMG or the High Mobility Group box (Mertin et al., 1999; Remenyi et al., 2003; Gontan et al., 2009). SRY or Sex Determining Region Y is a 46.7kDa transcription factor, which has been shown to be crucial in controlling expression of genes that initiate male sex determination. An aberration in SRY causes diseases in which sex determination and/or the reproductive capability are dramatically impaired. Sox 2 also consists of the HMG box for regulation of gene expression, especially during embryonic development stages. Sox 2 and Oct 4 work together in maintaining induced pluripotent cells (Takahashi et al., 2007).

These are proteins with a relatively simple fold and a structured HMG box, which has been structurally characterized (Remenyi et al., 2003, Murphy et al., 2001). When these proteins were expressed in bacteria, they gave severe degradation problems. This has been shown in the expression and purification profile for SRY (**figure 32**). A test expression was performed to optimise expression conditions for Sox 2 and SRY. This was done using TOP10F' and BL21DE3 for expression at 25°C and 37°C at 0.2 mM IPTG induction strength. Sox 2 expression can be noticed while comparing the before and after induction samples (bi, ai respectively) for all the constructs GST-Sox 2-His 2, GST-Sox 2-His 4 and GST-Sox 2-His 8. SRY however, did not show a clear expression level under any condition. Based on these experiments, 25°C seems to be a better condition for expression. This needs to be validated by a purification step which has been described in **fig 33**.

Figure 32A:

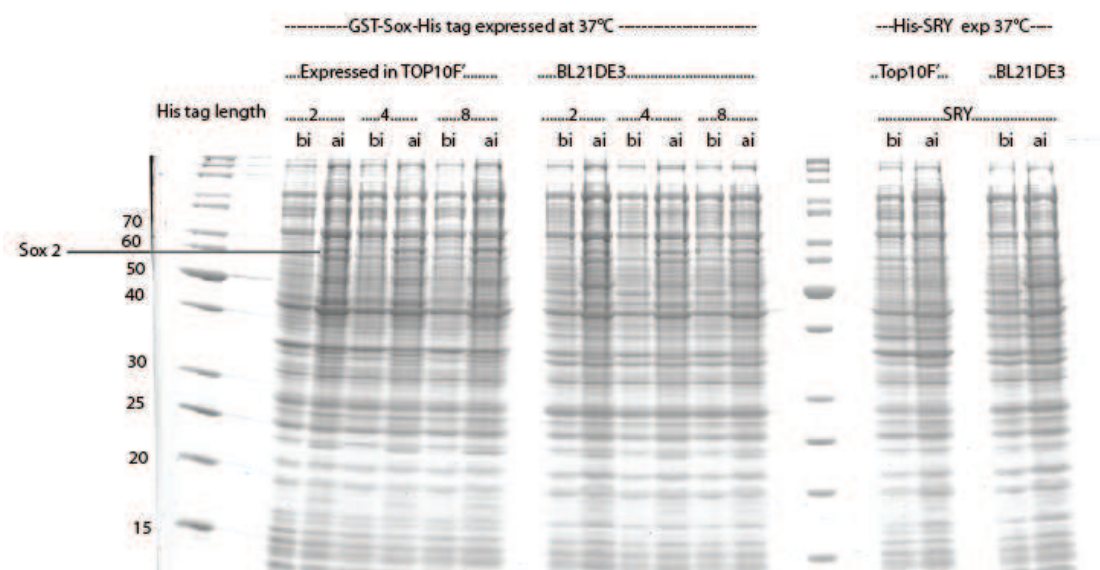


Figure 32B:

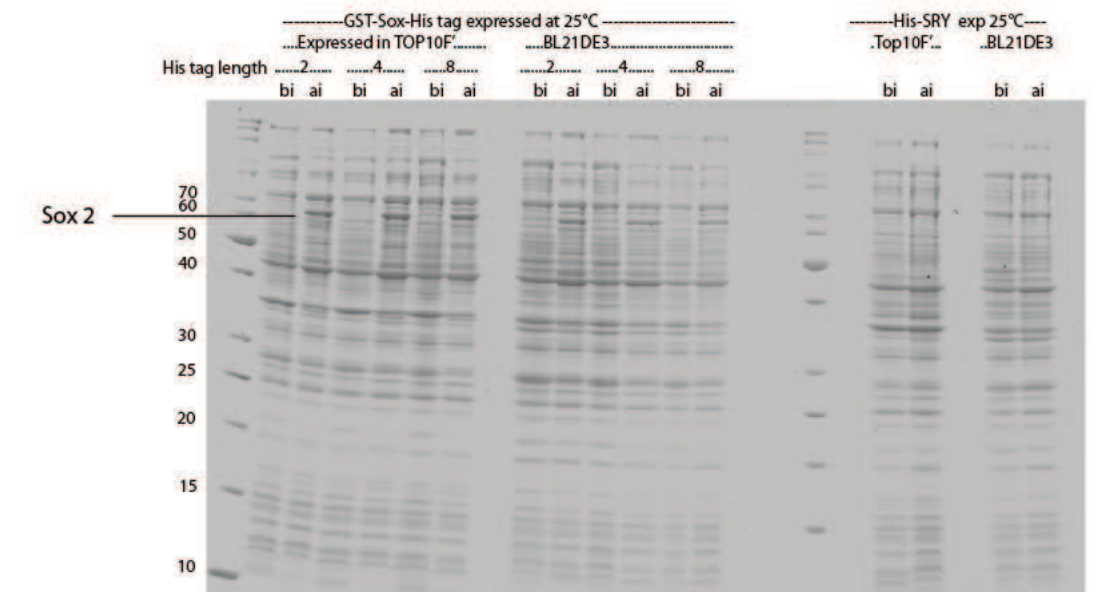
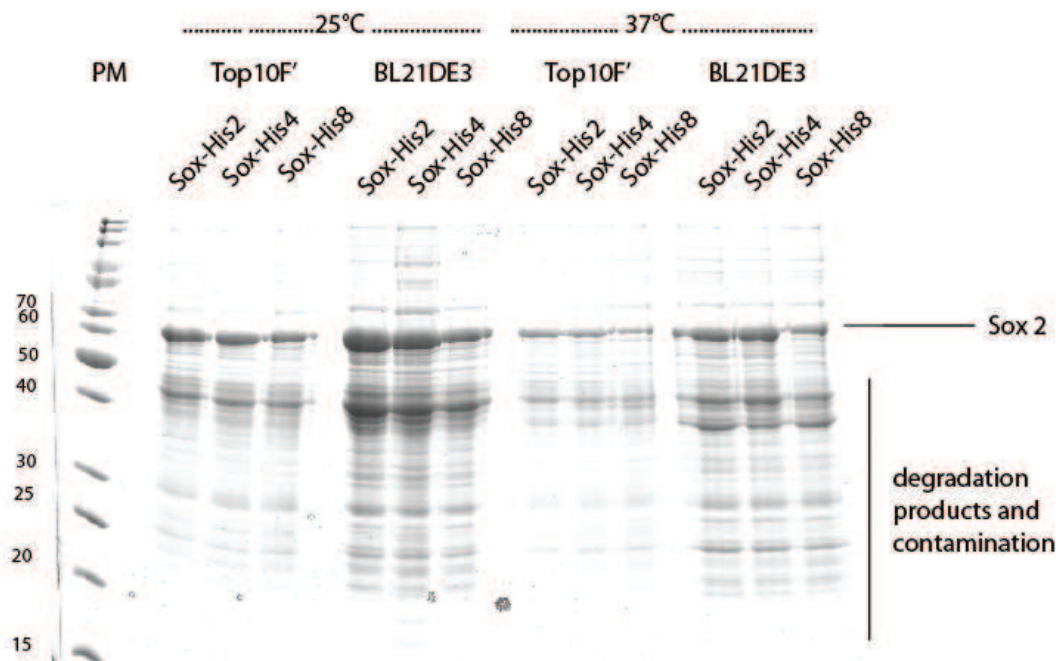


Figure 32 Optimisation of Sox 2 expression in TOP10F' and BL21DE3 with different temperatures for induction (25°C, 37°C).

Key: bi -before induction, ai- after induction.
2,4 and 8 indicate the length of the Histidine tag at the C terminus of Sox 2. Sox 2 is 61 kDa with the tags and SRY is 37.5 kDa with the tags. The proteins were expressed in TOP10F' and

BL21DE3 at 0.2 mM IPTG induction strength for 5 hours at 25°C (**figure B**) and 37°C (**figure A**). A small fraction of the bacterial culture was taken in a 1.5 ml epi, pelleted and resuspended in SDS loading buffer for lysis and denaturation (95°C, 5 minutes on a shaker). Equal amounts of the samples were loaded per lane for comparison.

A



B

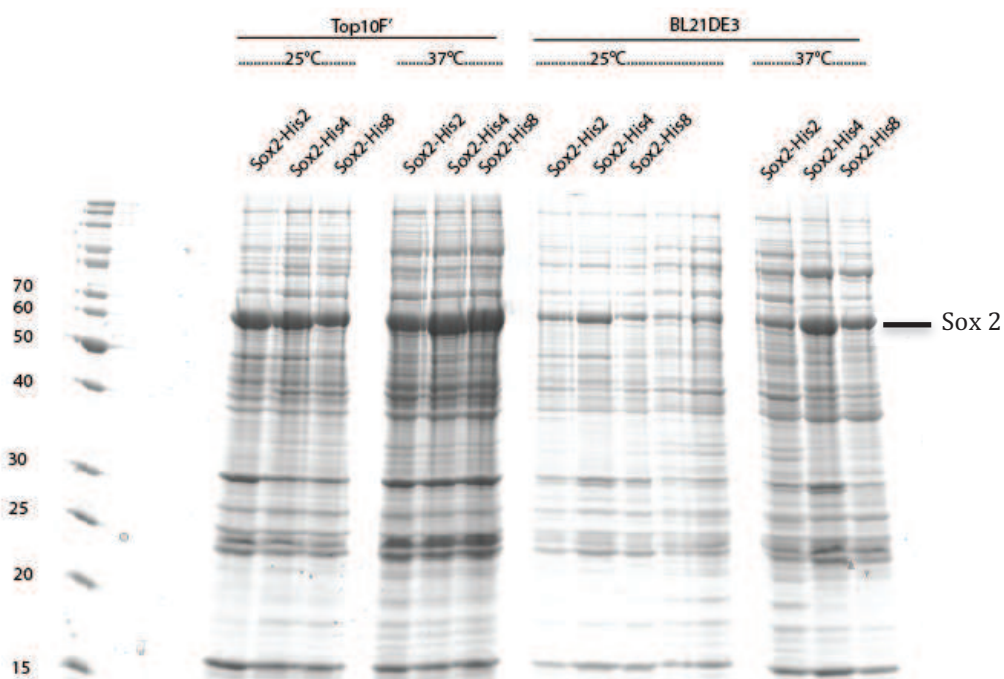


Figure 33 Purification of Sox 2 via via the GST tag at the N terminus (A) and the Histidine tags of different length at the C terminus (B).

