

Biochemical characterization of the Nup62·58·54
nucleoporin complex and mutational analysis of the
exportin CRM1

Dissertation

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

Hema Chug

born in

New Delhi, India

Göttingen, September 2013

Members of the thesis committee

Prof. Dr. Dirk Görlich
(Mentor and Referee) Dept. of Cellular Logistics
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Reinhard Lührmann
(Co-referee) Dept. of Cellular Biochemistry
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Peter Rehling Dept. of Biochemistry II
Georg August University Göttingen
Göttingen, Germany

Additional members of the Examination board

Prof. Dr. Holger Stark Research Group 3D Electron Cryo-Microscopy
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Heinz Neumann Research Group Applied Synthetic Biology
Georg August University Göttingen
Göttingen, Germany

Prof. Dr. Henning Urlaub Research Group Bioanalytical Mass Spectrometry
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Affidavit

I hereby declare that this dissertation "**Biochemical characterization of the Nup62·58·54 nucleoporin complex and mutational analysis of the exportin CRM1**" has been written independently with no other aids or sources than quoted.

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

Hema Chug

September 2013

Göttingen, Germany

Dedicated to my mother,

“The most exciting phrase to hear in science, the one that heralds the most discoveries, is not “Eureka!” (I found it!) but “That’s funny....”

---Isaac Asimov

Table of Contents

1. ABSTRACT	11
2. INTRODUCTION.....	13
2.1. NUCLEUS: A EUKARYOTIC HALLMARK	13
2.2. THE NUCLEAR PORE COMPLEX (NPC)	13
2.3. MECHANISM OF NUCLEO-CYTOPLASMIC EXCHANGE.....	14
2.3.1. NPC's permeability barrier.....	15
2.3.2. The RanGTP gradient.....	15
2.3.3. Nuclear transport receptors (NTRs)	16
2.4. NPC ARCHITECTURE.....	17
2.4.1. Structural organization of the NPC as observed in EM.....	18
2.4.2. Classification of Nups based on their function in the NPC.....	18
2.4.2.1. Transmembrane (TM) nucleoporins.....	18
2.4.2.2. Scaffold nucleoporins	20
2.4.2.3. Permeability barrier forming nucleoporins	20
2.4.3. Nucleoporin subcomplexes as the NPC building blocks.....	20
2.4.3.1. NPC disassembly and reassembly from nucleoporin subcomplexes.....	22
2.4.3.2. Nup107-160 complex (Y-complex)	22
2.4.3.3. Nup93-Nup205 complex.....	23
2.4.3.4. Nup214-88-62 complex.....	25
2.5. VARIOUS APPROACHES TO STUDY NUCLEOPORIN ARRANGEMENT IN THE NPC	25
2.5.1.1. Nucleoporins stoichiometry.....	25
2.5.1.2. High-resolution microscopy	26
2.5.1.3. X-ray crystallography.....	26
2.6. NUP62·58·54 SUBCOMPLEX	27
2.7. THE EXPORTIN CRM1	29
2.8. ABOUT THIS WORK	30
3. RESULTS	32
3.1. BIOCHEMICAL AND STRUCTURAL INVESTIGATION OF XLNUP62·58·54 COMPLEX.....	32
3.1.1. Coiled-coil domains in xINup62·58·54·93 complex.....	32
3.1.2. Δ FG-xINup62·58·54 complex.....	34
3.1.3. Co-expression with interaction partners is required for Δ FG-xINup54 solubility	36
3.1.4. xINup58 and xINup62 also engage in a direct interaction.....	38

3.1.5. A stable Δ FG-Nup62·58·54 complex can be purified under native conditions from <i>E. coli</i> after co-expression	40
3.1.6. Δ FG-xlNup62·58·54 is an elongated complex with 1:1:1 stoichiometry.....	41
3.1.7. Nup93 interaction with Δ FG-xlNup62·58·54 complex.....	43
3.1.8. xlNup93 N-terminus has weak affinity for Δ FG-xlNup62·58·54 complex.....	44
3.1.9. Crystallization attempts for Δ FG-xlNup62·58·54 complex.....	46
3.1.10. Nanobodies (Nbs) against Δ FG-xlNup62·58·54 complex.....	47
3.1.11. Δ FG-xlNup62·58·54 complex recognition by the nanobodies	48
3.1.12. Most nanobodies recognize Nup54 α/β domain in the Δ FG-xlNup62·58·54 complex	50
3.1.13. Nanobody 15 recognizes the trimeric coiled-coil interface of Δ FG-xlNup62·58·54 complex	51
3.1.14. Crystallization of Nanobody conjugated Δ FG-xlNup62·58·54 complex.....	52
3.1.15. Attempts to improve crystal diffraction	56
3.1.16. Limited proteolysis of Δ FG-xlNup62·58·54 complex.....	57
3.1.17. In-situ limited proteolysis.....	60
3.1.18. Cytosolic expression of nanobodies and Nup93n-Nb15 fusion	61
3.2. MUTATIONAL ANALYSIS OF THE EXPORTIN CRM1.....	65
3.2.1. Strong evolutionary conservation of 20 residues in CRM1	65
3.2.2. Analysis of CRM1 point mutants.....	65
3.2.3. CRM1 mutants stabilized in high affinity, cargo-binding state	68
3.2.4. CRM1 mutants stabilized in low affinity, cargo-free state	70
3.2.5. CRM1 mutants with a blocked hydrophobic cleft	71
4. DISCUSSION	76
4.1.1. Insights into the molecular behavior of the Nup62·58·54 complex	77
4.1.1.1. A clear 1:1:1 stoichiometry of the xlNup62·58·54 complex	77
4.1.1.2. Oligomerization of Nup62·58·54 complex- real or artifacts?	78
4.1.1.3. Interaction amongst the subunits of Nup62·58·54 complex	79
4.1.2. Nup62·58·54 complex anchorage to the NPC.....	79
4.1.3. Nanobodies as tools for co-crystallization of Nup62·58·54 complex	81
4.1.3.1. Nanobody mediated crystallization of Nup62·58·54 complex.....	82
4.1.3.2. Exploiting nanobodies for stabilizing Nup93 interaction with Nup62·58·54 complex	82
4.1.4. Mechanism of cargo-RanGTP cooperativity on CRM1	83
4.1.5. Various conformational states of CRM1 mutants.....	85

4.1.6. <i>Inherent structural flexibility forms the basis of NPC's architectural and functional machinery</i>	87
5. MATERIALS AND METHODS	89
5.1. MATERIALS.....	89
5.2. METHODS.....	93
5.2.1. <i>Standard techniques in Molecular Biology</i>	93
5.2.2. <i>General protein expression and purification</i>	96
5.2.3. <i>Specific purification protocols</i>	101
5.2.3.1. xLNup62·58·54 complex	101
5.2.3.2. xLNup62·58·54·93N complex.....	102
5.2.3.3. Nanobody expression and purification.....	103
5.3. NANOBODY CONJUGATION TO XLNUP62·58·54 COMPLEX	103
5.4. LIMITED PROTEOLYSIS OF XLNUP62·58·54 COMPLEX	104
5.5. MOLECULAR WEIGHT DETERMINATION BY STATIC LIGHT SCATTERING (SLS)	104
5.6. POLYDISPERSITY ANALYSIS USING DYNAMIC LIGHT SCATTERING (DLS).....	105
5.7. MASS SPECTROMETRY	105
5.8. PROTEIN CRYSTALLIZATION	105
5.8.1. <i>Sparse matrix screening with 96-well MRC sitting drop plates</i>	106
5.8.2. <i>Grid screening with 96-well MRC sitting drop plates</i>	106
5.8.3. <i>24-well Linbro hanging drop plates</i>	106
5.8.4. <i>Additive screen set up</i>	106
5.8.5. <i>Flash cooling in liquid nitrogen and data collection</i>	107
5.8.6. <i>In-situ proteolysis</i>	107
5.9. BINDING ASSAYS FOR CRM1 MUTANTS.....	107
5.10. LIVE EXPORT ASSAYS IN HELA NUCLEI	108
6. BIBLIOGRAPHY	109
7. ABBREVIATIONS	115
8. ACKNOWLEDGEMENTS	117
9. CURRICULUM VITAE	119

List of figures

Figure 1.1: The Nuclear pore complex (NPC).....	14
Figure 1.2: Overview of the nuclear import-export cycles and the RanGTPase system	16
Figure 1.3: Structural organization of the NPC.....	19
Figure 1.4: NPC subcomplexes	24
Figure 1.5: Structure of the CRM1-RanGTP-SPN1 export complex and unliganded CRM1.....	30
Figure 3.1: Coiled-coil domains in Nup62·58·54·93 complex.....	33
Figure 3.2: Secondary structure prediction for <i>Xenopus laevis</i> Δ FG-Nup62·58·54 complex.....	35
Figure 3.3: Solubility of Δ FG-xlNup62·58·54 complex subunits	36
Figure 3.4: Co-expression with interaction partners makes Nup54 soluble	38
Figure 3.5: Direct interaction between Nup62 and Nup58	39
Figure 3.6: Purification of Δ FG-xlNup62·58·54 complex	41
Figure 3.7: Stoichiometry of Δ FG-xlNup62·58·54 complex determined by MALS	42
Figure 3.8: Mapping N-terminus of xlNup93 for interaction with Δ FG-xlNup62·58·54 complex.....	44
Figure 3.9: xlNup93N-terminus has weak affinity for Δ FG-xlNup62·58·54 complex.....	45
Figure 3.10: Affinity measurement for Nup93 N-terminus interaction with Nup62·58·54 complex	46
Figure 3.11: Nanobodies against Δ FG-xlNup62·58·54 complex.....	48
Figure 3.12: All nanobodies bind Δ FG-xlNup62·58·54 complex with high affinity.....	49
Figure 3.13: Most nanobodies recognize xlNup54 α/β domain in the Δ FG-xlNup62·58·54 complex.....	50
Figure 3.14: Nb15 exclusively binds trimeric coiled-coil interface in Δ FG-xlNup62·58·54 complex	52
Figure 3.15: Scheme for large-scale purification of nanobody containing Δ FG-xlNup62·58·54 complex.....	53
Figure 3.16: Various nanobody - Δ FG-xlNup62·58·54 complexes screened in crystallization trials.....	54
Figure 3.17: Rod-shaped crystals of Nb15- Nup62·58·54 complex.....	57
Figure 3.18: Limited proteolysis of Δ FG-xl Nup62·58·54 complex with chymotrypsin	59
Figure 3.19: Gel-filtration of Nb15- Δ FG-Nup62·58·54 complex after limited proteolysis with chymotrypsin	60
Figure 3.20 In-situ limited proteolysis of Nb15 Δ FG-xlNup62·58·54 complex	61
Figure 3.21: Bacterial expression of disulfide-bond containing nanobodies.....	62
Figure 3.21: Nanobody 15 fusion to Nup93n (1-95) and optimization of spacer lengths	63
Figure 4.1: Sequence analysis of CRM1 from diverse eukaryotic species.....	66
Figure 4.2: Distribution of evolutionary conserved residues on mmCRM1	67
Figure 4.3: Cargo-binding of CRM1 mutants in the absence and presence of RanGTP.....	68
Figure 4.4: RanGTP binding of CRM1 mutants in the presence or absence of cargo	70
Figure 4.5: Comparison of Super-PKI and PKI binding of low affinity CRM1 mutants.....	71
Figure 4.6: Export assay in HeLa nuclei for high affinity, RanGTP independent CRM1 mutants	72
Figure 4.7: Export assay in HeLa nuclei for high affinity, RanGTP dependent CRM1 mutants	73
Figure 4.8: Export assay in HeLa nuclei for low affinity CRM1 mutants.....	73
Figure 4.9: Export assay in HeLa nuclei for CRM1 mutants with blocked hydrophobic cleft.....	74

List of tables

<i>Table 1: Domain architecture of different nucleoporins in the NPC</i>	<i>21</i>
<i>Table 2: Table showing properties of xINup62-58-54 complex as determined by MALS analysis.....</i>	<i>43</i>
<i>Table 3 Table summarising the properties of CRM1 mutants.....</i>	<i>75</i>

1. Abstract

Nuclear pore complexes (NPCs) are giant molecular machines embedded in the double membrane that surrounds the eukaryotic nucleus. Deciphering the molecular structure of the NPCs is critical to our understanding of both cellular architecture and the mechanism of nucleocytoplasmic transport. In less than a decade, atomic level structures of many nucleoporins (Nups) have been solved and the molecular picture of the NPC is becoming increasingly clearer. Nup62·58·54 subcomplex is a nucleoporin subcomplex in the NPC's central channel, the molecular structure of which, is not known so far. At a sequence level, the N-terminal half of all subunits in the complex contain intrinsically disordered phenylalanine-glycine (FG) repeat-motifs. The C-terminal half is structured into coiled-coil domains that engage in tight protein-protein interactions to hold the complex together and to anchor it to the NPC scaffold. In this project, I aimed to elucidate the molecular structure of the *X.laevis* ΔFG-Nup62·58·54 complex by X-ray crystallography. Poor solubility of some coiled-coil domain containing Nups, such as Nup54, in bacterial expression has been a challenge towards obtaining large amounts of nucleoporins required for crystallographic analysis. We established a system where all three subunits of xΔFGNup62·58·54 were co-expressed together in *E.coli*. The soluble amount of Nup54 was greatly enhanced, when co-expressed together with its interaction partners. Biochemical analysis of the recombinant ΔFG-Nup62·58·54 complex revealed a monomeric, non-spherical complex with a clear 1:1:1 subunit stoichiometry, which had been unclear so far. Also, in contrast to previously published reports, the ΔFG-Nup62·58·54 complex did not show any tendency to form protein-concentration dependent higher-order oligomers. However, extensive crystallization trials failed to crystallize ΔFG-Nup62·58·54 complex, probably due to inherent flexibility in its structure. Therefore, we generated and screened a series of single-domain antibodies (nanobodies, Nbs) against ΔFG-Nup62·58·54 complex and identified a few that stabilized the trimeric complex but did not recognize any of the individual subunits or dimeric interactions. Crystallization trials of ΔFG-xlNup62·58·54 complex conjugated to one such nanobody, Nb15, yielded rod-shaped crystals, which are currently limited in their diffraction potential to ~7.5Å. Thus, nanobody conjugation of the ΔFG-xlNup62·58·54 complex aided in its crystallization and can be used as a general approach to stabilize NPC subcomplexes. Furthermore, a coiled-coil interaction between xlNup93 N-terminus and xlNup62·58·54 complex is thought to recruit the Nup62·58·54 complex to the NPCs. This interaction in our experiments, however, turned out to be surprisingly weak. This now predicts additional interaction sites for a high avidity anchorage of the Nup62·58·54 complex to the NPC scaffold.