2,4 and 8 indicate the length of the Histidine tag at the C terminus of Sox 2. Sox 2 is 61 kDa with the tags. The proteins were expressed in TOP10F' and BL21DE3 at 0.2 mM IPTG induction strength for 4 hours at 25°C and 37°C. Small scale binding assays was performed with 700µl of lysate. In **figure 33A**, Sox 2 was purified using N terminal GST tag and in **figure 33B** the Sox 2 was purified using C terminal His tag. The beads were eluted with SDS loading buffer in 50µl and 10µl of sample was loaded per lane.

The expression and purification optimisation was started by focusing on Sox 2. Sox 2 (34.3 kDa - with tags 61 kDa) was expressed in TOP10F' and BL21DE3 at 25°C and 37°C with 0.2 mM IPTG induction strength. In **figure 33A**, Sox 2 was purified using the N terminus GST tag and in **figure 33B**, it was purified via the C terminus Histidine tag. In both cases the degradation is apparent and the degradation pattern also changes when Sox 2 is purified via the N terminus or C terminus end. This was a small scale binding assay and the final elution was done using SDS loading buffer resulting in background.

From the previous test, it was clear that Sox 2 is heavily degraded during expression and purification.

The next step was to establish a procedure to isolate full length protein. Sox 2 has a GST tag at the N terminus and a Histidine tag at the C terminal. The GST tag can be cleaved off by a TEV cleavage site, but the Histidine tag did not have a cleavage site. The length of the His tag was optimised to get the shortest length sufficient for purification of Sox2. His tags of the length of 2, 4 and 8 residues were tested. This is done since a long flexible tag can cause problems during crystallography analysis. The constructs were expressed in BLR with 0.2 mM IPTG at 25°C and a tandem purification was performed. In one case the protein was pulled out via the N terminus GST tag and following this, via the C terminal His tag. In the other case, the purification was performed in the reverse order. There was success in isolating the full length protein as seen in **fig A1, A2 and A3** (Appendix). It was clear that the amount of full length protein secured after a tandem purification was minimal. For crystallography trials it is essential to isolate large amounts of protein of high purity. Thus, a new strategy needed to be developed to generate Sox 2 suitable for a further crystallography approach.

3.16 Co-expression trials:

The next strategy was to co-express Exportin 4 and Sox 2 in bacteria. By co-expressing, we hypothesized that Exportin 4 would shield Sox 2 from the bacterial proteases. A His tag would be put on the smaller binding partner - Sox 2 and by using this tag the Exportin 4 - Sox 2 complex was expected to be pulled out. Two plasmids were chosen, which were compatible for co-expression in bacteria and one plasmid had a stronger expression than the other. His-tagged Sox2 and sumo-Exportin4 were co-expressed in E.Coli strain BLR. The co-expression was performed at three temperatures 18°C, 25°C and 37°C. Sumo-Exportin 4 did not express without a Histidine tag, although previously it was expressed well with Histidine + sumo tag. As a result, the co-expression of Sox 2 with Exportin 4 did not show any improvement in expression and isolation of full length Sox2.

In the other co-expression trial, the expression plasmids were reversed and untagged Exportin 4 was placed on the plasmid with the stronger expression capacity. In this trial, the untagged Exportin 4 was co-expressed with Sox2, Sox2 HMG domain, SRY and SRY HMG domain. We hypothesized that with a change in expression plasmid, Exportin 4 would show better expression. The samples were purified using the His tag on the transcription factors via Nickel affinity chromatography. Isolating full length Sox 2 was not possible when Sox 2 was expressed alone or with Exportin 4. This is true for SRY as well. When the Sox 2 HMG and SRY HMG domains were expressed alone, they give high yield but upon co-expression with Exportin 4 their yield dropped dramatically. In spite of changing the plasmid, the expression of Exportin 4 needs further optimisation because it did not show any improvement. For the success of the co-expression strategy, Exportin 4 expression without a tag would be the immediate issue to be addressed.

3.17 SRY

When SRY was expressed it gave severe degradation problems. The Hs SRY was expressed in BLR bacterial strain at 25°C and 0.2 mM IPTG for 4 hours. It was purified via its N terminal His tag and purified via Nickel affinity chromatography. The bacterial strain was also transformed with the empty plasmid (missing the insert) and treated in exactly the same manner as the bacteria transformed with the plasmid with SRY cloned into it. By comparing the elution from these two samples, we can clearly compare which protein bands come from non specific binding from the bacterial lysate to the Nickel beads and the bands that result from SRY degradation as shown in **Figure 34A**. The expression was optimised by expressing in TOP10F', BLR and MH1Alpha bacterial strains to reduce degradation (**Figure 34B**). In all cases, the degradation was evident and it was obvious that a new strategy was required to isolate full length of SRY for crystallographic analysis.

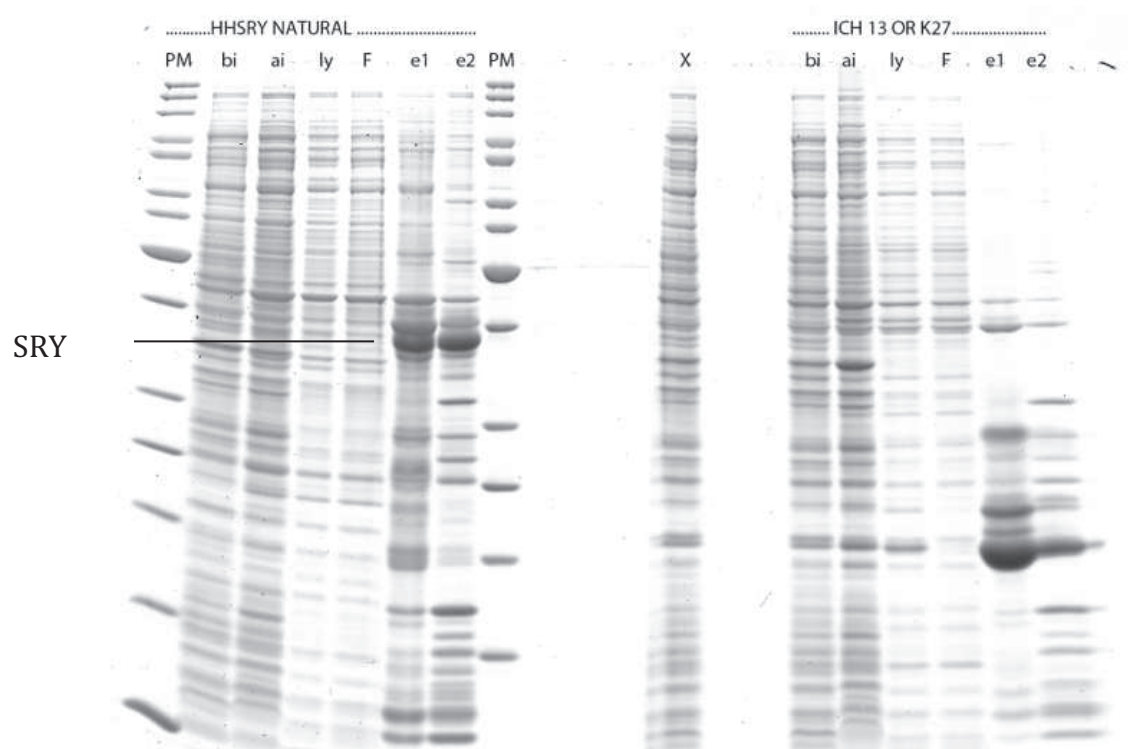


Figure 34 A: Expression of Hs SRY in BLR

Key: bi- before induction, ai- after induction, ly - lysate, sost - supernatant after lysis and centrifugation, F- flowthrough, E1 - elution with elution buffer (50mM Tris, 100m M NaCl, 500 mM Imidazole, 2 mM MgOAc2, 5 mM DTT), elution 2 - elution with SDS loading buffer.

HsSRY cloned into a Kanamycin vector called 'K27' was expressed in BLR bacterial strain. The expression condition was 0.2 mM IPTG, 25°C, 4 hours. This culture was harvested, lysed and purified by Nickle affinity chromatography. A small amount of the sample was diluted with SDS loading buffer in a 1:1 ratio and 20µl of this was loaded. The gel was stained by coomassie blue and as seen there are many degradtion products. SRY is 37.5 kDa (with the tags).

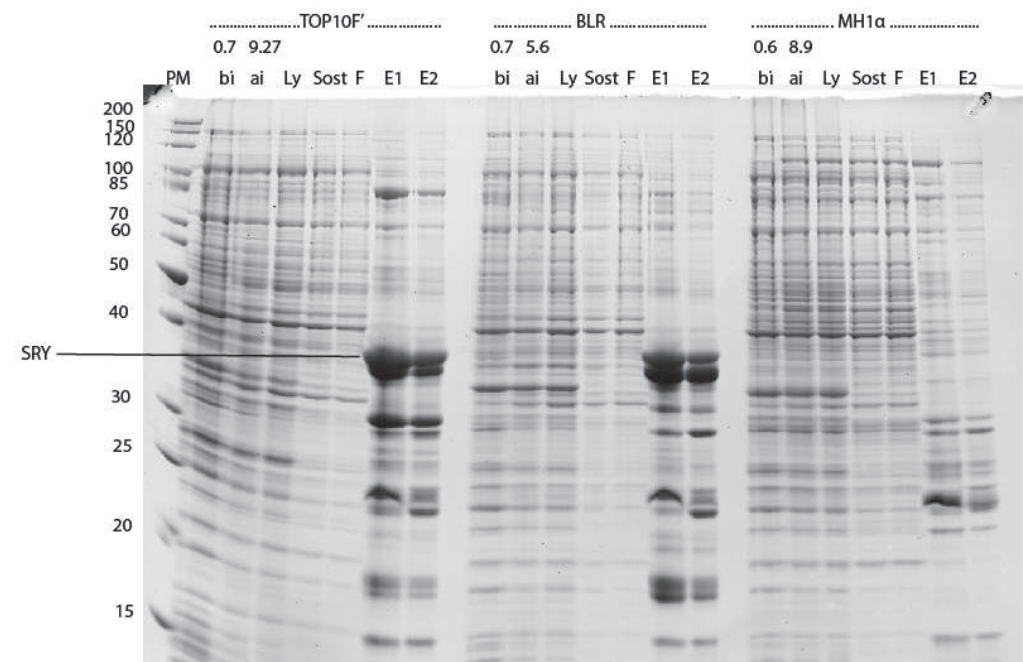


Figure 34 B: Expression of SRY in different bacterial strains

HsSRY cloned into a Kanamycin vector called 'K27' was expressed in BLR, TOP10F' and MH1alpha bacterial strains. The expression condition was 0.2 mM IPTG, 25°C, 4 hours. The purification procedure is similar to as described in Fig 34A.

3.18 Synthetic genes optimised for expression in bacteria:

To increase the yield of Sox 2 and SRY and to reduce degradation, synthetic genes were ordered with codons optimised for more efficient expression in

bacteria. It was clear by seeing the elution that there was degradation in the full length protein expressed from synthetic genes. The HMG box of Sox 2 has been crystallized in complex with DNA and the POU domain of Oct 4. The residues pertaining to the HMG box region were cloned from the synthetic Hs Sox 2 and synthetic Hs SRY were expressed in BLR bacterial strain and showed high yield and very little degradation (**figure 35 - A**).