1. Abstract

In a parallel project, we aimed to obtain insights into exportin CRM1's conformational states by analyzing its sequence conservation in a large number of diverse species, starting from lower eukaryotes such as *Trichomonas* till humans. Individual point mutations in 21 absolutely conserved CRM1 residues produced interesting phenotypes with either greatly enhanced or weakened cargo binding, suggesting that many of these evolutionary conserved residues are critical for stabilizing distinct conformational states of cargo-free or cargo-bound CRM1. These mutants will be exploited further in structural studies that explore the conformation space available to CRM1.

2. Introduction

2.1. Nucleus: A eukaryotic hallmark

Evolution of eukaryotic life from single-celled prokaryotic ancestors involved the development of a complex endomembrane system leading to compartmentalization of cellular processes. The defining feature of eukaryotes (*eu* meaning “true” and *karyo* meaning “kernel” or “seed”) is the presence of nucleus, which is absent in all prokaryotes. The nucleus is enclosed by a double membrane structure called the outer and the inner nuclear membrane (ONM and INM respectively) with ONM being continuous with the Endoplasmic reticulum (ER).

The containment of genetic information in the nucleus gave eukaryotes several advantages over their prokaryotic ancestors for a complete regulation of cellular functions that paved way for their differentiation into multicellular organisms. First among these advantages was the protection of DNA from mechanical shearing or highly reactive free oxygen radicals in the cytoplasm, thus imparting enhanced genetic stability. Second advantage was the increase in metabolic efficiency caused by increased local concentration of DNA replication, repair, transcription and RNA processing machinery inside the nucleus, and its separation from protein and fatty acid synthesis in the cytosol. This separation of translational and transcription machinery also added a layer of functional regulation by controlling the spatial and temporal access of specific activity factors in the nucleus and cytosol. Third advantage was the advent of introns and alternative splicing thereby compounding eukaryotic genomic diversity as various protein isoforms could arise from a single mRNA transcript. Exclusion of ribosomes from nucleus secured that unspliced mRNAs are not prematurely translated since only mature mRNAs are exported out of the nucleus (Cavalier-Smith, 1988; Görlich and Kutay, 1999).

2.2. The Nuclear pore complex (NPC)

The presence of nuclear pore complexes in the nuclear envelope allows exchange of macromolecules such as RNA and proteins between nuclear and cytoplasmic compartments (Feldherr, 1962). NPCs provide aqueous channels through the nuclear envelope (CALLAN et al., 1949) (BAHR and BEERMANN, 1954), formed at the sites of fusion between the outer and inner nuclear membranes (Watson, 1954) (*Figure 1.1a*).

Early electron microscopic (EM) observations of the nuclear membrane in amphibian oocytes revealed NPCs to be a cylindrical structure with octahedral symmetry (GALL, 1954; Gall, 1967) (*Figure 1.1b*). Metazoan NPCs have an estimated MW of ~125MDa (Reichelt et al., 1990), whereas the yeast counterpart is relatively smaller with a MW of ~66MDa (Rout and Blobel,

2. Introduction

1993). The average inner diameter of NPC is ~50 nm and the outer diameter ranges from 85-125 nm (Beck et al., 2004). Despite the differences in molecular dimensions of these giant assemblies, EM reconstructions of NPCs across evolutionary divergent species have established that the overall architecture is conserved (Grossman et al., 2012). The number and density of the NPCs in the nuclear envelope varies in different organisms depending on the nuclear volume with a modest ~200 in yeast (Rout and Blobel, 1993), to an intermediate ~2000-5000 in human cells and a high $\sim 5 \times 10^7$ NPC number in mature *Xenopus* oocytes (Cordes et al., 1995). Chemically, NPCs are entirely proteinaceous assemblies and compared to their large size, are comprised of relatively small number of distinct subunits i.e. ~30 (called nucleoporins or Nups), which are present, however, in multiple copies of eight (Rout et al., 2000; Cronshaw et al., 2002). The nucleoporins studied across species including mammals, amphibians and lower eukaryotes are structurally homologous thus accounting for strong evolutionary conservation of overall NPC architecture and function (Yang et al., 1998).

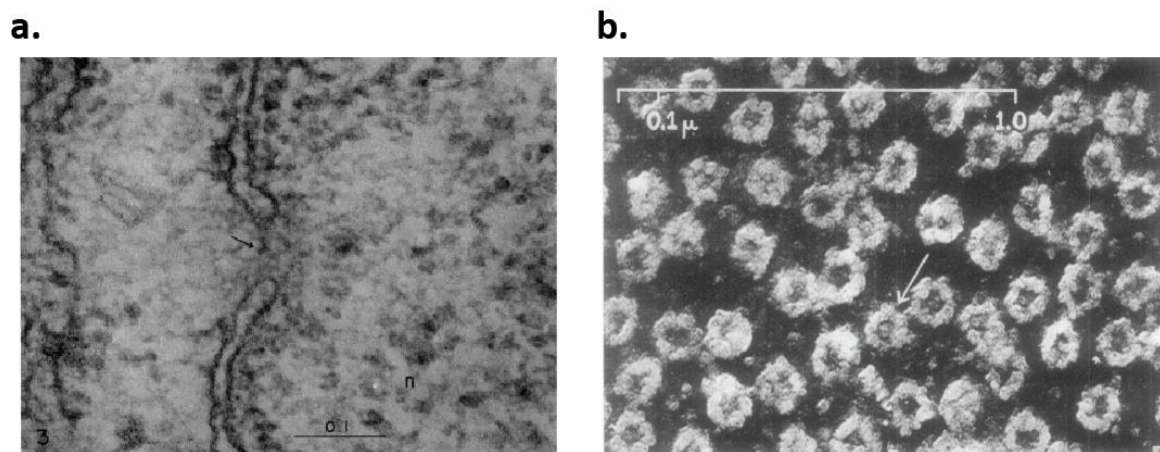


Figure 1.1: The Nuclear pore complex (NPC)

a. Early EM images showing NPCs at the junction of INM and ONM fusion (Watson, 1959). **b.** Surface view of amphibian nuclear membrane in EM showing octahedral NPCs (Gall, 1954).

2.3. Mechanism of nucleo-cytoplasmic exchange

NPCs are not mere aqueous channels in the NE but in fact, are the 'only' portals of exchange between the nucleus and the cytoplasm. The regulated macromolecular exchange at the NPC depends upon phenylalanine-glycine motif (FG-motif or repeat) containing nucleoporins that line the central NPC channel and form the 'permeability barrier' against free diffusion of macromolecules. Translocation across the NPC is a rapid, regulated process involving flow of nearly 100MDa mass per NPC per second that corresponds to ~ 1000 translocation events $\cdot \text{NPC}^{-1} \cdot \text{s}^{-1}$ (Ribbeck and Görlich, 2001). Small molecules with MW ≤ 30 kDa (or diameter ≤ 5 nm) can pass NPCs by simple diffusion; however, as the size of the molecule increases translocation

2. Introduction

across the NPC is restricted (Mohr et al., 2009). For crossing this diffusion barrier formed by NPC, macromolecules rely on binding to and subsequent facilitated translocation by dedicated shuttling factors called nuclear transport receptors (NTRs) (Görlich et al., 1994). The detailed mechanism and nature of macromolecular translocation through NPC is reviewed in (Görlich and Kutay, 1999) (Hetzer and Wente, 2009) and is briefly described below.

2.3.1. NPC's permeability barrier

The central transport channel of the NPC is filled with FG-repeats of FG-Nups that are crucial components for the formation of NPC's permeability barrier. The FG-repeat domains are intrinsically disordered regions (Denning et al., 2003) and contain repeat motifs of GLFG, FxFG or FG (where F is phenylalanine, G is glycine, L is leucine, x is any other residue) residues separated by hydrophilic spacers. About one-third of the total Nups in a NPC have FG-domains that provide binding sites for nuclear transport receptors (Iovine et al., 1995).

Due to intrinsic disorder in the FG-repeats, direct structural information about the arrangement of these motifs in the NPC central channel is not available. Based on the nature of interaction between the FG-repeats and the translocation mechanism, several models have been proposed to explain the molecular arrangement of FG-repeats in the central transport channel.

2.3.2. The RanGTP gradient

A concentration gradient of RanGTP across the nuclear envelope provides metabolic energy for driving active NPC transport (Görlich et al., 1996) (Izaurralde et al., 1997). In the nucleus, Ran exists predominantly in a GTP bound form whereas in the cytoplasm it is exclusively present as RanGDP. The proteins playing key role in maintaining this Ran gradient across the two compartments include Ran (Melchior et al., 1993; Moore and Blobel, 1993), Ran shuttling factor NTF2 (Moore and Blobel, 1994), Ran nucleotide exchange factor RCC1 (Bischoff and Ponstingl, 1991), RanGTP hydrolyzing protein RanGAP (Bischoff et al., 1994) and Ran binding proteins RanBP1 (Bischoff et al., 1995) and RanBP2 (Wu et al., 1995).

In the nucleus, nucleotide exchange on Ran from GDP to GTP is catalyzed by guanine nucleotide exchange factor RCC1. RCC1 is present exclusively in the nucleus due to its association with the chromatin (Ohtsubo et al., 1989). As a result, Ran is maintained in a GTP bound form inside the nucleus. On the other hand, RanGTPase-activating protein, RanGAP has an exclusive cytoplasmic localization. It acts in association with Ran binding proteins (RanBP1 or RanBP2) to hydrolyze RanGTP to RanGDP thereby forcing Ran to be present in a GDP bound form in the cytosol. NTF2 binds RanGDP in the cytoplasm and brings it to the nucleus (Ribbeck et al., 1998), thus maintaining a steady state level of Ran in the nucleus. The concentration gradient of RanGTP or

2. Introduction

RanGDP across the NE, thus formed, is used to provide directionality to the NPC transport (Figure 1.2). NTRs take advantage of this gradient as discussed below.