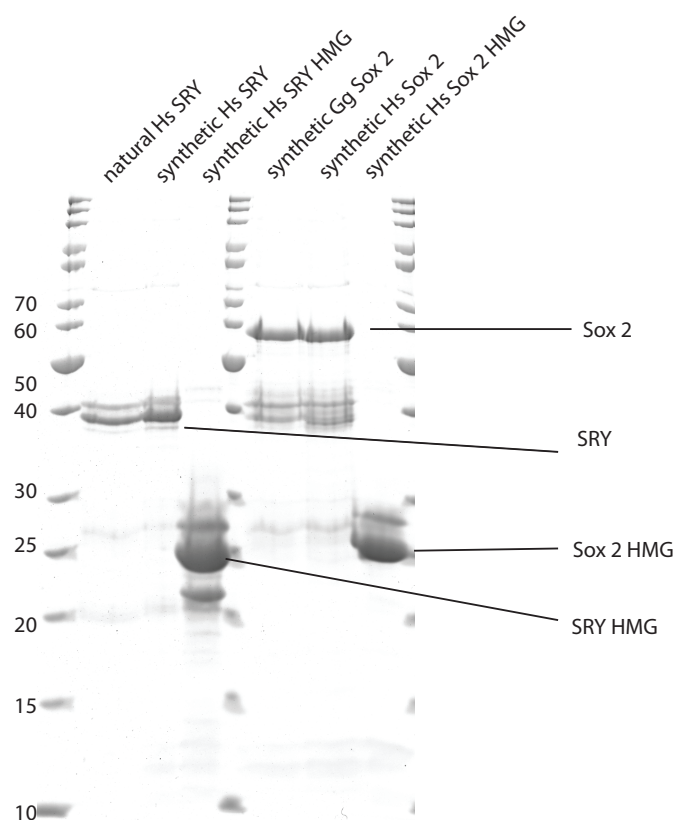


Figure 35-A Analysis of purified Sox 2 and SRY expressed from natural and synthetic genes (optimised for bacterial expression).

Synthetic genes optimised for Sox 2 and SRY expression in E.coli were cloned into a vector. This vector is optimised for expression in bacteria and has a His tag at the N terminus. The constructs were expressed in BLR at 25°C with 0.5 mM IPTG induction and purified via Nickle affinity chromatography. Two homologs of Sox 2 were expressed – from Gallus gallus (Gg Sox2) and homosapien (Hs). For SRY, Hs SRY from natural and synthetic sequences was expressed. The HMG domain of synthetic Hs Sox 2 and synthetic Hs SRY was also expressed and purified.

3.19 Smad3:

Along with synthetic genes for Sox 2 and SRY, Smad 3 gene was also ordered for optimal expression in bacteria. This is a transcription factor which is exported by Exportin 4. In a previous study, several constructs were made spanning the MH1 and MH2 domain of Smad 3 and their binding to Exportin 4 in the presence of RanGTP was tested (Kurisaki et al., 2006). Several high affinity patches were

found all along Smad 3. We must take into account that disrupting a natural fold can change the behaviour of the protein and expose hidden regions/disturb exposed surfaces. From this study, we can not conclude that there is a single stretch that binds best to Exportin 4 for export. But what we can conclude, is that there could be many high affinity patches all along the surface and Exportin 4 may interact with Smad 3 over a larger surface area. However this speculation can only be validated with a crystal structure. Smad 3 expression looks promising, it is stable and expresses in large amounts (**figure 36 A**). It had been cloned into a His tagged plasmid and it was purified using Nickel affinity chromatography. Next goal is to perform binding assays to test if it interacts with Exportin 4 with high affinity. Following this, stable complexes of Exportin 4 Smad 3 and RanGTP will be made for crystallographic trials.

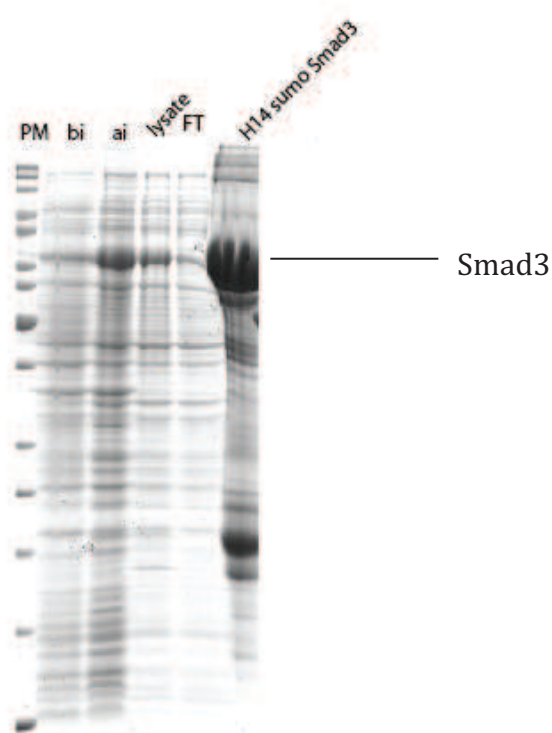


Figure 36 A - Synthetic genes were also ordered for Smad 3 and this shows stable expression.

The expression conditions are as follows: BLR strain for expression at 25°C with 0.5 mM IPTG for 4 hours. Standard buffers were used for purification:

Resuspension - 50T, 500N, 2 Mg, 25 I, pH 7.5;
Purification - 50T, 500N, 2 Mg, 25 I, 5 DTT pH 7.5; Elution - 50T, 500N, 2 Mg, 400 I, 10 DTT pH 7.5. Purification was performed via His tag on Smad 3 using Nickel affinity chromatography.

3.20 The expression of the HMG box domain of Sox 2 and SRY.

As shown in **figure 37**, the HMG box domain of Sox 2 expresses very well, is soluble and can be purified in large amounts. The constructs were expressed in BLR at 25°C for 4 hours and purified via Nickel affinity chromatography. This is ideal for crystallography. Since the behaviour of SRY in protein expression and purification matches Sox 2 closely, synthetic genes were also ordered for SRY and the HMG box of Hs synthetic SRY was also expressed. As seen in **figure 37**, SRY HMG domain expressed well, was soluble and was eluted in large amounts too. However, there is minor degradation, which can be removed by tag cleavage and reverse Nickel purification followed by size exclusion chromatography.

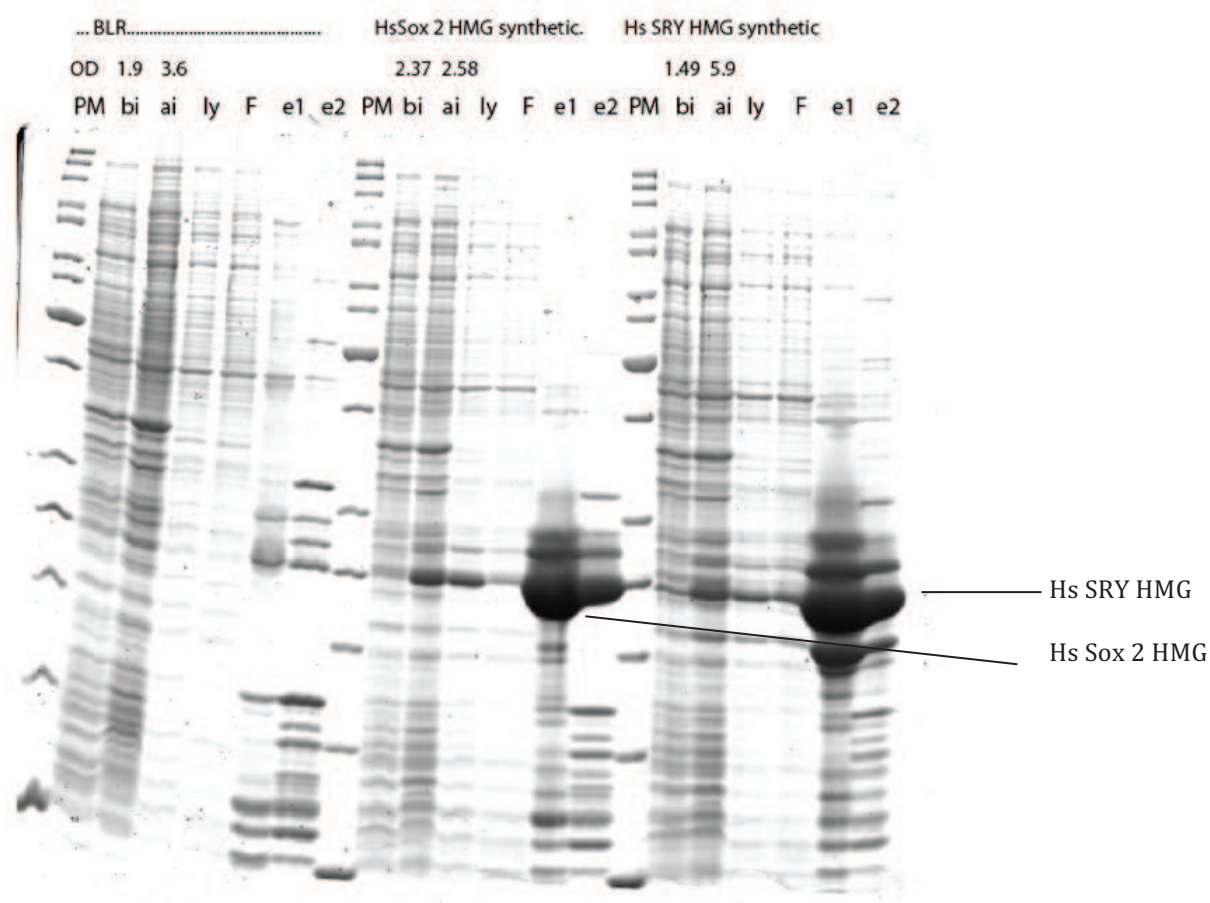


Figure 37: Expression of Sox 2 HMG and SRY HMG domains.

Figure 37: The HMG domain of Sox 2 and SRY were expressed in BLR bacterial strain (0.5 mM IPTG, 25°C induction temperature, 4 hours induction time) and purified via Nickel affinity chromatography. On the left the bacterial strain (un transformed) was treated exactly the same as the samples expressing the transformed constructs.

Bi – before induction, ai – after induction, ly – lysate post centrifugation, F – flowthrough, e1 – elution one, e2 is elution 2. For samples bi,ai,ly, F – 40 mOD loaded per lane. 10µl loaded for e1 and e2. This was a large scale preparation (700 ml).

If the HMG box regions are imported by Exportin 4, then a binary complex of Exportin 4 and the HMG domain should be formed in the absence of RanGTP. Nuclear Localisation Signals or NLS often have basic residues in the import signal. The HMG box also has basic residues and it is reasonable to suggest that this region has a dual function - an import signal to carry the transcription factors into the nucleus and act as regulators of gene expression inside the nucleus. Recently, it was showed that the HMG box region of Sox 9 is recognised by Exportin 4 for its import and this further supports our claim that HMG boxes have a dual function of regulating gene expression and carry signals for nuclear import.