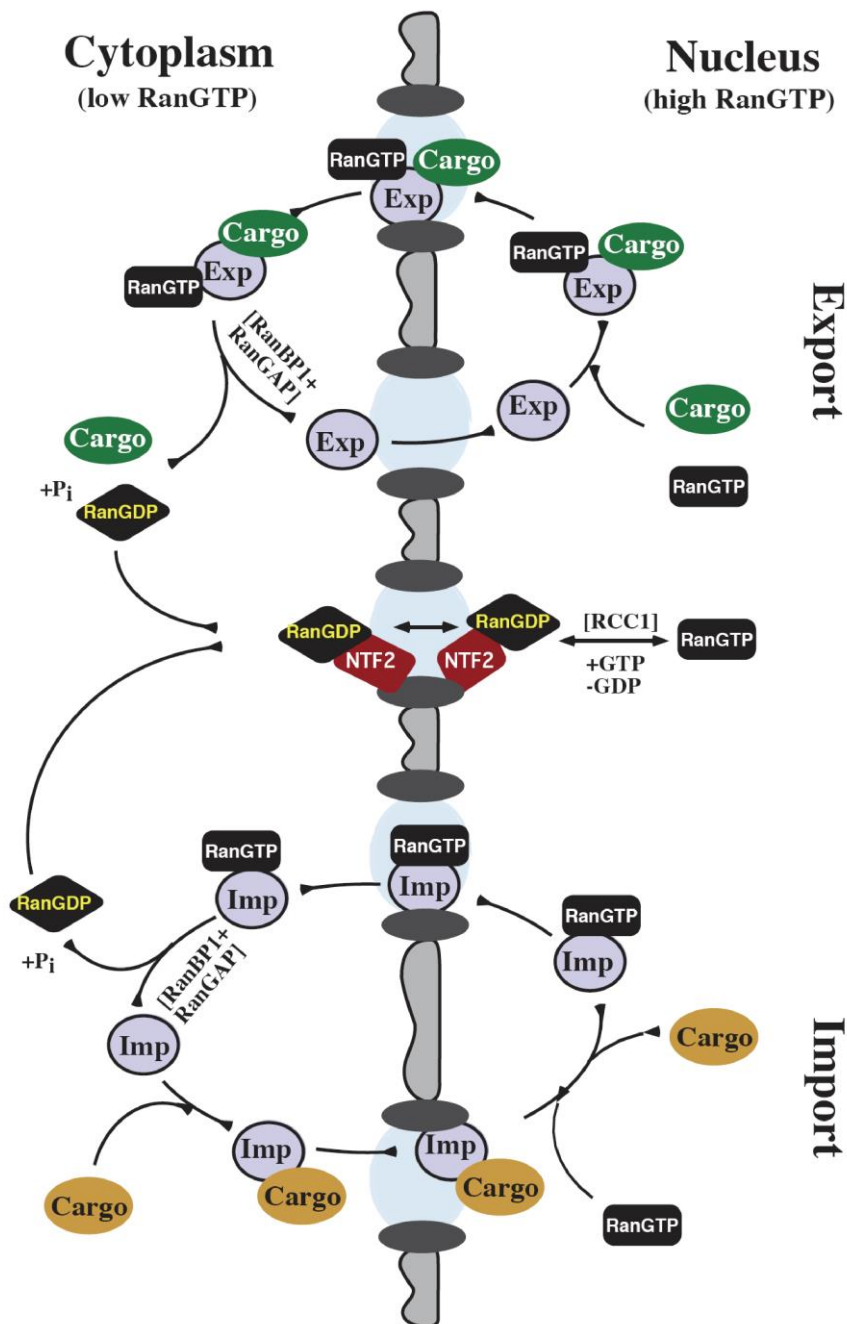


Figure 1.2: Overview of the nuclear import-export cycles and the RanGTPase system

Adapted from (Görlich and Kutay, 1999). See text for details.

2.3.3. Nuclear transport receptors (NTRs)

Nuclear transport receptors (NTRs) are crucial for facilitated transport of all macromolecules ≥ 40 kDa in size such as majority of transcription factors, histones, RNPs etc. across the nuclear envelope. Reviewed in (Görlich and Kutay, 1999) (Güttler and Görlich, 2011).

2. Introduction

The best-characterized NTRs belong to the Importin- β superfamily that contains large, superhelical, α -solenoidal receptors, named after the first receptor identified (Görlich et al., 1995) (Fornerod et al., 1997a). Structurally NTRs are characterized by stacked arrangement of multiple HEAT repeat elements, where a single HEAT repeat element is made up by two α -helices, 10-20 residues in length, in antiparallel orientation connected by a short linker. NTRs bind cargoes in the either cellular compartment through signal recognition mechanisms (Moore and Blobel, 1992) (Görlich et al., 1994) and have multiple hydrophobic binding pockets for docking to FG-repeats in the NPC (Iovine et al., 1995) (Radu et al., 1995). Cargo release at either side of the NPC is facilitated by RanGTP binding or hydrolysis (Rexach and Blobel, 1995) (Görlich et al., 1996) and NTRs recycle back to their original compartment for next round of cargo binding, NPC translocation and cargo release (*Figure 1.2*).

Based on the direction of cargo transport, NTRs can be subdivided into importins or exportins (some NTRs work in both directions). Some examples of importins include imp β (imports wide range of proteins including ribosomal proteins, HIV Rev, “classical” NLSs with imp α as adaptor) (Görlich et al., 1995) (Chi et al., 1995), transportin (imports hnRNPs) whereas CRM1 (exports wide variety of Leucine-rich NES containing cargoes) (Fornerod et al., 1997a), Exportin-6 (exports actin) (Stuven et al., 2003), CAS (exports Imp α) (Kutay et al., 1997), Exportin-t (exports mature tRNAs) (Arts et al., 1998) etc. are well known exportins. Importins bind their cargoes in the cytoplasm at low RanGTP levels and translocate across NPCs as importin-cargo complexes. At the nuclear side of NPC, competitive binding by RanGTP displaces cargo and importins return back to the cytoplasm across the NPC as importin-RanGTP complexes. At the cytoplasmic side, GTP in the importin-RanGTP complex is hydrolyzed, thereby dissociating the complex making importin freely available for the next round of cargo transport (Rexach and Blobel, 1995). The converse is true for exportins, which recognize nuclear export signals (NESs) to bind their cargoes in the nucleus in the presence of RanGTP, forming a ternary export complex that translocates across the NPC. RanGTP hydrolysis in the cytoplasm dissociates the export complex and exportins are recycled back into the nucleus (Kutay et al., 1997; Fornerod et al., 1997a).

2.4. NPC architecture

Modularity in the NPC architecture allows us to classify its components in at least three different possible ways. Accordingly, NPC structure can be subdivided based on a) structural modules visible in the EM images, b) functions of different nucleoporins in the NPC structure and c) biochemically stable nucleoporin subcomplexes that occur as soluble entities during NPC's mitotic disassembly and indicate stable nucleoporin interactions in the NPC.

2. Introduction

2.4.1. Structural organization of the NPC as observed in EM

A large amount of our current understanding of the NPC architecture in different organisms has been gathered from electro microscopic (EM) observations (Akey and Radermacher, 1993) (Goldberg and Allen, 1993). Using cryo-electron tomography (cryo-ET) techniques, high-resolution NPC structures from the nuclei of different organisms including *Saccharomyces*, *Dictyostelium*, *Xenopus* have been obtained (Beck et al., 2004) (Frenkiel-Krispin et al., 2010). Recently, a 6.6nm resolution structure of the human NPC was determined by cryo-ET in intact human fibroblast cells, without applying any purification steps, to maintain 'close-to-life' state of the NPC (Maimon et al., 2012) where critical details of sponge-like NPC structure have emerged. Traditionally, based on its visual inspection in the EM, NPC has been divided into a 'core NPC scaffold' and 'peripheral components' attached at either sides of the scaffold, called as the cytoplasmic filaments and the nuclear basket (*Figure 1.3*).

The NPC core is found to be a tripartite ring-structure consisting of three parallel ring-like structures viz. cytoplasmic ring, spoke ring and the nuclear ring along the plane of the nuclear envelope (Akey and Radermacher, 1993) (Goldberg and Allen, 1993). Along the axis perpendicular to the nuclear envelope, the NPC can be subdivided into eight radial arms or spokes that impart it with octagonal symmetry (Goldberg and Allen, 1995). Eight cytoplasmic filaments radiate from the cytoplasmic ring into the cytoplasm. Filaments extending from the nuclear ring into the nucleoplasm join to form a ring at the distal end called the nuclear basket (Beck et al., 2004). The translocation channel at the center of NPC is filled with a meshwork of FG-domains of nucleoporins and hence, does not show any distinct structural features (Beck et al., 2004). The EM technique is powerful for studying overall structure of the NPC but the current limitation in molecular resolution at 6nm does not allow resolution of nucleoporin locations, subcomplex boundaries or their relative orientation with respect to each other.

2.4.2. Classification of Nups based on their function in the NPC

Based on their structure, amino-acid sequence motifs and NPC location, nucleoporins can be classified as following (also see Table 1).

2.4.2.1. Transmembrane (TM) nucleoporins

Transmembrane Nups have large membrane-spanning domains that anchor the NPC to the nuclear envelope. In vertebrates, three nucleoporins namely, Ndc1, gp210 (Pom152 in yeast) and Pom121 (unique to vertebrates) with trans membrane α -helices have been identified whereas in yeast, scPom34, scNdc1 and scPom152 contain TM helices. The molecular structures of these Nups are not available and also the details about their interactions with other Nups in the NPC are not clearly understood. However, several studies have reported essential roles

2. Introduction

played by TM Nups during initial phases of NPC assembly and disassembly (Stavru et al., 2006b) (Stavru et al., 2006a).

a.

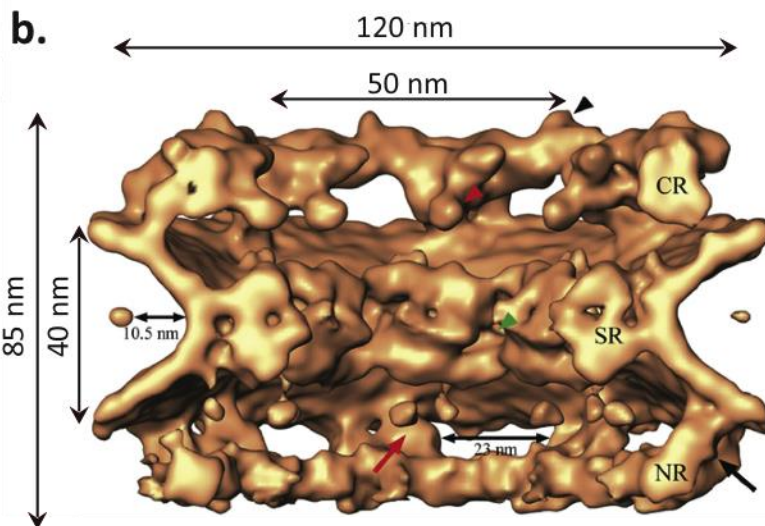
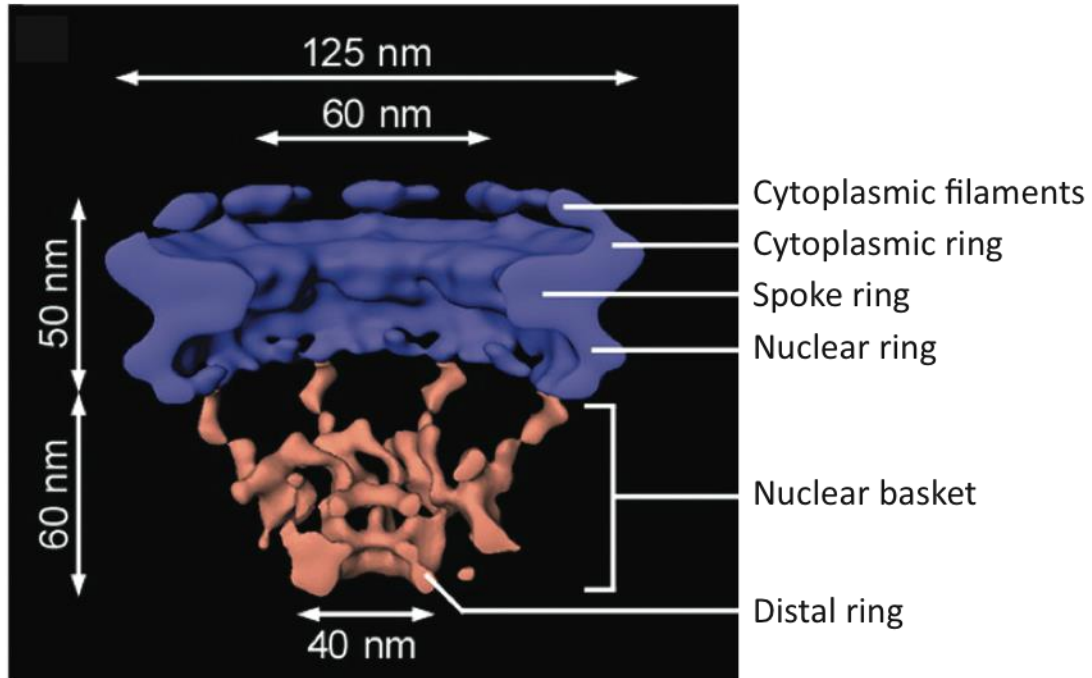


Figure 1.3: Structural organization of the NPC

Cryo-ET reconstructions of **a)** *Dictyostelium discoideum* NPC at 9nm resolution (Beck *et al*, 2004). **b)** Human NPC at 6nm resolution (Maimon *et al*, 2012). The canonical features of the NPC can be observed in the upper structure whereas details of tripartite ring-like NPC structures are easily visible in the bottom structure. CR- Cytoplasmic ring, SR- Spoke ring, NR- Nuclear ring. See text for details.

2. Introduction

2.4.2.2. Scaffold nucleoporins

About one-half of the total nucleoporins do not contribute directly to the permeability barrier and take part in forming the NPC's structural scaffold. These Nups are termed as scaffold Nups. These architectural Nups are symmetrically localized around the two-fold symmetry axis in the plane of the NE and in an eight-fold rotational symmetry along the central transport channel in the NPC. They form the essential skeleton of the NPC by stabilizing the curvature of the pore membrane and have been reported to have extremely long residence times in the NPC spanning the entire cell cycle (Rabut et al., 2004). Alpha-solenoids and beta-propeller domains were predicted to be the main structural components for the scaffold Nups (Devos et al., 2006) (*Table 1*). The solved molecular structures from various scaffold Nups show a high similarity to clathrin vesicle-coats and NTRs of importin- β superfamily and have been predicted to have a common evolutionary origin (Brohawn et al., 2009). Based on their location in the NPC scaffold, these Nups have also been classified as the outer ring (Nup107-160 complex) and the inner ring (Nup205-Nup93 complex) Nups, however, this arrangement is controversial and a current topic of research in the field (Alber et al., 2007) (Brohawn and Schwartz, 2009) (Debler et al., 2008) (Amlacher et al., 2011).