The next step is to perform binding assays with Exportin 4 and the HMG domain of Sox2 and SRY. This step will reveal whether this region does contain the NLS for nuclear import. After establishing that this region is successfully recognised by Exportin 4 and forms an import complex, the focus will be to generate pure and large amounts of homogenous complexes of Exportin 4 bound to the HMG domains of Sox 2 and SRY. These will then be carried forward for crystallography trials.

4 Discussion

4.1 Nuclear export by CRM1

CRM1 is the NTR with the most versatile cargoes and it can recognize two types of export signals - a) the simple linear sequences called the Leucine rich Nuclear Export Signals. These are found in PKI-NES, Rev-NES etc. This is a well defined classical Leucine rich signal or PKI type NES and it contains the consensus Φ 0 Φ 1 Φ 2 Φ 3 Φ 4 (Güttler et al., 2010). CRM1 also recognizes unconventional signal motifs as found in Snuporin1. In some cases, an adaptor may be required for the recognition and export of the cargo. So far, the binding of the CRM1 cargo has been detected to occur in the 'hydrophobic cleft'. The crystal structure of SNP1-CRM1-RanGTP gave clear insight into the mechanism of cargo interaction by CRM1 for export. Upon RanGTP binding, CRM1 is 'strained' and this opens the hydrophobic cleft, enabling the cargo's export signal to bind.

The hydrophobic cleft of CRM1 is rigid and narrow. The CRM1 residues surrounding the binding cleft and the acidic loop are highly conserved. This indicates that the cooperative action of RanGTP and cargo binding is a conserved mechanism. Hence, the binding cleft has retained its conformation and the different export signals adapt their structure to correctly position the hydrophobic Φ residues in the CRM1 binding cleft (Monecke et al., 2009, Güttler et al., 2010).

The classical Leucine rich consensus is well established and bioinformatic tools have been developed to predict this NESs in a submitted protein sequence for CRM1 binding. However, these can tend to be inaccurate and lead to detection of false positives or false negatives. If an algorithm is based on a primary sequence, it can falsely detect a sequence that could be a part of a structured fold or hidden in the interior of the protein. On example that emphasises this, is the NES analysis of the Abl Tyrosine Kinase. Here a Leucine rich NES was located in the Actin binding domain. However, when the structure was studied, it was found

that this sequence is hidden in the hydrophobic core of the protein. Thus, the identified sequence could not function as an NES.

Thus, disorder prediction, solvent exposure and structural characteristics are important factors to be included.

Moreover each Φ position favours Leucine, Isoleucine, Valine, Phenylalanine etc which introduces variability in the prediction of these residues. An important aspect to consider is that in nature, sequences of the NESs are deliberately attenuated for 'imperfect binding' to CRM1. This is essential because, if an NES has extremely high affinity, the disassembly of the export complex requires a large amount of energy to be disassembled and this is impractical.

An export signal should be such that the binding of the cargo is strong enough to be effectively exported and then easily disassembled in the cytosol. The deliberate attenuation of the export signal can be achieved by skipping of Φ residues, incorporation of atypical residues at Φ positions and imperfect spacing. Thus these are important features which should be considered while designing algorithms for NES prediction.

Moreover, the spacer residues also have additional functions which have not been addressed in previous studies. Defining the contribution of the spacer residues to the NES binding could provide valuable information for refining the bioinformatic tools for NES prediction.

There are two CRM1 cargoes, RanGAP 1 and eIF2 β where the nuclear export signal remain to be identified. During the course of my research work, I strived to identify the export signals in these CRM1 cargoes for nuclear export. The consensus used to define the classical Leucine rich NES found in CRM1 cargo (e.g. PKI-NES and Rev NES) is $x\Phi^0x(2)\Phi^1x(3)\Phi^2x(2-3)\Phi^3x\Phi^4$. We referred to this NES consensus in our analysis.

4.2 eIF2 β

eIF2 β is a translation factor which is a small sized protein that can passively diffuse into the nucleus or be accidentally encompassed in the nucleus during the membrane reformation post cell division. This accidental entry needs to be corrected by constant exclusion from the nucleus. When the dependence of nuclear export of eIF2 β on CRM1 was tested by blocking CRM1 using Leptomycin B, a rapid nuclear accumulation of eIF2 β was observed. This suggests that eIF2 β is actively imported since passive diffusion could not sustain such a rapid rate of nuclear accumulation. CRM1 could be correcting this accidental import.

eIF2 β has poly Lysine tracks in its N terminal domain. This could be accidentally recognised as an import signal. Moreover, the importin could also serve the role of a chaperone. Such charged proteins have a tendency to aggregate and masking the charged region could be one way to prevent that. Thus, an accidental binding of eIF2 β to an importin could be one way to explain its rapid nuclear accumulation when CRM1 function was inhibited.

Nuclear exclusion of translation factors (e.g. eIF2 β) is one of the ways employed by the cell to stop translation in the nucleus. During the process of generation of a fully processed and spliced transcript, it is crucial that the translation machinery is spatially separated. If unspliced transcripts or incompletely processed transcripts are translated, this can lead to formation of proteins that can have pre termination codons, missense mutations, behave as dominant negative mutants or they may be misfolded. These in turn can have a dramatic effect on cell metabolism and can prove to be lethal (Bohnsack et al., 2002).

The goal was to investigate the molecular mechanism for recognition and consequent export of eIF2 β by CRM1.

When the sequence of eIF2 β was scanned to find a familiar Leucine rich NES no positive hits were found. This could be because of 2 reasons: the NES prediction results were unreliable or eIF2 β has a novel NES and uses a new mechanism for

CRM1 binding and nuclear export. An unbiased approach was undertaken to locate the export signals of eIF2 β .

4.2.1 eIF2 β interacts with CRM1 in the hydrophobic cleft and contains a classical PKI type Leucine rich NES

We addressed the question if eIF2 β interacts with CRM1 in the hydrophobic cleft. Based on the crystal structure of SNP1-RanGTP-CRM1, the mechanistic details of formation of this export complex were clarified. There was an evident cooperation in the cargo and RanGTP binding to CRM1. Based on this, some CRM1 mutants were designed to break this cooperative binding of RanGTP and cargo. A541E CRM1 mutant was designed such that the cargo can not bind to CRM1, irrespective of the presence or absence of RanGTP. This behaviour can be seen in cargo possessing the classical Leucine rich signal e.g. the PKI-NES. We compared the behaviour of eIF2 β with PKI-NES in binding to A541E CRM1 mutant. We saw a clear similarity, and this suggested that eIF2 β does interact with CRM1 in the binding cleft. This also confirmed that eIF2 β could have a classical Leucine rich NES and the prediction by the NES prediction tools was imprecise.

4.2.2 Deciphering the eIF2 β NES

An unbiased approach was undertaken by generating truncations of eIF2 β and we located a sequence that functions as the NES in the unstructured N terminal region of the protein. This is reasonable, since this region is solvent exposed and flexible. Hence, it can fit into the narrow and rigid binding cleft of CRM1 with ease. We then performed a deeper analysis of this stretch of residues and predicted residues that could fit the classical Leucine rich NES consensus. We confirmed this prediction by biochemical analysis. The hydrophobic Φ residues were mutated and their interaction with CRM1 was analyzed by binding assays and export capacity was tested in an in vivo system by transfections in HeLa cells. An atypical residue - Alanine, was found at the Φ 1 position of the NES. We

showed that this residue actively contributes to influencing the affinity of the eIF2 β NES to CRM1.

The discovered NES is spanning residues 89 to 103 as shown below. The identified Φ residues have been highlighted in yellow.

I	F	D	I	D	E	A	E	E	G	V	K	D	L	K	I	E
87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103

4.2.3 An atypical Φ 1 hydrophobic residue was discovered in the eIF2 β NES

During this work, we identified that an Alanine occupies the Φ 1 hydrophobic residue in the eIF2 β NES. This was an atypical residue to be found at a Φ position and this probably made detection of the NES by bio informatic tools difficult. It was apparent that this position is very sensitive to the type of amino acid introduced. For instance, when mutated to an Arginine or Glycine the binding went down dramatically, but when a Leucine was introduced, the binding went up significantly compared to that of the wild type NES. Hence, apart from its function in translation, this residue also plays a crucial role in modulating the affinity of the eIF2 β NES to CRM1. This makes sense from the biological perspective. If the NES binds to CRM1 with extremely high affinity, then it will remain bound to CRM 1, even after being exported to the cytosol, because disassembling the export complex will require too much energy. This will result in a futile cycle because the cargo will not be released in its destined location. By doing a homology search, it was apparent, that this residue is highly conserved, with the exception that Threonine was found at Φ 1 in the *Xenopus Tropicalis* eIF2 β NES. We modified the *homo sapien* NES to match the *Xenopus Tropicalis* NES and found that a Threonine can be tolerated at this position, but there is a slight decrease in the export efficiency in the A Φ 1T mutant of Hs eIF2 β NES. We wanted to address the issue of incorporation of atypical residues at Φ 1 position in a classical PKI type NES and we wanted to use the eIF2 β NES as a reference. This will also help us understand which residues have a less probability of being permitted at this Φ position in the classical PKI type NES from the biological perspective.

This Φ 1 Alanine was mutated to bulky, hydrophobic, polar, charged and rigid residues successively and the interaction with CRM1 was analysed. It was clear that hydrophobic residues like Isoleucine and Valine were tolerated well, by virtue of their side chain properties in comparison to the other amino acids. A bulky Tyrosine, a polar Glutamine or Serine, a charged Lysine, Histidine or a conformationally restricted Proline are not allowed at the hydrophobic Φ 1 position because of dramatic weakening of the NES affinity to CRM1. These GFP tagged mutants were expressed in HeLa cells by transfection, and their sub cellular distribution validated the conclusions based on the previously conducted binding assays.

The PKI NES was also tested for incorporation of atypical Φ residues. Φ 0 was found to be most tolerant and can accommodate the widest variety of residues compared to the other Φ positions. Φ 1 and Φ 3 showed variable behaviour depending upon the type of residue that was introduced. This is common for the PKI and eIF2 β NES - Φ 1 residue in the eIF2 β NES also showed variation in interaction with CRM1 for export, depending upon the residue introduced.

4.2.4 Defining the spacer residues in the eIF2 β NES

In this study, we investigated the role of the spacer residues that surround the Φ positions in the context of the eIF2 β NES and their contribution to the binding of the NES to CRM1 in the binding cleft. We first addressed the contribution of these spacer residues from a structural perspective. We successively mutated the spacer residues to Prolines and Glycines and tested them for export by CRM1 by transfections and direct interaction with CRM1+RanGTP in binding assays. Here, their contribution to help the eIF2 β NES to adopt the correct conformation to fit inside the rigid binding cleft of CRM1 was tested.

The mutants showed reduced export when Glycines were introduced at spacer residues surrounding Φ 1, (E92G, E94G, E95G). A stronger reduction in export efficiency was observed at positions close to Φ 2 (K98G) and when Glycine was introduced at K101, which is adjacent to Φ 3 and Φ 4.

Upon introduction of Prolines, the fall in export efficiency was much more dramatic compared to that of Glycine mutations. This is also probable because of the conformationally restricted nature of Proline could hinder the fitting of the NES in the hydrophobic cleft. Thus it was not surprising that the Proline mutation of the spacer residues at all positions dramatically reduced export. Minimal export was observed when Prolines were introduced at positions D89 and D91 which are adjacent to $\Phi 0$.