2.4.2.3. Permeability barrier forming nucleoporins

The FG-repeats from FG-Nups radiate into the central channel and constitute the NPC permeability barrier (as discussed in section 2.4). These FG-Nups are incorporated into the NPC central channel through direct interaction with one or more scaffold Nups. The structured regions of FG-Nups are composed of protein interaction domains such as coiled-coils domains (Nup62-58-54 complex, Nup62-88-214) or autoproteolytic domains (Nup98) or RRM (RNA recognition motif) domains (Nup53) that mediate their interaction with the scaffold Nups (Devos et al., 2006) (*Table 1*).

2.4.3. Nucleoporin subcomplexes as the NPC building blocks

During mitotic disassembly of nuclear envelope, nucleoporins reside in the cytoplasm as biochemically defined subcomplexes and can be isolated as stable interaction assemblies from HeLa cells (Matsuoka et al., 1999) and mitotic *Xenopus* egg extract (Macaulay et al., 1995). Protein interaction studies between nucleoporins by antibody labeling (Pante et al., 1994), yeast genetic screens (Doye and Hurt, 1995) and biochemical fractionations (Schlaich et al., 1997) (Allen et al., 2002) (Bailer et al., 2000) etc. have revealed that Nups in the assembled NPCs also exist as multi-protein subcomplexes. These sub complexes get arranged around a cylindrical axis to give NPC its characteristic octahedral symmetry. Thus, nucleoporin subcomplexes act as architectural modules for NPC formation.

2. Introduction

NPC localization	Vertebrate Nup	Yeast Nup	Structural elements	Relative* stoichiometries
Cytoplasmic filaments	Nup214	Nup159		0.5
	Nup88	Nup82		0.5
	Nup62	Nsp1		~1.25
	Gle1	Gle1		0.25
	Nup358/ RanBP2	-		1
Outer scaffold ring	Nup133	Nup133		1
	Nup160	Nup120		1
	Nup107	Nup84		1
	Nup85	Nup85		1
	Nup96	Nup145C		1
	Seh1	Seh1		0.75
	Sec13	Sec13		1
	Nup37	-		1
	Nup43	-		1
	ELYS	-		0.5
	Aladin	-		0.75
Inner scaffold ring	Nup205	Nup192		~1
	Nup188	Nup188		0.5
	Nup93	Nic96		~1.5
	Nup155	Nup157/ 170		~1.75
	Nup35	Nup53/ 59		1
Central channel	Nup62	Nsp1		~1.25
	Nup58	Nup49		~0.75
	Nup54	Nup57		~0.75
	Nup98	Nup100/Nup116/ Nup145N		~1.5
Transmembrane nups	Ndc1	Ndc1		~0.75
	Pom121	-		~0.4
	gp210	Pom152		~1
	-	Pom34		-
Nuclear basket	Tpr	Mlp1		~1.25
	Nup153	-		1
	Nup50	Nup2		~0.75
		Nup1		-
	Nup60		-	

Transmem- Transmembrane, RRM- RNA recognition motif, AP- Autoproteolytic, RB-Ran-binding

Table 1: Domain architecture of different nucleoporins in the NPC

Adapted from (Brohawn *et al*, 2009). Relative stoichiometries of Nups compared to Nup107-160 complex are indicated according to (Ori *et al*, 2013). See text for details.

2. Introduction

2.4.3.1. NPC disassembly and reassembly from nucleoporin subcomplexes

In higher eukaryotes, NPC disassembly occurs during NE breakdown at the early stages of mitosis. Several studies have shed light to series of events leading to NPC disassembly and reassembly but the process is not understood with complete clarity yet. Phosphorylation is known to play a key role in NPC disassembly. Nucleoporins exist in hyper-phosphorylated state in M-phase *Xenopus* extract compared to S-phase (Miller et al., 1999). Mitotic disassembly is brought into effect by action of several kinases particularly CDK1 kinases (supported by Nek kinases), which specifically phosphorylate Serine and Threonine residues in the Nups. Recently, it was reported that phosphorylation of Nup98 is an important step leading to mitotic disassembly (Laurell et al., 2011). The order of NPC reassembly post mitosis is more clearly understood than its disassembly. It starts by recruitment of Nup107-160 complex on chromatin, which binds through its DNA binding subunit ELYS. This is followed by recruitment of transmembrane Nup Pom121 and Nup93-205 complex (Dultz et al., 2008) For review see (Antonin et al., 2008; D'Angelo and Hetzer, 2008). NPC biogenesis during interphase requires the insertion of NPC in intact NE. Also, the lower eukaryotes with closed mitosis require NPC insertion into an intact NE, which follows different operating mechanisms than the post mitotic NPC assembly. Assembly of interphase NPCs studied using time-lapse microscopy revealed that new pores do not contain components from pre-existing pores, thus are synthesized de novo (D'Angelo et al., 2006). Live cell microscopy analysis showed that interphase NPC assembly is slower compared to post mitotic assembly (Dultz and Ellenberg, 2010). Thus, mechanistic differences exist between two processes and are a topic of ongoing research (Imamoto and Funakoshi, 2012)

As depicted in *Figure 1.4a*, the major biochemically-defined subcomplexes present in the NPC include:

2.4.3.2. Nup107-160 complex (Y-complex)

The metazoan Nup107-160 complex consists of nine subunits – Nup133, Nup96, Sec13, Nup107, Nup85, Seh1, Nup160, Nup37 and Nup43 while the homologous yeast Nup84 complex has seven subunits in total comprising of Nup84, Nup120, Nup85, Nup145C, Sec13, Seh1 and Nup133 (Boehmer et al., 2008). The complex appears to be present symmetrically in the NPC scaffold on both sides of the NE and is the most well characterized NPC subcomplex. Metazoan Nup107-160 complex is thought to be an essential component for the NPC assembly as immunodepletion of Nup107-160 complex during in-vitro nuclear assembly reactions results in NE lacking NPCs (Walther et al., 2003). Both natively isolated and in-vitro constituted scNup84 complex show a Y-shaped structure in EM (Siniosoglou et al., 2000) and the complex therefore, is also referred to as the Y-complex (*Figure 1.4b*). Both yeast and metazoan Y-complex share some structural

2. Introduction

elements with COPII vesicle coat machinery such as α -helical structural folds and common components like Sec13. As a result a common evolutionary origin has been proposed (Brohawn et al., 2008).

Most of the subunits of γ Nup84 complex have been crystallized and have been shown to contain α -helical solenoids or β -propeller domains. Although the arrangement of the subunits within the complex is clear, in the absence of inter-complex molecular interaction data, various models for the arrangement of γ Nup84 complex in the NPC have been proposed (Bilokapic and Schwartz, 2012). The arrangement predicted by the Lattice model is based on shared similarity of scNup84 complex with COP II vesicle coats. According to this model, two parallel rings of scNup84 complex are arranged in a head-to-head fashion spanning the entire width of the NPC.

This molecular arrangement of γ Nup84 complex stabilizes the membrane curvature of the NE in a manner similar to COPII vesicle coat machinery; however, the orientation and interaction of the inner ring complex in this model is unclear (Brohawn and Schwartz, 2009). The Fence-pole model based on the packing of subunits in the crystallized fragments of scNup84 complex proposes a head-to-tail arrangement of the complex in four concentric rings (Debler et al., 2008). However, the model argues for promiscuous interactions between the subunits of the complex and proposes a dynamic arrangement of NPC scaffold where conformational changes in the NPC structure would occur in response to cargo passage (Hoelz et al., 2011). A third, computational model, based on the integration of available positional and interaction data of various Nups, predicts head to tail arrangement of the scNup84 complex in the NPC in two concentric rings. The model proposes presence of two outer rings and two concentric inner rings with linker Nups, scNic96 and scNup82 connecting the two rings (Alber et al., 2007).

2.4.3.3. *Nup93-Nup205 complex*

Some of the largest Nups that do not contain FG-repeats and are not a part of the Y-complex include Nup205, Nup188, Nup155, Nup93 and Nup53. Stoichiometric interactions have been reported between some of these Nups based on biochemical fractionations and yeast genetic screens (Aitchison et al., 1995) (Marelli et al., 1998) but the molecular interaction of the subunits with each other is much less understood than the Y-complex. These Nups connect with trans membrane Nups, peripheral Nups as well as outer ring Nups, therefore, the complex, altogether, is also referred to as the linker complex. Crystal structures from scNic96 (vertebrate Nup93), Nup170 (vertebrate Nup155), scNup192 (vertebrate Nup205) and scNup188 have revealed a stacked alpha-helical solenoid arrangement in these Nups, with flexible domains similar to NTRs (Schrader et al., 2008b) (Sampathkumar et al., 2013) (Andersen et al., 2013). Recently, ctNup53 (*Chaetomium thermophilum* Nup53) and ctNic96 were shown to interact with large Nups

2. Introduction

ctNup192 and ctNup170 with short, flexible motifs. Such short, flexible motifs connecting the nucleoporins in a subcomplex are proposed to be responsible for the overall plasticity of the NPC (Amlacher et al., 2011).

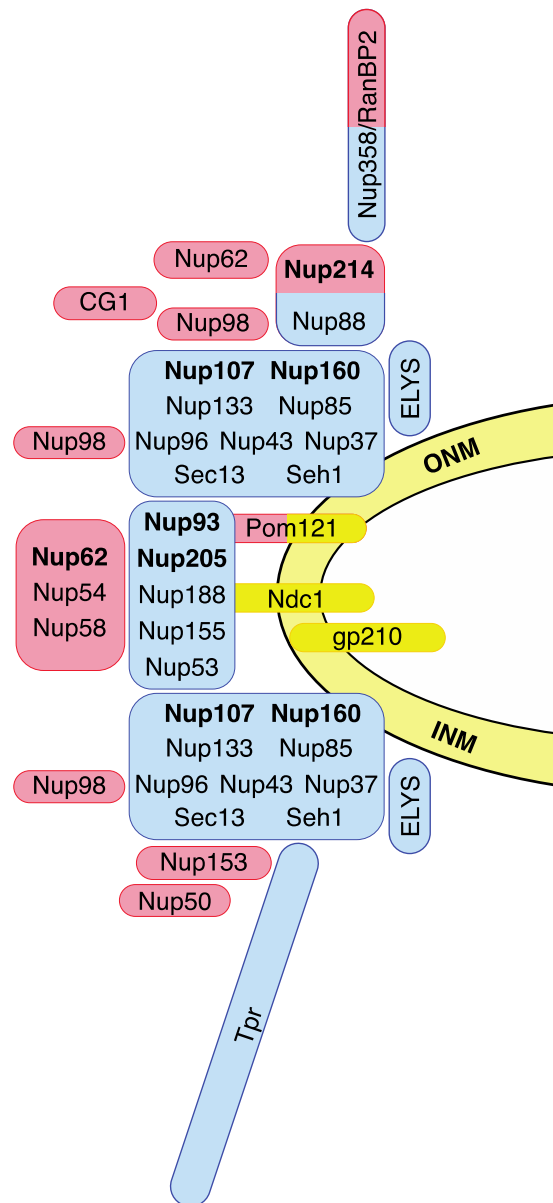


Figure 1.4: NPC subcomplexes

Current view of subcomplex arrangement in the NPC. In blue- scaffold Nups, red- FG Nups, yellow- Transmembrane Nups, ONM- outer nuclear membrane, INM- inner nuclear membrane (Image prepared by Dr. Aksana Labokha).

2. Introduction

In yeast, scNic96 co-purifies with Nsp1·Nup49·Nup57 complex (vertebrate Nup62·58·54 complex) (Grandi et al., 1993) where its N-terminal heptad repeats anchor Nsp1·Nup49·Nup57 complex in the NPC presumably through coiled-coil interactions (Grandi et al., 1995a). In vertebrates, however, the majority of Nup93 exists in a stable complex with Nup205 (Grandi et al., 1997).

2.4.3.4. *Nup214·88·62 complex*

Nup214·88·62 complex (scNup159·Nup82·Nsp1) is exclusively localized at the cytoplasmic side of the NPC (Fornerod et al., 1997b). Both Nup62 and Nup214 contain FG-repeats and all three subunits are also predicted to contain coiled-coil domains similar to Nup62·58·54 complex (*Table 1*). Nup214 plays an important role in the terminal steps of CRM1-cargo-RanGTP disassembly (Hutten and Kehlenbach, 2006) and contains a beta-propeller domain that provides interaction sites for mRNA export factors (Napetschnig et al., 2007). The complex has not been crystallized so far but presence of coiled-coil domains point towards a similar interaction as found in Nup62·58·54 complex.

2.5. Various approaches to study nucleoporin arrangement in the NPC

Studying the NPC in an atomic detail by any single technique has not been feasible mainly because of the challenge posed by its large size and hundreds of individual components. A high degree of modularity in the NPC structure, however, has allowed us to break down this large molecular puzzle into smaller pieces. As a result, solving the high-resolution structures of individual Nups and their subcomplexes by X-ray crystallography or NMR and fitting them in high confidence cryo-ET maps seems to be the most logical solution to the problem. The multidisciplinary approach followed in the field for structural characterization of NPC has been very successful and is discussed below.