To further understand the importance of the spacer residues - the amino acids had been mutated to bulky, less acidic or polar and hydrophobic residues.

Introduction of bulky Tyrosines at any position in the spacer region almost abolished binding since the NES has to fit inside a narrow and rigid binding cleft. Tyrosines allowed weak export when introduced at spacer positions close to $\Phi 0$ (D91) and $\Phi 3$ (D99) and the spacer residue immediately following $\Phi 4$ (E103).

When the acidic nature of the NES was reduced by successive mutations to Glutamine, there was an apparent decrease in the binding to CRM1 at specific positions. This effect was more pronounced at particular locations e.g. Glutamine introduced at spacer residues surrounding $\Phi 2$ (G96 and K98) and $\Phi 3$ (K101) abolished export by CRM1. A dramatic decrease was also evident in the D89G mutant - suggesting that an interaction close to $\Phi 0$ was disrupted by this mutation. When Serine was introduced, the effect was similar to that of introduction of Glutamines.

Lastly, the aim was to replace the acidic spacer residues by hydrophobic Alanine. This was well tolerated at many spacer positions and showed export similar to the wildtype NES. When the spacer residues surrounding $\Phi 2$ (G96 and K98) and $\Phi 3$ and $\Phi 4$ (K101) were mutated to Alanine a decrease in nuclear export of the mutants was observed.

In summary, spacer residues surrounding hydrophobic positions $\Phi 2$ and $\Phi 3$ are most sensitive to the mutations. These are positions deep inside the binding cavity of CRM1. $\Phi 0$ seems to be more resilient to the introduction of the different mutations. This reduced sensitivity to introduction of various residues at $\Phi 0$ position of the NES was also observed in the PKI-NES. We also confirm that the residue immediately adjacent to $\Phi 4$ also plays a crucial role in influencing the affinity of the eIF2 β NES to CRM1. Hence the consensus should be expanded to include this residue as well (additional spacer residue is shown below in red).

x $\Phi 0$ x(2) $\Phi 1$ x(3) $\Phi 2$ x(2-3) $\Phi 3$ x $\Phi 4$ x

Hence, by systematic mutations of the residues surrounding the Φ positions to bulky (Tyrosine), hydrophobic (Alanine), polar (Glutamine), conformationally restricted (Proline) and flexible (Glycine) residues, the significance of the spacer residues in influencing the interaction of the NES in the CRM1 binding cleft and helping the NES acquire the correct conformation was evident.

This information has defined which residues can be tolerated at the spacer positions. We have also analysed which atypical residues can be tolerated in the context of the $\Phi 1$ position of the eIF2 β NES. By using this information we can generate an accurate and reliable algorithm for identification of classical Leucine rich Nuclear Export Signals for export by CRM1.

4.3 RanGAP 1

RanGAP 1 is crucial in maintaining the RanGTP gradient across the nuclear envelope and in disassembly of export complexes as they exit the nucleus. RanGAP 1 can be accidentally included in the nucleus and it has been shown that it is excluded from the nucleus by CRM1.

If RanGAP 1 is present in the nucleus, export complexes disassembly will be disrupted in the cytosol and their assembly will be disturbed in the nucleus. The RanGTP gradient will be dramatically affected and inhibit nuclear transport.

The mechanism of export of RanGAP 1 by CRM1 is unknown and deciphering the export signal in RanGAP 1 for its recognition by CRM1 is the goal of my project.

4.3.1 Bioinformatic tools to predict the RanGAP 1 NES

When the RanGAP 1 sequence was submitted in bioinformatic servers to detect a Leucine rich NES that fits the defined consensus, several hits were located. However, this is a Leucine rich protein and many hits were false positives e.g. NESs that follow the consensus to bind in the CRM1 need to be flexible to fit into the rigid binding cleft of CRM1. Finding hits in structured regions like an alpha helix or regions hidden in the protein folds is improbable.

4.3.2 Trials to locate an NES containing fragment of RanGAP 1

An unbiased approach to locate the NES was undertaken. Truncated constructs were created all along the protein. The interaction with CRM1 was tested in the presence and absence of RanGTP.

The initial set of truncated constructs indicated that the best binding to CRM1 was seen in the constructs spanning the N terminal LRR region. This interaction was only observed in the presence of RanGTP. Further constructs were designed spanning the full length protein and two stretches of residues (1 to 77 and 1 to 369 amino acids) that lie in the N terminal Leucine Rich Region of RanGAP 1 were found to interact with CRM1 in the presence of RanGTP with high affinity in the presence of RanGTP.

The constructs were designed after analysis of structured folds, conserved regions and disorder prediction. The LRR region could be folded in one continuous stretch or in discrete groups with interspersed flexible linkers. Seeing the SMART protein prediction of the domain structure, the LRR were depicted as being in roughly three groups with flexible linkers in between. These linker regions were targeted for cutting to creating smaller constructs in the LRR

region. However, if the LRR were continuous, then cutting in a continuous stretch can expose hidden regions that could act as false positives.

Usually transport signals lie in the ends of domains and exposed regions. Of course there are also cases where a conformational change is introduced upon interaction and consequently opening up a binding domain/sequence. A strong way to validate if the identified sequences are positive hits for an NESs, the approach would be to mutate the predicted hydrophobic residues in the identified NES in these regions and test affinity for CRM1 and RanGTP in binding assays. To ensure that this mutation does not affect the folding and activity of RanGAP 1, these mutants can also be tested for GTPase activity. Another effective way to validate if the conformation of RanGAP 1 is unaffected by the mutations, is to test its interaction with RanGTP and RanBP1. If this trimeric complex is formed with the mutated version of RanGAP 1, then the mutations do not affect the structure. An *in vivo* validation must be done by transfecting mutants expressed with a GFP tag into HeLa cells and tracing the localisation of the various mutants.

So far the crystal structures of RanGAP 1-RanGTP-RanBP1, CRM1-RanGTP-RanBP1 (yeast homologs of these proteins) and CRM1-RanGTP-SNP1 have been solved and these structures have given valuable insight into the function of RanGAP 1 and RanBP1 and interaction surfaces of CRM1 with its cargo and RanGTP. It would be very interesting to study how RanGAP 1 is exported by CRM1 without hydrolysis of RanGTP in its own export complex. CRM1 may bind RanGTP and RanGAP 1 in a manner to spatially segregate them or bind RanGAP 1 in a conformation that disables its influence on RanGTP. Hence, another goal will be to crystallize RanGAP 1-CRM1-RanGTP complex to get mechanistic insights into the nuclear export of RanGAP 1 by CRM1 and clearly see how the NES interacts with CRM1.

4.4 Structural analysis of Exportin 4 and its cargo

There are some transport receptors e.g. Importin 13 and Exportin 4 which have been shown to have a dual function as an importin and an exportin. Exportin 4 was of interest to me and this transporter was shown to export eIF5A1 and Smad 3 (Lipowski et al., 2000; Kurisaki et al., 2006). In 2009, Exportin 4 was shown to import Sox 2 and SRY as well (Gontan et al., 2009). The goal was to do a structural analysis of Exportin 4 bound to its export and import cargo. By seeing the structure the regions that come into contact would be clearly identified. The residues of the cargo that form the transport signal are clarified this can be used to define recognition motifs. This information can be used further on to develop bioinformatic tools to detect similar motifs in other proteins and possibly identify new interacting proteins as well. In the case of export cargo, valuable information regarding the influence of RanGTP on the structure and consequently, function of the transport receptor and the steps involved in building the export complex also get clarified.

A useful strategy while studying these proteins is to crystallize them when they are bound to their cargo. This will contribute stability to these flexible proteins. Moreover, it is also more informative to crystallize the complexes in their functional states to understand the mechanism of cargo recognition and consequent transport. Other important prerequisites for crystallography include generating large amounts of stable and homogenous complexes.

When Exportin 4 was expressed in bacteria, it showed good expression levels. It was soluble and the isolated protein showed minor degradation which can be removed easily by further purification steps.

The major challenge of crystallizing Exportin 4 bound to its export and import cargo is isolating homogenous, stable and large amounts of Exportin 4 cargo.

4.4.1 Hypusination strategy for eIF5A1

The biological function of eIF5A is under speculation and this is the only known protein to have the post translation modification called Hypusine. This is essential for viability and is found in in some archea and eukaryotes, and hypusination does not occur in bacteria (Lipowski et al., 2000). It was also observed that Exportin 4 has maximum affinity for eIF5A when it is fully hyousinated. Thus the goal was to generate large amounts of fully hypusinated eIF5A.

Hypusination is carried out by Deoxyhypusine synthase (DHPS) and Deoxyhypusine hydroxylase (DOHH). To generate a fully modified eIF5A, an in vivo strategy was adopted. E.Coli is a standard expression system of choice since it is fast, cost effective and convenient to express proteins. However, hypusination is not found in prokaryotes. The strategy was designed in a way that the DOHH and DHPS enzymes were put on one plasmid and eIF5A1 was put on the other plasmid. These were co-transformed and the hypusination was allowed to proceed in vivo. This is a more reliable approach compared to an in vitro method because it occurs in a more natural environment which can facilitate the enzymatic process.

During co-expression of the hypusinating enzymes, the amount of DHPS and DOHH expressed and purified was almost negligible. eIF5A, which usually gives a robust expression and high yield, also showed a significant decrease in its expression when co expressed with the enzymes. One of the reasons for this could have been excess demand of the bacterial chaperones because of recombinant expression of three proteins simultaneously. Expression parameters were optimised to find the appropriate expression conditions. Temperature for expression (18°C, 25°C, 37°C), duration (4 hours for 25°C and 37°C and overnight for 18°C), IPTG induction strength (0.05mM, 0.1mM, 0.5 mM and 1mM IPTG) and bacterial strains (BLR, DE3Star, TOP10F') were tested systematically, but the expression of the enzymes did not show any marked preference for better expression, both when expressed with and without eIF5A.

DOHH is an Iron containing enzyme and this is important for its catalytic activity. It has been shown in the yeast homolog of DOHH, that this catalytically important Fe is also crucial for proper folding of the enzyme. There was a possibility that enough Iron was not available during expression of DOHH and this leads to improper folding during its recombinant expression in bacteria. Iron was supplied (with chelating agent NTA) in the media for bacterial growth and DHPS and DOHH showed a significant improvement in expression. When eIF5A was co-expressed with the hypusinating enzymes in Iron (FeSO₄ with NTA) containing media, there was a dramatic increase in its expression as well. To confirm if the hypusination succeeded, eIF5A will be sent for ESI - Electrospray Ionization where the mass of the full length protein will be analyzed accurately. eIF5A expressed with and without the hypusinating enzymes will be compared to conclude if a modification is present.

After confirming that homogenous amounts of fully hypusinated eIF5A is being formed, its interaction with Exportin 4 and RanGTP to form an export complex will be tested. Conditions will be optimised to attain a large amounts of a pure and stable complex. Following this, the export complex will be taken further for crystallography analysis. This structure is in particular interesting because this will also explain how the presence of hypusine increases the affinity of eIF5A for Exportin 4 in the export complex.