2.5.1.1. *Nucleoporins stoichiometry*

The number of different nucleoporins present in the NPCs had been determined by mass-spectrometry of enriched NPC fractions from yeast and rat hepatocytes almost a decade ago. In both studies, identical set of nucleoporins was found to be present (~30) (Rout et al., 2000; Cronshaw et al., 2002). However, accurate knowledge of the number of copies of each Nup per NPC is limited with figures ranging from 8-48 copies per Nup in the NPC. As a result several models of Nup arrangement in the NPC exist. The picture is complicated by various other roles

2. Introduction

attributed to the Nups besides the components of NPC architecture such as in gene regulation, DNA replication and so on (Bukata et al., 2013) loss of weakly interacting peripheral Nups during isolation procedures. Recently, cell-type dependent variations in NPC composition were addressed and relative stoichiometries of individual Nups in human NPCs were quantified more accurately using targeted mass spectrometry (Ori et al., 2013) (listed in Table 1). More such targeted studies are required to accurately decipher the copy numbers of different Nups in the NPC and their heterogeneity if any, in response to cellular regulation.

2.5.1.2. High-resolution microscopy

Super-resolution microscopy applied to fluorescently labeled NPC components has the potential of revealing intricate NPC details that are not visible so far by any other means. Lately, NPCs have been studied by various high-resolution microscopic techniques, such as three-dimensional structural illumination microscopy (3D-SIM), which showed distinct differential localization of Nups within the NPC (Schermelleh et al., 2008). Polarized fluorescence anisotropy measurements of GFP-tagged Nups in live cells has revealed relative orientation of the nucleoporins, based on which a certain model of Y-complex arrangement in the NPC has been confirmed (Kampmann et al., 2011). Recently, using single molecule particle averaging methods in super-resolution microscopy, a precise arrangement of the Y-complex in the NPC were reported (Szymborska et al., 2013). Similar studies when applied to other nucleoporin complexes can ultimately reveal many other relative nucleoporin arrangements in the NPC.

2.5.1.3. X-ray crystallography

High-resolution crystal structures are now available for a majority of Nups and their interaction complexes. ScNic96 (Nup93) (Schrader et al., 2008a), scNup192 (Nup205) (Sampathkumar et al., 2013), scNup188 (Nup188) (Andersen et al., 2013) are among the largest Nups present in the inner scaffold ring whose structures have recently been solved successfully. All these protein exhibit alpha-helical solenoid like folds and flexible domains comparable to NTRs (*Figure 1.4*).

All components of heptameric scNup84 complex have been crystallized (Bilokapic and Schwartz, 2012). The complex is shown to be entirely built up from alpha-helical and beta-propeller domains with flexible-hinge regions. What remains to be seen is the configuration in which these proteins are arranged in the NPC.

In this arrangement (*Figure 1.4a*), the structure of Nup62·58·54 complex is still missing. In fact, it is the least understood NPC subcomplex both at a biochemical as well as the atomic level and is described below.

2. Introduction

2.6. Nup62·58·54 subcomplex

Nup62 was the first nucleoporin described as a structural component of the NPC more than 25 years ago when a punctate nuclear rim staining with a monoclonal antibody (mAB414) that recognizes Nup62 among other FG-containing Nups, was observed in the immunofluorescence microscopy (Davis and Blobel, 1986). The protein was later found to exist in a stable complex with two other FG-repeat containing NPC components, Nup58 and Nup54 (Finlay et al., 1991) (Kita et al., 1993).

The Nup62·58·54 complex (homologous complex in yeast Nsp1·Nup49·57 complex) lines the NPC's central transport channel (Grote et al., 1995). All three subunits of the complex contain FG-repeat domains at their N-terminus. In addition, Nup58 also has C-terminal FG-repeat domains. All FG-domains of the Nup62·58·54 complex contribute synergistically towards formation of the vertebrate NPC permeability barrier (Labokha et al., 2013). The FG-repeat domains of Nup62 are modified heavily by addition of N-acetyl glucosamine *in vivo* (Davis and Blobel, 1987) and as a result Nup62 is one of the main nucleoporins that binds WGA (Wheat germ agglutinin). The effect of this posttranslational modification on the structure of Nup62 was analyzed in baculoviral expression, where glycosylated Nup62 was shown to have the same structure as non-glycosylated Nup62 (Bailer et al., 1995). Thus, posttranslational modifications are not crucial for maintaining structure of Nup62 and have more functional roles such as regulating the NPC's permeability barrier as in case of Nup98 (Labokha et al., 2013).

The C-terminal coiled-coil domains of all three subunits in the Nup62·58·54 complex interact strongly with each other (Buss and Stewart, 1995). The coiled-coil domains are critical in mediating interactions between the subunits with each other as well as anchoring the complex to the NPC. Single amino acid substitutions in the heptad repeat in these coiled-coil domains were found to cause growth defect in yeast (Nehrbass et al., 1990). In yeast, the Nsp1·Nup49·Nup57 complex shows a strong interaction with scNic96 (yeast homolog of vertebrate Nup93). The regions responsible for this interaction have been mapped at the coiled coil domain containing N-terminus of scNic96 (Grandi et al., 1995b).

Recently, *Xenopus* nuclei reconstitution experiments showed that depleting xNup62·58·54 complex from the NPC leads to formation of nuclei having slightly deviated morphology and a reduced nuclear transport (Hulsmann et al., 2012). It is also one of the last components assembled into the NPC's structural backbone (Dultz et al., 2008). Thus, Nup62·58·54 complex has more functional roles in the NPC than imparting structural integrity. The complex being present at the inner most part of the NPC's central channel can probably assume a more flexible

2. Introduction

conformation compared to the more rigid scaffold Nups in the vicinity. This might be one of the reasons behind failed attempts to crystallize Nup62·58·54 complex so far.

However, a partial *rattus norvegicus* Nup62·58·54 complex was assembled by Melcak and coworkers (Melcak et al., 2007), which formed protein concentration and salt dependent oligomeric assemblies in solution. These multimeric protein assemblies were proposed to be physiologically relevant in the context of the NPC. Small, 50-60 residues long fragments of Nup58 and Nup54 have been crystallized (Melcak et al., 2007) (Solmaz et al., 2013) that showed crystal packing into homo oligomeric assemblies in different conformations. The same fragments also could be crystallized in hetero-oligomeric assemblies with each other (Nup54·58) (Solmaz et al., 2011). Thus, a promiscuous interactions between the coiled-coil domains of Nup62·58·54 complex subunits was shown that has been explained in terms of a 'ring dilation and contraction' model of the NPC to explain their physiological relevance (Solmaz et al., 2011) (Solmaz et al., 2013). The model proposes dynamic instability of the Nup62·58·54 complex inside the nuclear pore complex, resulting in dynamic association and dissociation of Nup62·58·54 complex subunits from each other in order to dilate and constrict the NPC ring. According to this mechanism, the changes in the NPC diameter brought about by such dynamic assemblies of Nup62·58·54 complex subunits can bring about changes in the NPC diameter in the range of ~20 nm sufficient to contract and dilate the NPC ring.

Severe problems with this model exist. First, the authors do not take into account the extremely 'chopped off' nature of the complex studied by them that does not contain all the predicted structured regions in the subunits. Since, coiled-coil domains are commonly known to engage in promiscuous interactions in the absence of right interaction partners, it is possible that the observed homo and hetero-oligomeric assemblies of the fragments are simply a consequence of missing interaction partners. Second, such a model is difficult to reconcile with the arrangement of scaffold Nups in the NPC that do not undergo such large structural arrangements. Third, the copy number and stoichiometry of Nup62·58·54 complex proposed by the model (128 copies of Nup62, 32 copies of Nup58 and 64 copies of Nup54) does not match the copy numbers actually observed by many of the proteomic studies (which generally predict 16-48 copies of each subunit) (Cronshaw et al., 2002) (Ori et al., 2013). Fourth, the regions covered in the structure constitute only about 40% of the total structured regions in the Nup62·58·54 complex and the contribution or arrangement of other coiled-coil domains in the complex is not taken into account. Hence, the structure from a complete Nup62·58·54 complex containing all the structured elements will be closer to the real Nup62·58·54 complex structure within the NPC, and is the focus of this work.

2. Introduction

2.7. The exportin CRM1

CRM1 (or Xpo1) is an essential NTR belonging to importin- β superfamily, which exports a wide variety of cargoes out of the nucleus (Fukuda et al., 1997) (Stade et al., 1997; Fornerod et al., 1997a). The range of cargo molecules exported by CRM1 ranges from small proteins such as snurportin 1 (SPN1) to large complexes such as 60S pre-ribosomal subunits and HIV genomic RNA (Fornerod et al., 1997a; Kehlenbach et al., 1998) (Güttler and Görlich, 2011).

CRM1 recognizes short, leucine-rich peptide motifs as export signals in its cargoes, termed as leucine rich repeats (LRRs) (Kutay and Guttinger, 2005). The crystal structure of CRM1-RanGTP-snurportin complex (Monecke et al., 2009) elucidated that during cargo export, CRM1 forms a toroid ring with its N- and C-terminus in close proximity to each other. The ring encapsulates RanGTP at its concave surface towards the N and C-terminal HEAT repeats, which engage in extensive electrostatic interactions with Ran. In addition, a β -hairpin loop (called the acidic loop) originates from HEAT repeat 9 and extends into the middle of the CRM1 ring to dock RanGTP in its position. SPN1 binding occurs at the outer surface of CRM1 ring where alpha-helical NES of SPN1 docks in the hydrophobic cleft (HC) formed by HEAT repeats 11 and 12 of the CRM1 ring (*Figure 1.5*). Thus, in contrast to other exportins where cargo and RanGTP binding occurs in close proximity to each other, CRM1 has spatially separated sites for binding of the two. The binding of the cargo outside the CRM1 ring also explains the basis behind recognition of huge variety of cargoes by CRM1 compared to other transport receptors (Monecke et al., 2009). Recently, the structural basis behind ability of CRM1 to bind wide variety of cargoes was deciphered (Güttler et al., 2010). Accordingly, the nature and spacing of the Φ residues in the NES is the determinant of the cargo recognition by CRM1.

It is known that RanGTP binding to CRM1 increases its affinity for cargo binding by approx. 1000 folds (Monecke et al., 2009). In the absence of a direct interaction between cargo and RanGTP in the export complex, this increase in affinity could only be explained by an allosteric mechanism. Therefore, it was proposed that conformational changes induced in CRM1 on RanGTP binding lead to structural changes in the hydrophobic cleft resulting in the cargo binding and *vice-versa* (Monecke et al., 2009). Later, crystal structures of CRM1 in disassembly complex with RanBP1 and RanGTP revealed an intermediate CRM1 conformation where RanGTP binding induced conformational changes in the CRM1 hydrophobic cleft were observed (Koyama and Matsuura, 2010). In this conformation, CRM1's acidic loop adopted a compact conformation leading to closing of the hydrophobic cleft in the disassembly complex. Recently, structures of unliganded CRM1 from *S.cerevisiae* (Saito and Matsuura, 2013) and *C.thermophilum* (Monecke et al., 2013)

2. Introduction

were reported, showing an overall extended conformation of free CRM1, a compact acidic loop and a closed hydrophobic cleft (*Figure 1.5*). In these structures, C-terminal helix of scCRM1 was shown to regulate CRM1 transition from extended to compact conformation on RanGTP binding. Molecular dynamics simulations have revealed that CRM1 in extended conformation is highly flexible and binding of RanGTP makes it rigid and compact (Dolker et al., 2013).

So far, it is clear from the structural studies that CRM1 undergoes many conformational changes during its function as an export receptor. However, the molecular details of events leading to these structural and conformational changes are not understood. As a part of this project, we aimed to identify critical CRM1 residues that may be involved in stabilizing 'cargo-free' and 'cargo-bound' CRM1 conformations. For this, we looked into the evolutionary conservation of CRM1 and identified the residues that have been thoroughly conserved across eukaryotic clades. Mutating these residues will shed light on various conformational states that CRM1 adopts during its function as a transport receptor.