4.4.2 Expression and purification of Smad 3

Smad3 is a transcription factor that is exported by CRM1. The expression and purification of Smad 3 has been established. It expressed very well and can be purified in large amounts with high purity. The next step is to see if it forms a stable complex with Exportin 4 and RanGTP. Following this, the complex will be structurally analysed.

4.4.3 Isolating a stable HMG domain of Sox 2 and SRY for interaction with Exportin 4

Sox 2 and SRY belong to the Sox family of transcription factors. They contain a homeobox DNA binding domain that interacts with the target gene, binds it and controls its expression. These proteins have a relatively simple structural fold and the HMG box has been structurally studied. When the full length protein was expressed in bacterial strains, heavy degradation was observed for both Sox 2 and SRY. The degradation was seen after expression, which indicates that bacterial proteases attack the protein being expressed. Expression conditions and bacterial strains were optimised but did not prove to be sufficient to prevent degradation. The next strategy used to secure full length protein was designed using 2 tags - an N terminal GST tag and a C terminal Histidine tag. The expressed protein was put through two consecutive rounds of purification by pulling out the protein from one terminal at a time. The full length protein secured was pure, but not in enough amount for crystallographic trials. A co-expression trail was also performed to secure the Exportin 4-Sox2 and Exportin4-SRY complexes in vivo. Here the transcription factors with a Histidine tag were co-expressed with Exportin 4 (without a tag). During expression, it was clear that Exportin 4 did not express without the tag and thus this strategy needs further optimisation.

Following this, synthetic genes, which were codon optimised for expression in bacteria, were used to give better yield and reduce degradation. However, even with the synthetic genes, there was little improvement regarding degradation.

The protein structure of the HMG box of Sox 2 and SRY has been analysed. This contains basic residues and in many members of the Sox family has been discovered to contain Nuclear Localisation Signals. This suggests, that the HMG box region could serve as a DNA binding region and as a NLS for Sox 2 and SRY as well. The HMG box was cloned and expressed. The expression and purification for these HMG domains of Sox2 and SRY was established. These proteins express well and can be purified in large amounts. A binding assay will be

performed to validate a direct interaction between the purified HMG domains of Sox2 and SRY with Exportin 4. For an import complex, this complex can not form in the presence of RanGTP.

Once a stable, stoichiometric import complex is formed between the Sox2-HMG domain with Exportin4 and SRY-HMG domain with Exportin 4, the purified complexes can be taken further for crystal trials to elucidate their structure.

5 Appendix

5.1) Figures A1, A2 and A3: Tandem purification of dual tagged Sox 2 to isolate full length protein

Here, a small scale binding assay was performed with 700µl of lysate. For each Sox 2 construct there were 2 tandem purifications performed - once in the order Nickel affinity purification (shown in black print) followed by GST purification (shown in red print). The second tandem purification was done in reverse order. Both these samples were loaded in parallel on the SDS gel. The samples from the different purification steps have been loaded. The different steps include flowthrough after binding to the beads, wash, elution with elution buffer, followed by elution with SDS loading buffer (to remove all protein from the beads). It is clear that the amount of full length protein secured after a tandem purification was minimal and impure. For crystallography trials it is essential to isolate large amounts of protein of high purity. Thus, a new strategy needed to be developed to generate Sox 2 suitable for a further crystallography approach.

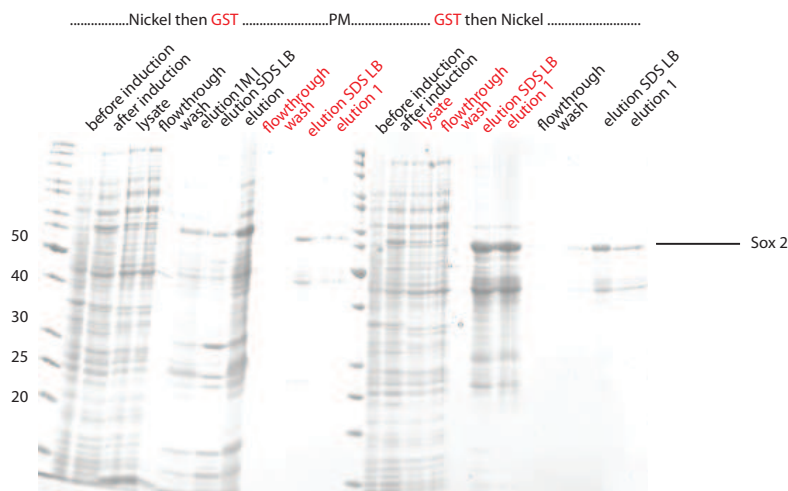


Fig A1
GST-Sox 2 -His 2

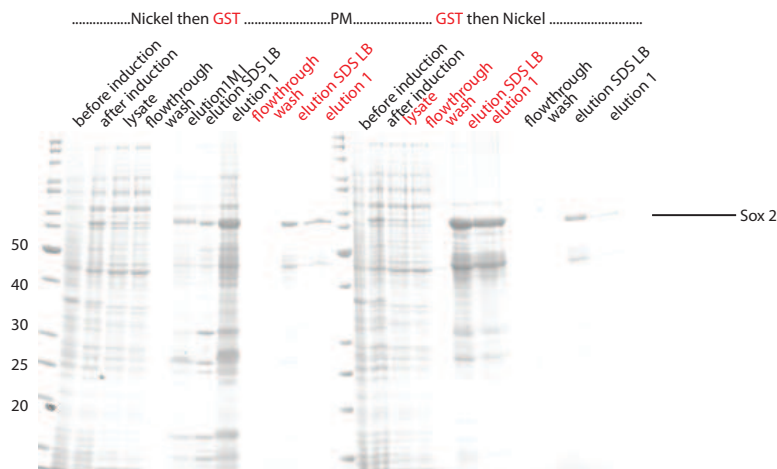


Fig A2
GST-Sox2-His 4

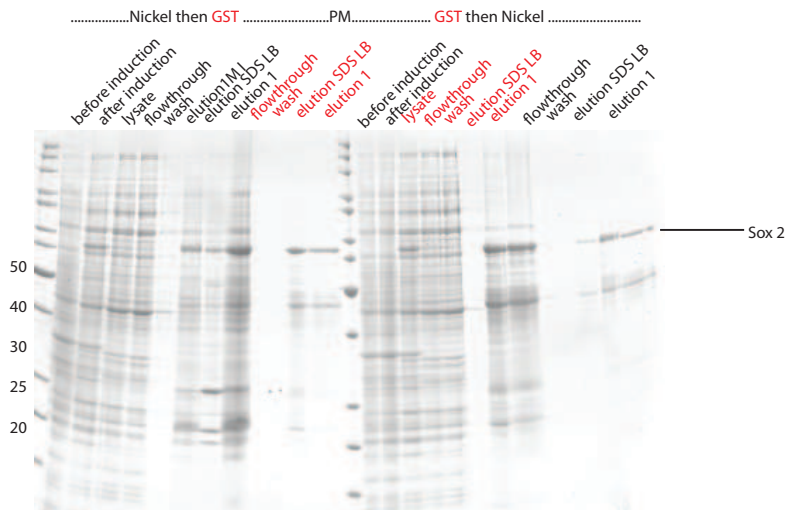


Fig A3
GST-Sox2-His 8

Figure A1,A2 and A3: Tandem purification of Sox 2 using the GST and Histidine tags.

Sox 2 was expressed with a N terminal GST tag and a C terminal Histidine tag. The length of the Histidine tag varied in length – 2,4 and 8 Histidine residues. This was a tandem purification where Sox 2 was first pulled out via the His tag by Ni²⁺ affinity chromatography and for the second round Sox 2 was pulled out via the GST tag using Glutathione Sepharose beads (left side

of the gels). This tandem purification was then performed in the reverse order (loaded on right side of the gel). The GST purification steps are shown in red and the Nickel purification is shown in black. The expression was done in BL21DE3 Star at 25°C with 0.2 mM IPTG. This was performed on a small scale with 700µl lysate. The lysate was in a buffer - 50mM Tris, 500 mM NaCl, 2 mM MgOAc2 pH 7.5. This was taken for purification rounds and in GST elution, the buffer used was 50 mM Tris, 100 mM NaCl, 5 mM MgOAc2, 10 mM DTT and 20 mM Glutathione reduced, pH 7.5) and for Nickel purification, the elution buffer was 50mM Tris, 500 mM Imidazole, 2 mM MgOAc2, 5 mM DTT, pH 7.5. This was elution 1 is the both the purification schemes. Following this the beads were eluted once again with SDS loading buffer to remove all binding protein. (in case of Nickel purification, there were three rounds - elution 1 with elution buffer, elution 2 with 1 M Imidazole and elution 3 with SDS loading buffer).

Bibliography

- Adachi Y, Yanagida M (1989): Higher order chromosome structure is affected by cold-sensitive mutations in a *Schizosaccharomyces pombe* gene *crm1+* which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. *J. Cell Biol.* 108, 1195–1207
- Ader C, Frey S, Maas W, Schmidt HB, Görlich D, Baldus M (2010): Amyloid-like interactions within nucleoporin FG hydrogels. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6281–6285
- Arnautov A, Dasso M (2003): The Ran GTPase regulates kinetochore function. *Dev. Cell* 5, 99–111
- Arts GJ, Kuersten S, Romby P, Ehresmann B, Mattaj IW (1998): The role of exportin-t in selective nuclear export of mature tRNAs. *EMBO J.* 17, 7430–7441
- Avis JM, Clarke PR (1996): Ran, a GTPase involved in nuclear processes: its regulators and effectors. *J. Cell. Sci.* 109 (Pt 10), 2423–2427
- BAHR GF, BEERMANN W (1954): The fine structure of the nuclear membrane in the larval salivary gland and midgut of *Chironomus*. *Exp. Cell Res.* 6, 519–522
- Bayliss R, Littlewood T, Stewart M (2000): Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* 102, 99–108
- Bayliss R, Ribbeck K, Akin D, Kent HM, FELDHERR CM, Görlich D, Stewart M (1999): Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J. Mol. Biol.* 293, 579–593
- Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD (2002): Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* 108, 345–356
- Bischoff FR, Görlich D (1997): RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors. *FEBS Lett.* 419, 249–254
- Bischoff FR, Krebber H, Kempf T, Hermes I, Ponstingl H (1995): Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport. *Proc. Natl. Acad. Sci. U.S.A.* 92, 1749–1753
- Bischoff FR, Ponstingl H (1991): Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* 354, 80–82
- Bischoff FR, Ponstingl H (1991): Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10830–10834
- Bischoff FR, Ponstingl H (1995): Catalysis of guanine nucleotide exchange of Ran by RCC1 and stimulation of hydrolysis of Ran-bound GTP by Ran-GAP1. *Meth. Enzymol.* 257, 135–144
- Bohnsack MT, Regener K, Schwappach B, Saffrich R, Paraskeva E, Hartmann E, Görlich D (2002): Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO J.* 21, 6205–6215