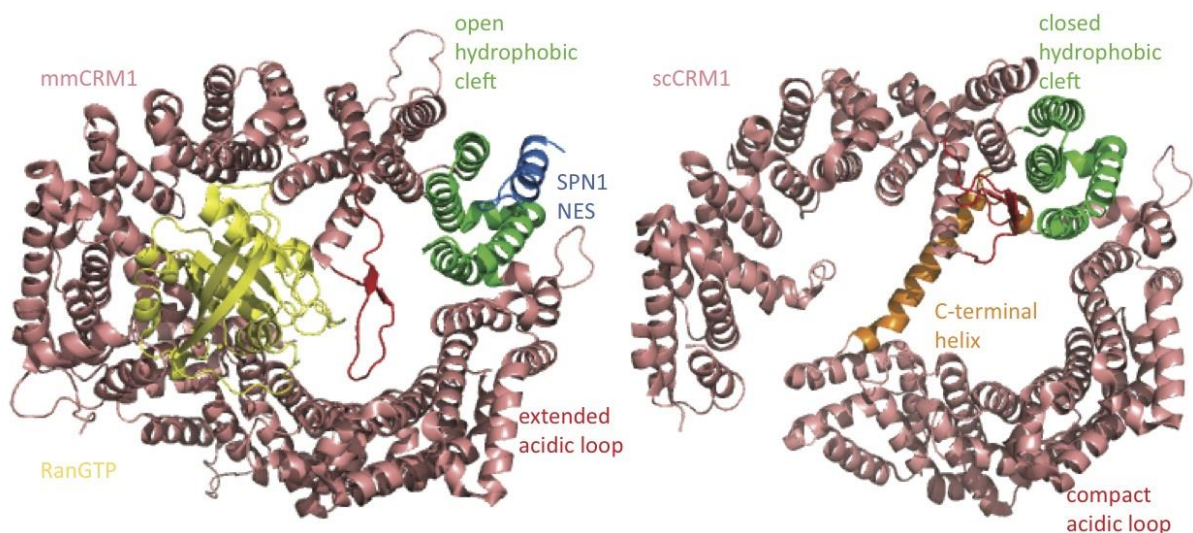


Figure 1.5: Structure of the CRM1-RanGTP-SPN1 export complex and unliganded CRM1

CRM1 conformation in the ternary CRM1-RanGTP-SPN1 export complex (left) and unliganded form (right) as observed in the respective crystal structures. Open hydrophobic cleft (green) and an extended acidic loop (red) in the CRM1 toroid ring (salmon) are observed in the CRM1-RanGTP-SPN1 export complex whereas a closed hydrophobic cleft and a compact acidic loop are present in the unliganded CRM1. For clarity, only the NES of bound SPN1 (blue) is shown. RanGTP (yellow) binds the inner surface of CRM1 ring. Adapted from Monecke *et al*, 2009 (left) and Saito *et al*, 2013 (right).

2.8. About this work

In this project, I studied the essential elements of the NPC architecture and function by biochemical analysis of Nup62·58·54 complex; a major NPC structural element and the functional characterization of the exportin CRM1, a major player in facilitated NPC translocation. The first

2. Introduction

chapter describes our data on biochemical properties of xINup62·58·54 complex and subsequent crystallization attempts for its structure determination. Here, we first addressed many conflicting questions in the literature about Nup62·58·54 complex's stoichiometry in the NPC, published reports about its oligomerization behavior and the inter-subunit interactions. We then describe its interaction with Nup93, another nucleoporin thought to anchor Nup62·58·54 complex into the NPC. The last part describes our approach to crystallize Nup62·58·54 complex by conjugating it to the nanobodies, and the initial crystallization trials.

The second chapter describes the functional features of the exportin CRM1, an essential nuclear transport receptor in all eukaryotes. The X-ray crystallographic structures of CRM1 in different functional complexes are already available but the overall mechanism by which CRM1 adopts various conformations in these complexes is unclear. We adopted a mutagenesis approach to study this phenomenon where we mutated CRM1 residues that are strictly conserved during eukaryotic evolution. Here, we describe the effects of these mutations on CRM1 function through different functional assays that can be related to stabilization of one or other CRM1 conformations.

3. Results

3.1. Biochemical and structural investigation of xINup62·58·54 complex

3.1.1. Coiled-coil domains in xINup62·58·54·93 complex

It has been reported since a long while ago that the amphipathic α -helices of all three subunits of Nup62·58·54 complex contain heptad repeats (HR) characteristic of the coiled-coil domains (Grandi et al., 1995b). Coiled-coil domains are the special interaction motifs in proteins build up of HR units, which are typically comprised of seven residues numbered *a-g*. In double-stranded coiled-coils, heptad positions *a* and *d* are occupied by hydrophobic amino acids whereas the rest of the positions contain polar residues. Prediction algorithms such as COILS (Lupas et al., 1991) score each heptad position for coiled-coil forming propensity by comparing it to the sequences of known coiled-coil proteins and thereby predict coiled-coil domains in the protein sequences. We used COILS algorithm to predict coiled-coil domains in *Xenopus laevis* Nup62, Nup58, Nup54 and Nup93. *Figure 3.1a* shows coiled-coil probability vs residue positions for three scanning window sizes (14, 21 and 28 residues) where short, intermediate and long stretches of alpha helices participating in coiled-coil interactions can be observed. In xINup62·58·54 complex, almost the entire C-terminal structured region is predicted to form coiled-coils whereas in xINup93, the extreme N-terminus shows weak coiled-coil forming propensity (along with other regions spread across the protein sequence). MultiCoil is another coiled-coil prediction algorithm that uses the similar approach to differentiate three-stranded coiled-coils from two-stranded coiled-coils (Wolf et al., 1997). We also analyzed all four subunits of xINup62·58·54·93 complex for their trimer forming potential using MultiCoil.

For xINup62, the entire structured region (342-547) can be divided into five sub domains, CC1-CC5, containing heptad repeats separated by flexible linker regions. The two central subdomains, CC2 and CC3, are longer (>28) and have high probability for coiled-coil interactions whereas the peripheral domain CC1 is shorter (~21 residues) but still has high probability for coiled-coil domain interactions. The C-terminal end of Nup62 comprises of two small regions, CC4 and CC5, with weak propensity to form coiled-coils. MultiCoil predicted the second coiled-coil domain CC2 to have strong trimeric coiled-coil propensity (*Figure 3.1b*). xINup58 was predicted to have two large domains, CC1 and CC2, both with high probability to form coiled-coils.

3. Results

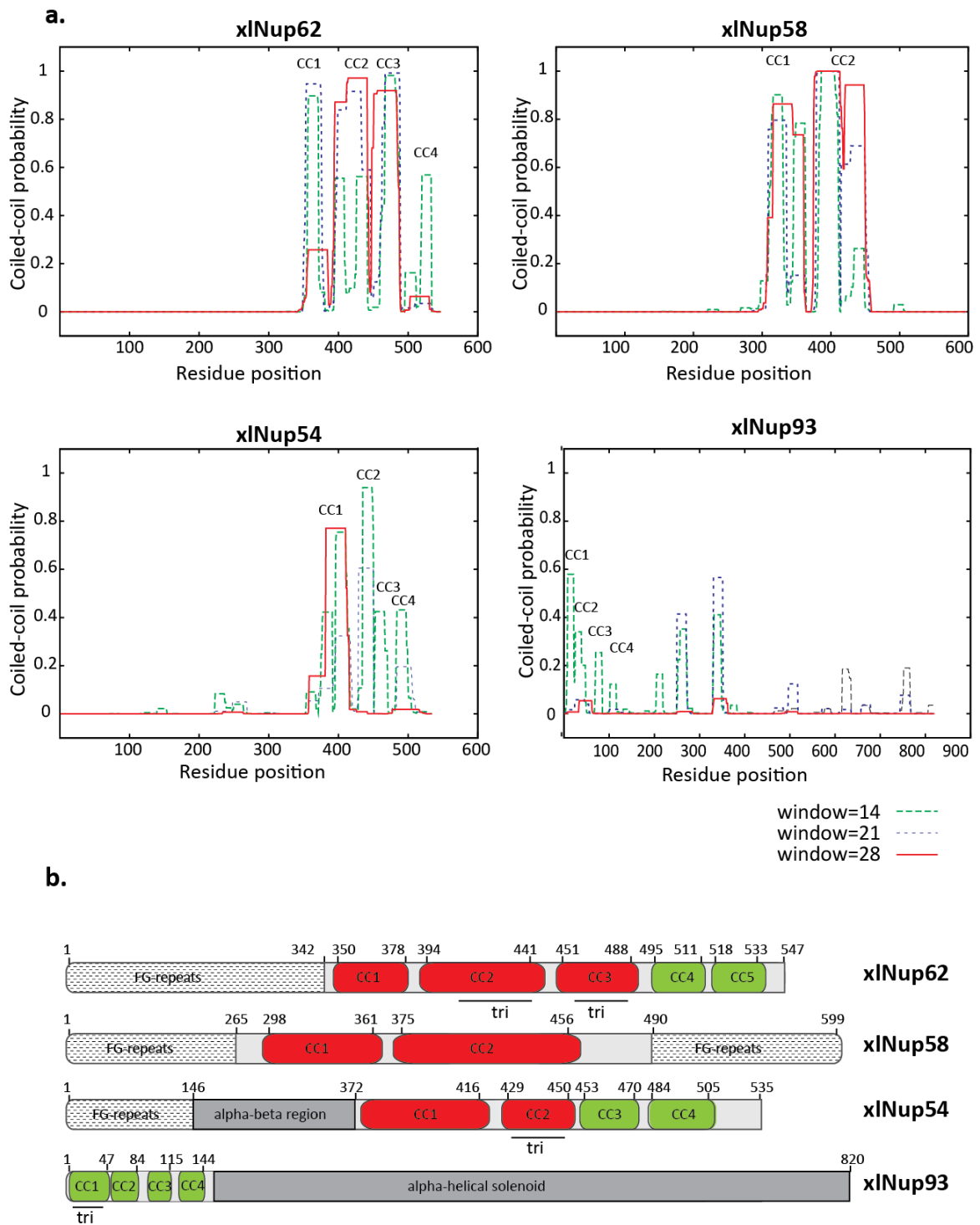


Figure 3.1: Coiled-coil domains in Nup62·58·54·93 complex

a. Prediction of coiled-coil domains in *Xenopus laevis* Nup62, Nup58, Nup54 and Nup93 using COILS algorithm (Lupas et al, 1991). The program calculates coiled-coil forming probability per residue, which is plotted against the residue position in the sequence. 14, 21 and 28 residue window lengths are shown in green, blue and red respectively. **b.** Domain organization of *Xenopus laevis* Nup62, Nup58, Nup54 and Nup93 showing FG-repeat, coiled-coil (CC), alpha-beta and alpha-solenoid domains. Coiled-coil domains with strong and weak coiled-coil forming propensity are depicted in red and green respectively. Regions in Nup62, Nup54 and Nup93 with strong trimeric coiled-coil forming propensity, as predicted by MultiCoil (Wolf et al, 1997) are underlined.

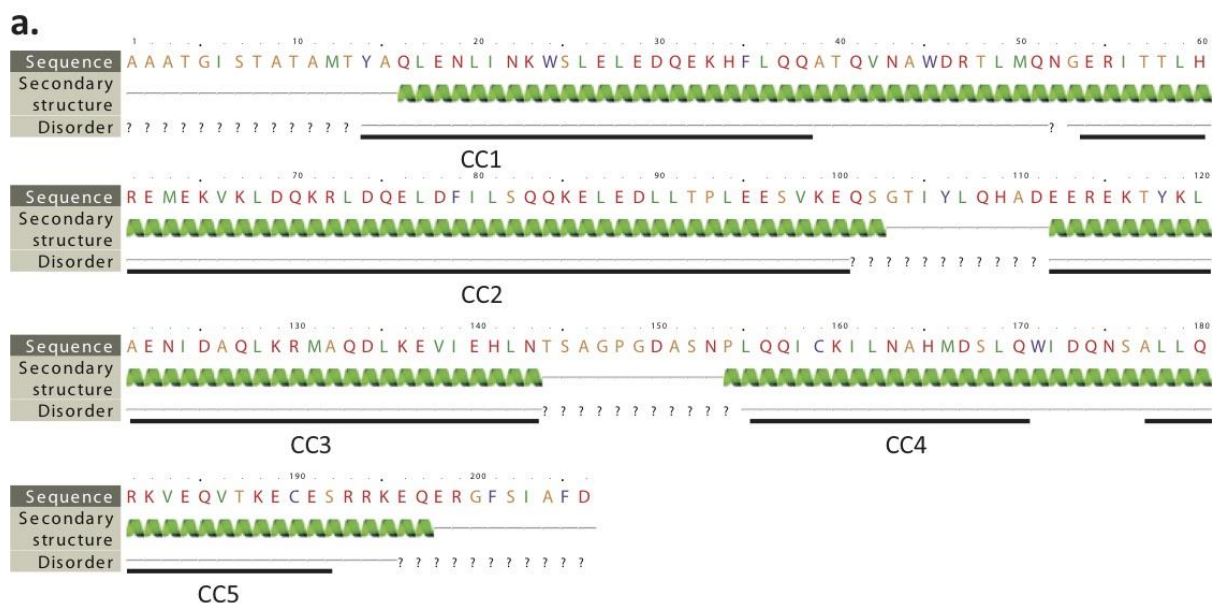
3. Results

xINup54 on the other hand, contains one large and one small domain (CC1 and CC2 respectively) with high coiled-coil forming propensity and two C-terminal domains with low coiled-coil forming potential which is similar to xINup62. Multicoil algorithm predicted a strong trimer forming potential for Nup54 CC2. xINup93 N-terminus was predicted to contain four small N-terminal domains, CC1-CC4, with low coiled-coil forming probability. The trimer forming potential was strong for Nup93 CC1 and CC2. Thus, all coiled-coil regions in xINup62-58-54-93 complex are not similar in their predicted interaction strength and oligomerization state.

3.1.2. Δ FG-xINup62-58-54 complex

It was observed in our lab before that FG-repeats in xINup62-58-54 complex are not necessary for the complex formation between its subunits and that a complex lacking FG-repeats is fully capable of incorporation into in-vitro assembled NPCs in *Xenopus* nuclei (Bastian Hülsmann, unpublished data). Based on this observation, the N-terminal boundaries of Nup62 and Nup54 were determined and the sequence essential for complex formation was determined to be Nup62 (342-end) and Nup54 (146-end) respectively. As Nup58 also contains C-terminal FG-repeats, its N-terminal boundary was determined by in-vitro incorporation assay while C-terminus FG-region was removed based on sequence analysis resulting in the boundaries Nup58 (265-490). xINup62 and xINup58 boundaries determined by in-vitro incorporation assay contain the entire coiled-coil structured regions predicted by COILS algorithm. In case of Nup54 an additional domain stretch of residues not predicted to form coiled-coils (146-372) was also required for complex formation. Secondary structure analysis of Nup54 predicts this region to be composed of α -helices and β -sheets (from here on referred to as Nup54 α/β domain).

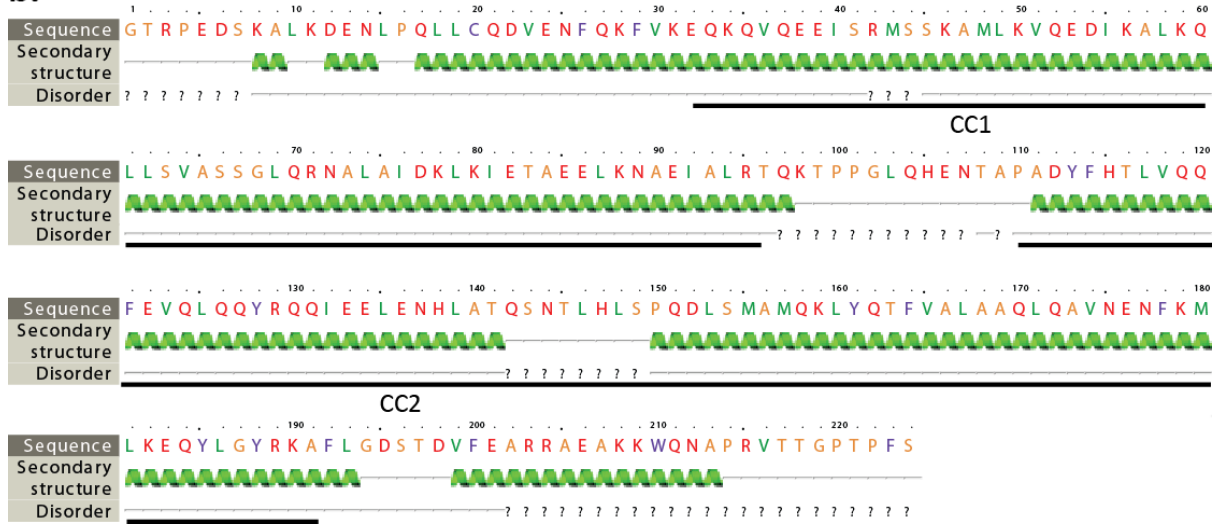
Δ FG-xINup62 (342-547)



3. Results

Δ FG-xINup58 (265-490)

b.



Δ FG-xINup54 (146-535)

c.

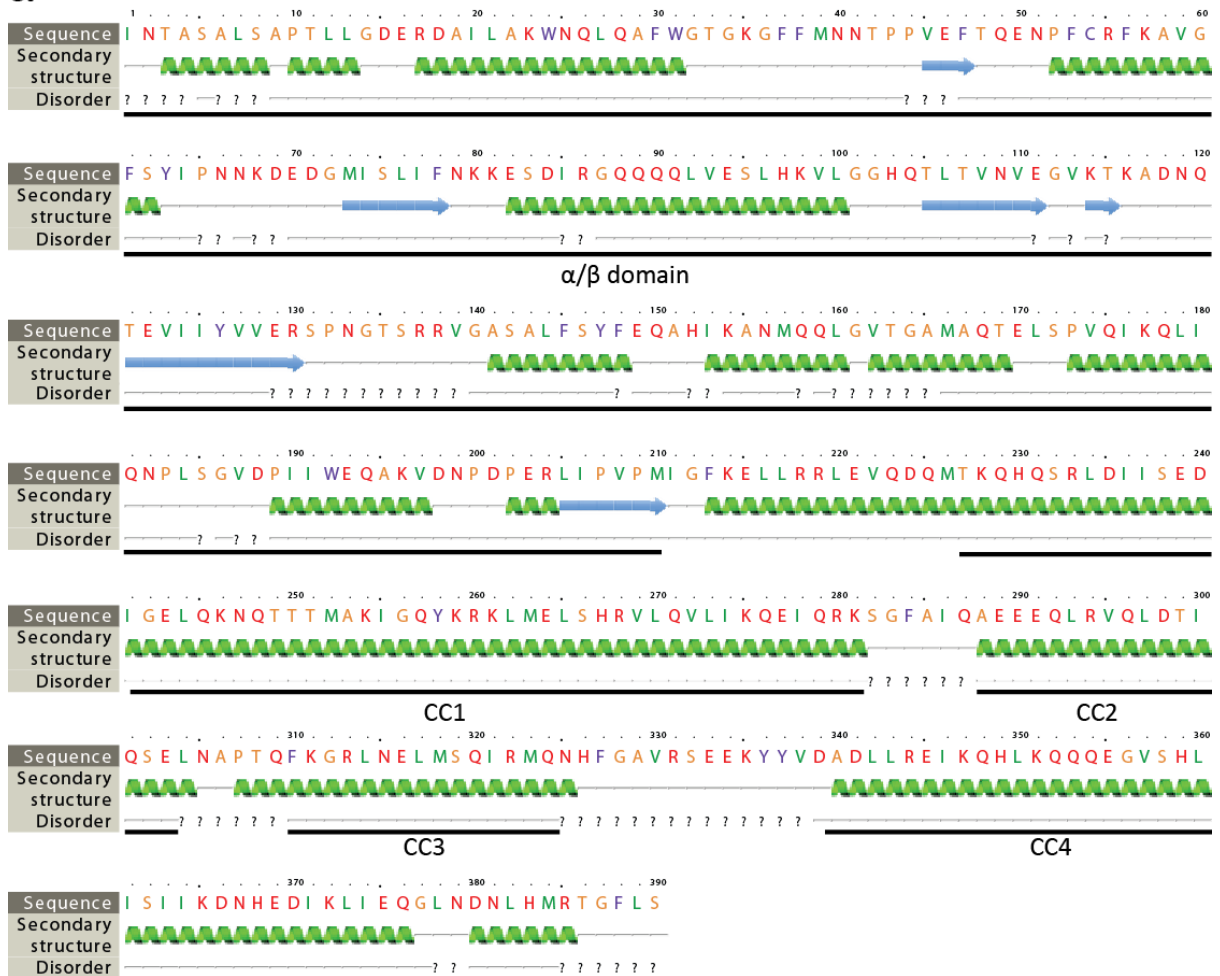


Figure 3.2: Secondary structure prediction for *Xenopus laevis* Δ FG-Nup62-58-54 complex

a. Δ FG-xINup62 (342-547) b. Δ FG-xINup58 (265-490) c. Δ FG-xINup54 (146-535). Secondary structure and disordered regions in Nup62-58-54 complex were predicted using Phyre2 secondary structure prediction tool. The program predicts protein secondary structure and disordered regions in the protein sequence

3. Results

based on PSI-BLAST output. The program predicts all Nup62·58·54 complex subunits to be comprised of alpha helices (green) with interspersed disordered linker regions (question marks). Coiled coil domains predicted from COILS algorithm are underlined. In addition to coiled-coil forming alpha helices, Nup54 also contains an additional region made up of alpha helices and beta-strands towards its N-terminus, termed as Nup54 α/β domain.

The secondary structure and disorder prediction for xINup62 (342-547), xINup58 (265-490) and xINup54 (146-590) (from here on referred to as, Δ FG-xINup62·58·54 complex) are shown in *Figure 3.2*. All subunits appear to contain coiled-coil domains interspersed with disordered regions. The extreme C-terminal regions of xINup58 and xINup54 although structured, do not contain coiled-coil domains and were predicted to be flexible. For crystallization, Δ FG-xINup62·58·54 complex was expressed in *E.coli*.

3.1.3. Co-expression with interaction partners is required for Δ FG-xINup54 solubility

To purify Δ FG-xINup62·58·54 complex for crystallographic analysis, all subunits were cloned as His₁₄-tagged constructs with protease cleavage sites in between and expressed in *E.coli*.

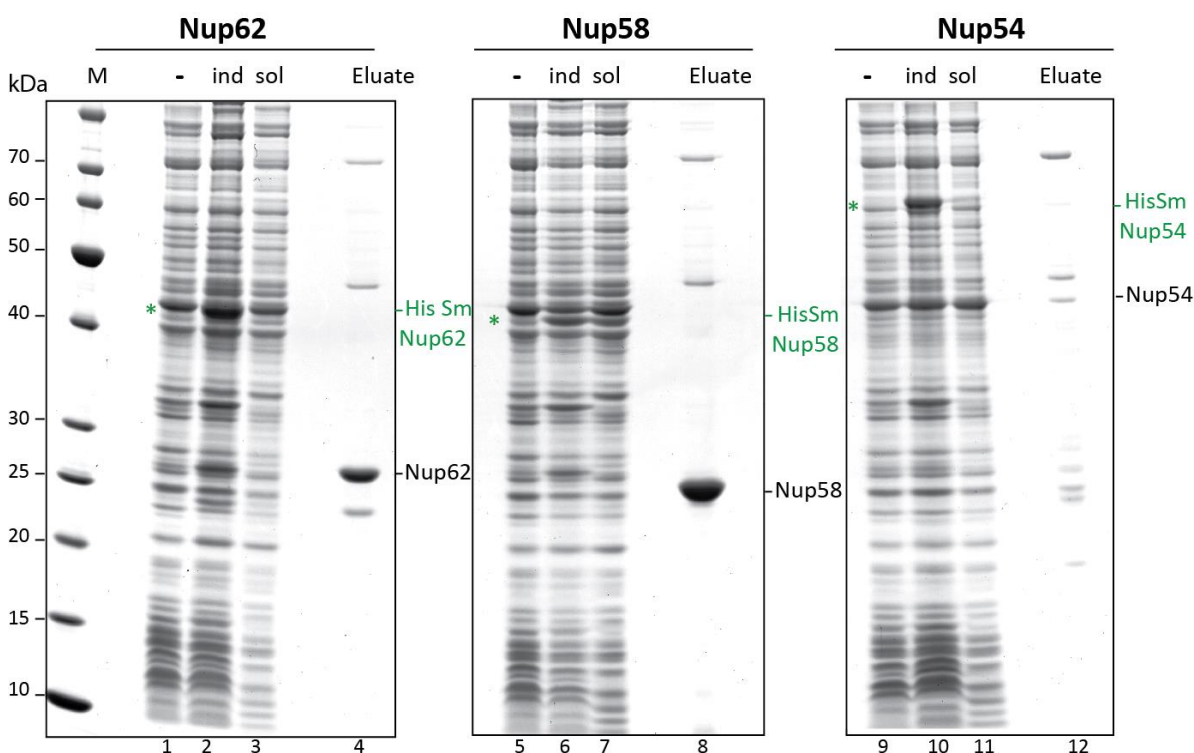


Figure 3.3: Solubility of Δ FG-xINup62·58·54 complex subunits

Nup62 (342-547), Nup58 (265-490) and Nup54 (146-535) were expressed individually in *E.coli* and purified by Nickel-affinity chromatography followed by on-column protease elution. Δ FG-xINup62 and Δ FG-xINup58 were moderately soluble (green asterisk marked bands in Lane 3, 7) while Nup54 was mostly insoluble (Lane 11). Purified proteins were obtained by on-column protease cleavage. The ~15kDa shift in expressed and eluted protein size is due to removal of the His-Sumo-tag. Equal amounts nickel eluents were loaded on the SDS gel. "-", Uninduced; "ind", induced; "sol", soluble fraction; "El"- eluate after on-column protease cleavage.

3. Results

Δ FG-Nup62 and Δ FG-Nup58 were found to be moderately soluble on expression and could be purified in sufficient amounts. Δ FG-Nup54 on the other hand, was found to be mostly insoluble (*Figure 3.3*). All attempts to obtain Δ FG-xINup54 in a soluble form such as growth at low temperature (16°C), inclusion of solubility enhancing tags (ZZ, MBP) etc. did not enhance its solubility.

In general, insolubility of heterologous proteins in *E.coli* can be attributed to several factors such as absence of eukaryotic folding machinery required for proper protein folding and stability, aggregation prone exposed hydrophobic surfaces on the protein of interest or increase in protein solubility limit during overexpression. Co-expression of subunits of a multi-protein complex or a protein with a chaperone has been shown before to have an effect on increasing protein solubility and stability in *E.coli* (Li et al., 1997) (Bross et al., 1993). Therefore, we designed multi-cistronic expression constructs for co-expression of Δ FG-Nup54 with Δ FG-Nup62 and Δ FG-Nup58 in *E.coli*. We used a double-tag affinity purification method where the insoluble partner Δ FG-Nup54 was His₁₄-tagged while its interaction partners Δ FG-Nup62 and Δ FG-Nup58 had a N-terminal ZZ-tag. Both His₁₄ and ZZ tags were followed by protease recognition sites (of brSUMO protease) to cleave off the affinity tags directly from the Ni⁺² affinity matrix after protein binding. Binding the co-expressed proteins to Ni⁺² column using the His₁₄-tag on Δ FG-Nup54 followed by on-column protease elution, led to specific elution of either Δ FG-Nup54 or proteins directly bound to Δ FG-Nup54 from the Ni⁺² column.