- Brow DA (2002): Allosteric cascade of spliceosome activation. *Annu. Rev. Genet.* 36, 333–360
- Brownawell AM, Macara IG (2002): Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J. Cell Biol.* 156, 53–64
- Calado A, Treichel N, Müller E, Otto A, Kutay U (2002): Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *EMBO J.* 21, 6216–6224
- Caraglia M, Park MH, Wolff EC, Marra M, Abbruzzese A (2011): eIF5A isoforms and cancer: two brothers for two functions? *Amino Acids*
- Chai J, Wu J, Yan N, Massagué J, Pavletich NP, Shi Y (2003): Features of a Smad3 MH1-DNA complex. Roles of water and zinc in DNA binding. *J. Biol. Chem.* 278, 20327–20331
- Cho S, Hoffman DW (2002): Structure of the beta subunit of translation initiation factor 2 from the archaeon *Methanococcus jannaschii*: a representative of the eIF2beta/eIF5 family of proteins. *Biochemistry* 41, 5730–5742
- Conti E, Izaurralde E (2001): Nucleocytoplasmic transport enters the atomic age. *Curr. Opin. Cell Biol.* 13, 310–319
- Cook A, Bono F, Jinek M, Conti E (2007): Structural biology of nucleocytoplasmic transport. *Annu. Rev. Biochem.* 76, 647–671
- Cook AG, Conti E (2010): Nuclear export complexes in the frame. *Curr. Opin. Struct. Biol.* 20, 247–252
- Cronshaw JM, Krutchinsky AN, Zhang W, Chait BT, Matunis MJ (2002): Proteomic analysis of the mammalian nuclear pore complex. *J. Cell Biol.* 158, 915–927
- Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD (1988): The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J. Cell Biol.* 107, 841–849
- Dingwall C, Sharnick SV, Laskey RA (1982): A polypeptide domain that specifies migration of nucleoplasmin into the nucleus. *Cell* 30, 449–458
- FELDHERR CM (1962): The nuclear annuli as pathways for nucleocytoplasmic exchanges. *J. Cell Biol.* 14, 65–72
- Feng W, Benko AL, Lee JH, Stanford DR, Hopper AK (1999): Antagonistic effects of NES and NLS motifs determine *S. cerevisiae* Rna1p subcellular distribution. *J. Cell. Sci.* 112 (Pt 3), 339–347
- Fischer U, Huber J, Boelens WC, Mattaj IW, 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, Mattaj IW (1997): CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90, 1051–1060
- Frey S, Görlich D (2007): A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* 130, 512–523
- Frey S, Görlich D (2009): FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties. *EMBO J.* 28, 2554–2567

- Frey S, Richter RP, Görlich D (2006): FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* 314, 815–817
- Fried H, Kutay U (2003): Nucleocytoplasmic transport: taking an inventory. *Cell. Mol. Life Sci.* 60, 1659–1688
- Fu S, Imai K, Horton P (2011): Prediction of leucine-rich nuclear export signal containing proteins with NESsential. *Nucleic Acids Res.* 39, e111
- Fukuda M, Asano S, Nakamura T, Adachi M, Yoshida M, Yanagida M, Nishida E (1997): CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* 390, 308–311
- Gall JG (1967): Octagonal nuclear pores. *J. Cell Biol.* 32, 391–399
- Gilbert W (1978): Why genes in pieces? *Nature* 271, 501
- Gontan C, Güttler T, Engelen E, Demmers J, Fornerod M, Grosveld FG, Tibboel D, Görlich D, Poot RA, Rottier RJ (2009): Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *J. Cell Biol.* 185, 27–34
- Görlich D, Dabrowski M, Bischoff FR, Kutay U, Bork P, Hartmann E, Prehn S, Izaurralde E (1997): A novel class of RanGTP binding proteins. *J. Cell Biol.* 138, 65–80
- Görlich D, Kutay U (1999): Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* 15, 607–660
- Görlich D, Mattaj IW (1996): Nucleocytoplasmic transport. *Science* 271, 1513–1518
- Görlich D, Seewald MJ, Ribbeck K (2003): Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J.* 22, 1088–1100
- Görlich D, Panté N, Kutay U, Aebi U, Bischoff FR (1996): Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J.* 15, 5584–5594
- Güttler T, 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, 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
- Hastings ML, Krainer AR (2001): Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol.* 13, 302–309
- Heldin C, Moustakas A (2012): Role of Smads in TGF β signaling. *Cell Tissue Res.* 347, 21–36
- Hellmuth K, Lau DM, Bischoff FR, Künzler M, Hurt E, Simos G (1998): Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Mol. Cell. Biol.* 18, 6374–6386
- Hetzer MW, Walther TC, Mattaj IW (2005): Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Annu. Rev. Cell Dev. Biol.* 21, 347–380
- Hillig RC, Renault L, Vetter IR, Drell T, Wittinghofer A, Becker J (1999): The crystal structure of rna1p: a new fold for a GTPase-activating protein. *Mol. Cell* 3, 781–791

- Ho JH, Kallstrom G, Johnson AW (2000): Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *J. Cell Biol.* 151, 1057–1066
- Iovine MK, Watkins JL, Wente SR (1995): The GLFG repetitive region of the nucleoporin Nup116p interacts with Kap95p, an essential yeast nuclear import factor. *J. Cell Biol.* 131, 1699–1713
- Izaurralde E, Kutay U, Kobbe C von, Mattaj IW, 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
- Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW (1995): A cap-binding protein complex mediating U snRNA export. *Nature* 376, 709–712
- Kaffman A, O'Shea EK (1999): Regulation of nuclear localization: a key to a door. *Annu. Rev. Cell Dev. Biol.* 15, 291–339
- Kaffman A, Rank NM, O'Neill EM, Huang LS, O'Shea EK (1998): The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature* 396, 482–486
- Kahle J, Baake M, Doenecke D, Albig W (2005): Subunits of the heterotrimeric transcription factor NF-Y are imported into the nucleus by distinct pathways involving importin beta and importin 13. *Mol. Cell. Biol.* 25, 5339–5354
- Kaiser A (2012): Translational control of eIF5A in various diseases. *Amino Acids* 42, 679–684
- Kalderon D, Richardson WD, Markham AF, Smith AE (1984): Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 311, 33–38
- Kashimada K, Koopman P (2010): Sry: the master switch in mammalian sex determination. *Development* 137, 3921–3930
- Kaur G, Delluc-Clavieres A, Poon IKH, Forwood JK, Glover DJ, Jans DA (2010): Calmodulin-dependent nuclear import of HMG-box family nuclear factors: importance of the role of SRY in sex reversal. *Biochem. J.* 430, 39–48
- Kaur G, Jans DA (2011): Dual nuclear import mechanisms of sex determining factor SRY: intracellular Ca²⁺ as a switch. *FASEB J.* 25, 665–675
- Kelberman D, Castro SCP de, Huang S, Crolla JA, Palmer R, Gregory JW, Taylor D, Cavallo L, Faienza MF, Fischetto R et al. (2008): SOX2 plays a critical role in the pituitary, forebrain, and eye during human embryonic development. *J. Clin. Endocrinol. Metab.* 93, 1865–1873
- Kerscher B, Nzukou E, Kaiser A (2010): Assessment of deoxyhypusine hydroxylase as a putative, novel drug target. *Amino Acids* 38, 471–477
- Kim KK, Hung LW, Yokota H, Kim R, Kim SH (1998): Crystal structures of eukaryotic translation initiation factor 5A from *Methanococcus jannaschii* at 1.8 Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* 95, 10419–10424
- Kudo N, Wolff B, Sekimoto T, Schreiner EP, Yoneda Y, Yanagida M, Horinouchi S, Yoshida M (1998): Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1. *Exp. Cell Res.* 242, 540–547

- Koopman P (1999): Sry and Sox9: mammalian testis-determining genes. *Cell. Mol. Life Sci.* 55, 839–856
- Kuersten S, Ohno M, Mattaj IW (2001): Nucleocytoplasmic transport: Ran, beta and beyond. *Trends Cell Biol.* 11, 497–503
- Kurisaki A, Kose S, Yoneda Y, Heldin CH, Moustakas A (2001): Transforming growth factor-beta induces nuclear import of Smad3 in an importin-beta1 and Ran-dependent manner. *Mol. Biol. Cell* 12, 1079–1091
- Kurisaki A, Kurisaki K, Kowanetz M, Sugino H, Yoneda Y, Heldin C, Moustakas A (2006): The mechanism of nuclear export of Smad3 involves exportin 4 and Ran. *Mol. Cell. Biol.* 26, 1318–1332
- Kutay U, Bischoff FR, Kostka S, Kraft R, Görlich D (1997): Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* 90, 1061–1071
- Kutay U, Güttinger S (2005): Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol.* 15, 121–124
- Kutay U, Lipowsky G, Izaurralde E, Bischoff FR, Schwarzmaier P, Hartmann E, Görlich D (1998): Identification of a tRNA-specific nuclear export receptor. *Mol. Cell* 1, 359–369
- La Cour T, Kierner L, Mølgaard A, Gupta R, Skriver K, Brunak S (2004): Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng. Des. Sel.* 17, 527–536
- Laurino JP, Thompson GM, Pacheco E, Castilho BA (1999): The beta subunit of eukaryotic translation initiation factor 2 binds mRNA through the lysine repeats and a region comprising the C2-C2 motif. *Mol. Cell. Biol.* 19, 173–181
- Leighton F, Poole B, Beaufay H, Baudhuin P, Coffey JW, Fowler S, Duve C de (1968): The large-scale separation of peroxisomes, mitochondria, and lysosomes from the livers of rats injected with triton WR-1339. Improved isolation procedures, automated analysis, biochemical and morphological properties of fractions. *J. Cell Biol.* 37, 482–513
- Lipowsky G, Bischoff FR, Schwarzmaier P, Kraft R, Kostka S, Hartmann E, Kutay U, Görlich D (2000): Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *EMBO J.* 19, 4362–4371
- Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U (2004): Nuclear export of microRNA precursors. *Science* 303, 95–98
- Macara IG (2001): Transport into and out of the nucleus. *Microbiol. Mol. Biol. Rev.* 65, 570-94, table of contents
- Maier B, Tersey SA, Mirmira RG (2010): Hypusine: a new target for therapeutic intervention in diabetic inflammation. *Discov Med* 10, 18–23
- Maquat LE (2004): Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* 5, 89–99
- Matsuura Y, Stewart M (2004): Structural basis for the assembly of a nuclear export complex. *Nature* 432, 872–877

- Melchior F, Paschal B, Evans J, 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, Weber K, Gerke V (1993): A functional homologue of the RNA1 gene product in *Schizosaccharomyces pombe*: purification, biochemical characterization, and identification of a leucine-rich repeat motif. *Mol. Biol. Cell* 4, 569–581
- Mertin S, McDowall SG, Harley VR (1999): The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Res.* 27, 1359–1364
- Mingot J, Bohnsack MT, Jäkle U, Görlich D (2004): Exportin 7 defines a novel general nuclear export pathway. *EMBO J.* 23, 3227–3236
- Mingot JM, Kostka S, Kraft R, Hartmann E, Görlich D (2001): Importin 13: a novel mediator of nuclear import and export. *EMBO J.* 20, 3685–3694
- Mohr D, Frey S, Fischer T, Güttler T, Görlich D (2009): Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J.* 28, 2541–2553
- Monecke T, Güttler T, Neumann P, Dickmanns A, Görlich D, Ficner R (2009): Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* 324, 1087–1091
- Moore MS, Blobel G (1994): A G protein involved in nucleocytoplasmic transport: the role of Ran. *Trends Biochem. Sci.* 19, 211–216
- Murphy EC, Zhurkin VB, Louis JM, Cornilescu G, Clore GM (2001): Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. *J. Mol. Biol.* 312, 481–499
- Nakielnny S, Dreyfuss G (1998): Import and export of the nuclear protein import receptor transportin by a mechanism independent of GTP hydrolysis. *Curr. Biol.* 8, 89–95
- Nemergut ME, Mizzen CA, Stukenberg T, Allis CD, Macara IG (2001): Chromatin docking and exchange activity enhancement of RCC1 by histones H2A and H2B. *Science* 292, 1540–1543
- Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW (2000): PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* 101, 187–198
- Ohtsubo M, Okazaki H, Nishimoto T (1989): The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. *J. Cell Biol.* 109, 1389–1397
- Ossareh-Nazari B, Bachelier F, Dargemont C (1997): Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* 278, 141–144
- Paine PL, FELDHERR CM (1972): Nucleocytoplasmic exchange of macromolecules. *Exp. Cell Res.* 74, 81–98
- Paraskeva E, Izaurralde E, Bischoff FR, Huber J, Kutay U, Hartmann E, Lührmann R, Görlich D (1999): CRM1-mediated recycling of snurportin 1 to the cytoplasm. *J. Cell Biol.* 145, 255–264
- Park JH, Dias CAO, Lee SB, Valentini SR, Sokabe M, Fraser CS, Park MH (2011): Production of active recombinant eIF5A: reconstitution in *E.coli* of eukaryotic

hypusine modification of eIF5A by its coexpression with modifying enzymes. *Protein Eng. Des. Sel.* 24, 301–309

Park MH, Nishimura K, Zanelli CF, Valentini SR (2010): Functional significance of eIF5A and its hypusine modification in eukaryotes. *Amino Acids* 38, 491–500

Partridge JR, Schwartz TU (2009): Crystallographic and biochemical analysis of the Ran-binding zinc finger domain. *J. Mol. Biol.* 391, 375–389

Patel SS, Belmont BJ, Sante JM, Rexach MF (2007): Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* 129, 83–96

Paz EA, Garcia-Huidobro J, Ignatenkos NA (2011): Polyamines in cancer. *Adv Clin Chem* 54, 45–70

Peat TS, Newman J, Waldo GS, Berendzen J, Terwilliger TC (1998): Structure of translation initiation factor 5A from *Pyrobaculum aerophilum* at 1.75 Å resolution. *Structure* 6, 1207–1214

Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI, Hellen CU (2001): Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7029–7036

Price N, Proud C (1994): The guanine nucleotide-exchange factor, eIF-2B. *Biochimie* 76, 748–760

Qin BY, Lam SS, Correia JJ, Lin K (2002): Smad3 allostery links TGF- β receptor kinase activation to transcriptional control. *Genes Dev.* 16, 1950–1963

Reichelt R, Holzenburg A, Buhle EL, Jarnik M, Engel A, Aebersold 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

Reményi A, Lins K, Nissen LJ, Reinbold R, Schöler HR, Wilmanns M (2003): Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* 17, 2048–2059

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

Ribbeck K, Görlich D (2001): Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* 20, 1320–1330

Ribbeck K, Lipowsky G, Kent HM, Stewart M, Görlich D (1998): NTF2 mediates nuclear import of Ran. *EMBO J.* 17, 6587–6598

Richards SA, Lounsbury KM, Carey KL, Macara IG (1996): A nuclear export signal is essential for the cytosolic localization of the Ran binding protein, RanBP1. *J. Cell Biol.* 134, 1157–1168

Robbins J, Dilworth SM, Laskey RA, Dingwall C (1991): Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* 64, 615–623

Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT (2000): The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* 148, 635–651

- Scheffzek K, Klebe C, Fritz-Wolf K, Kabsch W, Wittinghofer A (1995): Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. *Nature* 374, 378–381
- Seewald MJ, Körner C, Wittinghofer A, Vetter IR (2002): RanGAP mediates GTP hydrolysis without an arginine finger. *Nature* 415, 662–666
- Shi Y (2001): Structural insights on Smad function in TGFbeta signaling. *Bioessays* 23, 223–232
- Shi Y, Massagué J (2003): Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* 113, 685–700
- Shibata S, Sasaki M, Miki T, Shimamoto A, Furuichi Y, Katahira J, Yoneda Y (2006): Exportin-5 orthologues are functionally divergent among species. *Nucleic Acids Res.* 34, 4711–4721
- Smith A, Brownawell A, Macara IG (1998): Nuclear import of Ran is mediated by the transport factor NTF2. *Curr. Biol.* 8, 1403–1406
- Smith JM, Koopman PA (2004): The ins and outs of transcriptional control: nucleocytoplasmic shuttling in development and disease. *Trends Genet.* 20, 4–8
- Solsbacher J, Maurer P, Vogel F, Schlenstedt G (2000): Nup2p, a yeast nucleoporin, functions in bidirectional transport of importin alpha. *Mol. Cell. Biol.* 20, 8468–8479
- Stade K, Ford CS, Guthrie C, Weis K (1997): Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* 90, 1041–1050
- Strawn LA, Shen T, Shulga N, Goldfarb DS, Wentz SR (2004): Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat. Cell Biol.* 6, 197–206
- Stüven T, Hartmann E, Görlich D (2003): Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. *EMBO J.* 22, 5928–5940
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007): Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872
- Tong Y, Park I, Hong B, Nedyalkova L, Tempel W, Park H (2009): Crystal structure of human eIF5A1: insight into functional similarity of human eIF5A1 and eIF5A2. *Proteins* 75, 1040–1045
- Thomas F, 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
- Vetter IR, Nowak C, Nishimoto T, Kuhlmann J, Wittinghofer A (1999): Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. *Nature* 398, 39–46
- Walker P, Doenecke D, Kahle J (2009): Importin 13 mediates nuclear import of histone fold-containing chromatin accessibility complex heterodimers. *J. Biol. Chem.* 284, 11652–11662
- Wen W, Meinkoth JL, Tsien RY, Taylor SS (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, 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
- Xiao Z, Liu X, Lodish HF (2000): Importin beta mediates nuclear translocation of Smad 3. *J. Biol. Chem.* 275, 23425–23428
- Xu D, Farmer A, Collett G, Grishin NV, Chook YM (2012): Sequence and structural analyses of nuclear export signals in the NESdb database. *Mol. Biol. Cell* 23, 3677–3693
- Yamada J, Phillips JL, Patel S, Goldfien G, Calestagne-Morelli A, Huang H, Reza R, Acheson J, Krishnan VV, Newsam S et al. (2010): A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol. Cell Proteomics* 9, 2205–2224
- Yao M, Ohsawa A, Kikukawa S, Tanaka I, Kimura M (2003): Crystal structure of hyperthermophilic archaeal initiation factor 5A: a homologue of eukaryotic initiation factor 5A (eIF-5A). *J. Biochem.* 133, 75–81
- Yi R, Qin Y, Macara IG, Cullen BR (2003): Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* 17, 3011–3016
- Yoshida K, Blobel G (2001): The karyopherin Kap142p/Msn5p mediates nuclear import and nuclear export of different cargo proteins. *J. Cell Biol.* 152, 729–740
- Zanelli CF, Valentini SR (2007): Is there a role for eIF5A in translation? *Amino Acids* 33, 351–358
- Zurdo J, Parada P, van den Berg A, Nusspaumer G, Jimenez-Diaz A, Remacha M, Ballesta JP (2000): Assembly of *Saccharomyces cerevisiae* ribosomal stalk: binding of P1 proteins is required for the interaction of P2 proteins. *Biochemistry* 39, 8929–8934

Acknowledgement

I would like to thank my supervisor and members of the department of cellular logistics with whom I gained insight into the process of nuclear trafficking.

I would like to thank my thesis committee members Prof. Dr. Rodnina and Prof. Dr. Ficner for helpful advice and guidance. I would also convey my thanks to the members of my examination board, Prof. Dr. Doenecke, Dr. Hans Dieter Schmitt and Prof. Dr. Thumm. I would especially like to thank Dr. Steffen Burkhardt who has given incredible support during my MSc and PhD studies and I am so grateful for that.

I convey heart felt thanks to Heinz-Jürgen for help with the transfections and wonderful company through these years. I would also like to thank Susanna for help with the cloning and the kitchen staff for perfect maintenance of the lab infrastructure. Thanks also go to Matthias, Bastian, Thomas, Koray, Michael and Haruki for good discussions.

I would like to thank the MSc/PhD IMPRS Molecular Biology Program for giving me an opportunity to get fantastic exposure to work in the top tier research laboratories.

Most importantly, I thank my parents for their constant love and affection. I would also like to thank Manoj who has brought so much happiness to me.

Curriculum Vitae

First Name: Chandini
Last Name: Kadian
Date of Birth: 3rd September 1985
Gender: Female
Nationality: Indian
Email: ckadian@gwdg.de
Telephone: +49 551 201 – 2403
Address: Am Fassberg 11, Dept. Cellular Logistics, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

EDUCATION

2008 - 2012	Dr. rer. nat. in Dept. Cellular Logistics, Max-Planck-Institute for Biophysical Chemistry, Prof. Dr. Dirk Görlich (Integrated MSc-PhD)	Göttingen, Germany
2007- 2009	Master of Science – International Max-Planck-Research School (MSc/PhD Molecular Biology Program), University of Göttingen <ul style="list-style-type: none">Awarded 2007 – 2009 Stipend of the Excellence Foundation of the Max Planck Society	Göttingen, Germany
2003-2006	Bachelor of Science – Delhi University – Sri Venkateswara College, Biochemistry honours <ul style="list-style-type: none">Graduated with a first class degree (with honours)Awarded the Dhirubhai Ambani Scholarship (merit based)	New Delhi, India
2002- 2003	All India Senior School Certificate Examination (Class 12), Central Board of Secondary Education, Delhi Public School, Vasant Kunj, New Delhi <ul style="list-style-type: none">Received Proficiency Certificate for overall performanceGraduated with 91% (Physics, Chemistry, Biology and Math aggregate)	New Delhi, India
2000-2001	All India Secondary School Examination (Class 10), Central Board of Secondary Education, Delhi Public School, Vasant Kunj, New Delhi <ul style="list-style-type: none">Graduated with 98% in Science	New Delhi, India

PUBLICATIONS

Traffic. 2009 Oct;10(10):1543-59.

Endosomal fusion upon SNARE knockdown is maintained by residual SNARE activity and enhanced docking.

Bethani I, Werner A, Kadian C, Geumann U, Jahn R, Rizzoli SO.

Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, 37077 Göttingen, Germany.

Identification of a nuclear export signal in a transcription factor for CRM1 mediated export Chandini Kadian et al.,
2012 - in preparation.

CONFERENCE

Weizmann Institute of Science, Israel, Intracellular Trafficking Conference – May 2010

EXTRA CURRICULAR ACTIVITIES

- Organizer of Horizons in Molecular Biology international symposium 2008-09, 2009-10, 2010-11
- Tamanna Special School – Assistant teacher for arts and crafts to mentally challenged children (2006-07)