Interestingly, we found that Δ FG-Nup54 becomes greatly soluble on co-expression with Δ FG-Nup62 or Δ FG-Nup58 indicating that these subunits provide interaction sites for Δ FG-Nup54, thus preventing its aggregation into the inclusion bodies. Both Δ FG-Nup54-62 and Δ FG-Nup54-58 complexes were pulled out on Ni⁺² affinity matrix by using the His-affinity tag on Nup54 (*Figure 3.4*). The complexes were eluted by on-column protease cleavage using brSUMO protease. Roughly 1:1 stoichiometry of subunits was observed for Δ FG-Nup54-62 complex and 1:2 stoichiometry for the Δ FG-Nup54-58 complex on the SDS-gel. Thus, co-expression can be used as a tool for expressing Δ FG-Nup54 in a soluble form, which also leads to formation of stable complexes between the subunits that can be purified efficiently from *E.coli*.

3. Results

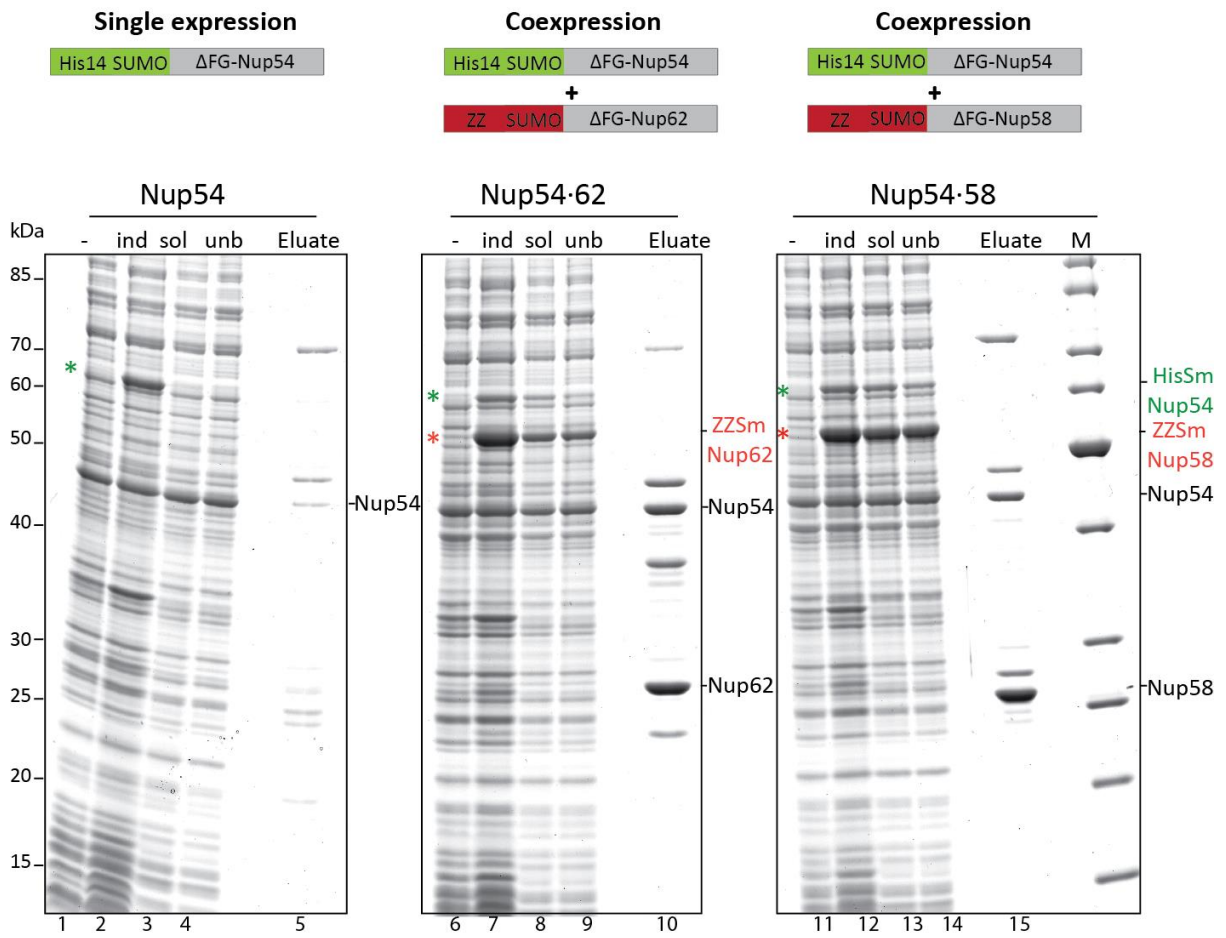


Figure 3.4: Co-expression with interaction partners makes Nup54 soluble

His₁₄ brSumo-Nup54 (HisSm-Nup54) was expressed singly or co-expressed with ZZ brSumo-Nup62 (ZZSm-Nup62) or ZZ brSumo-Nup58 (ZZSm-Nup58) on a compatible plasmid vector in *E. coli*. The bands marked with asterisks are the expressed proteins. As seen in lane 3, Nup54 is insoluble when expressed alone, but gets soluble when co-expressed with either Nup62 or Nup58 (lane 8 and 13 respectively). Nup54 elutes as a complex with Nup62 (Nup54·62) and with Nup58 (Nup54·58) after Ni²⁺ pull down and on-column protease cleavage. The shift in the band sizes in expressed and eluate fractions is due to the removal of corresponding His₁₄-Sumo and ZZ-Sumo tags. “-”, Uninduced; “ind”, induced; “sol”, soluble fraction; “unb”, unbound fraction; El- eluate after on-column protease cleavage.

3.1.4. *xlNup58* and *xlNup62* also engage in a direct interaction

In various studies before, native Nup62·58·54 complex has been isolated from NE extracts of different cell types using various isolation procedures resulting in different stoichiometries of the isolated complex. The reported stoichiometries range from 4:1:4 (Finlay et al., 1991) to 1:1:2 (Kita et al., 1993) (Buss and Stewart, 1995) for Nup62·58·54 subunits. A 1:1:1:1 (Guan et al., 1995) stoichiometry has also been reported for Nup62·58·54·45 complex isolated from rat liver nuclei (Nup45 being alternative spliced version of Nup58 in rats). In yeast, however, using a combination of genetic tagging and deletion studies, a more consistent 1:1:1 stoichiometry of Nsp1·Nup49·57 complex has been reported (Schlaich et al., 1997; Grandi et al., 1995b) (Grandi et al., 1997). Individual subunits of scNsp1·Nup49·57 complex have also been recombinantly

3. Results

expressed and purified from bacteria under denaturing conditions to reconstitute the complex (Bailer et al., 2001). Strong interactions observed between Nup54 and Nup62 (and the corresponding scNup57, scNsp1 subunits) in the reconstitution experiments under denaturing conditions have led to the view that Nup54 binds Nup62 and Nup58 directly (Buss and Stewart, 1995). However, no direct proof of interaction between Nup62 and Nup58 have been reported so far.

We co-expressed Nup62 and Nup58 in *E.coli* using the double affinity tag method (Figure 3.5). The purification was done under two different salt concentrations, at 50mM NaCl (low salt or ls) or 500mM NaCl (high salt or hs) in the buffer to see effect of salt concentration, if any, on this interaction.

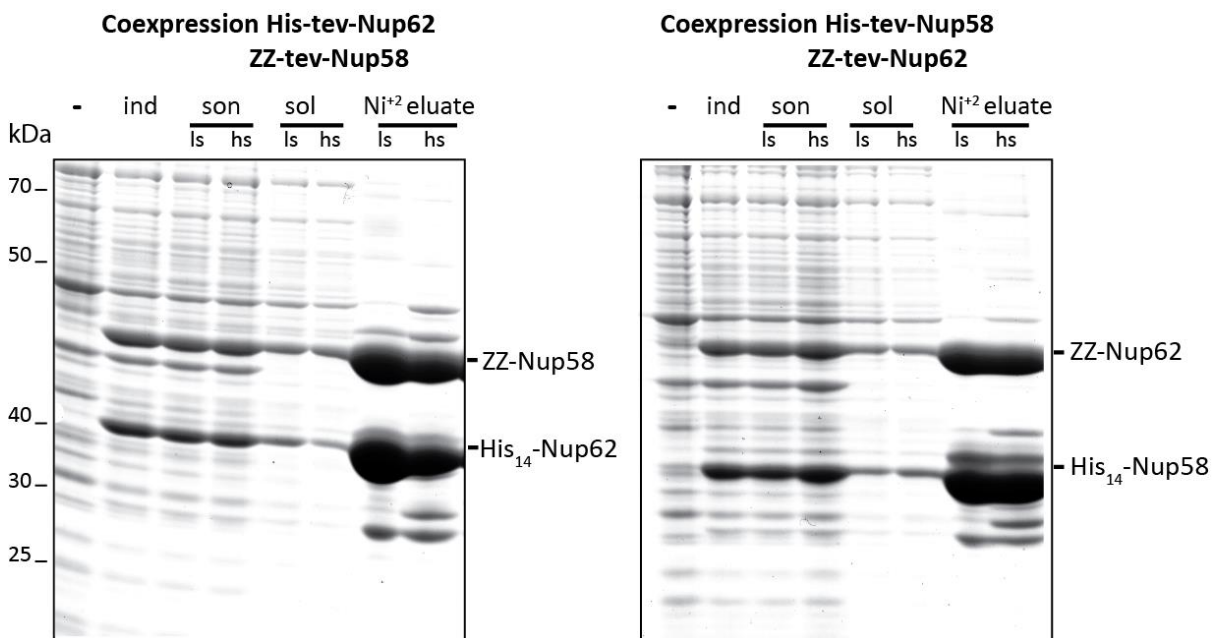


Figure 3.5: Direct interaction between Nup62 and Nup58

Direct interaction between Nup62 (342-547) and Nup58 (265-490) was observed after co-expression using different tags under different salt conditions. The Nup62·58 complex could be pulled out using affinity tag on either subunit – His₁₄-tag on Nup62 and ZZ-tag on Nup58 (left) or in an alternate tag system using His₁₄-tag on Nup58 and ZZ-tag on Nup62 (right) under two different salt conditions. “-”, Uninduced; “ind”, induced; “son”, sonicated fraction; “sol”, soluble fraction; “Ni²⁺ eluate”, fraction after imidazole elution from Ni₂ column; “ls”, low salt buffer 50mM NaCl; “hs”, high salt buffer 500mM NaCl.

Under all the conditions tested, Nup62 and Nup58 interacted and formed a complex that could be pulled out using tag on either of the subunit. Salt concentration in the buffer (50mM or 500mM) did not have any visible effect on the interaction indicating that the interaction was specific and stable. Therefore, we could establish that a direct interaction between Nup62 and Nup58 exists and that both are able to form a binary complex. However, on a prolonged incubation at 4°C for 2-3 days, the binary Nup62·58 complex showed a visible aggregation. In a

3. Results

similar manner, although, Nup54 could be obtained in a soluble form by co-expression and binary complexes between Nup54-62 and Nup54-58 could be purified in high yields, these complexes were prone to form oligomers during gel-filtration. This tendency of binary complexes to aggregate and oligomerise indicated exposed hydrophobic residues on the surface of these subunits that most probably are the interaction sites for the absent third subunit. This also shows that in the absence of interaction partner the Nup62-58-54 complex subunits undergo promiscuous interactions leading to higher-order oligomers.

3.1.5. A stable Δ FG-Nup62-58-54 complex can be purified under native conditions from *E.coli* after co-expression

Next, we co-expressed all three subunits of Δ FG-xlNup62-58-54 complex together in *E.coli* to obtain a soluble complex for crystallization. Δ FG-Nup54 was present in a His-tagged form on a low copy vector (p15A; 10-12 copies/cell) while Nup62 and Nup58 were present as ZZ-tagged on a bicistronic vector with an intermediate copy number (30-40 copies/cell). On co-expression, Nup54 was present in the soluble fraction. Ni⁺²-affinity purification pulled out a trimeric Δ FG-xlNup62-58-54 complex (*Figure 3.6*). The complex showed a clear 1:1:1 stoichiometry on SDS-gel. Thus, we could obtain large amounts of Δ FG-xlNup62-58-54 complex under native conditions from *E.coli* using co-expression.

Since in these experiments, the amount of soluble Nup54 was limited by the presence of Nup62 and Nup58 (which were present in excess over Nup54 due to solubility-enhancing ZZ-tag and presence on a high copy vector), virtually the entire Nup54 that was pulled out was in complex with Nup62 and Nup58. This was confirmed by gel-filtration and ion exchange chromatography, where protease cleaved Ni⁺² eluent was directly loaded on the chromatography columns. The complex migrated as a single peak both in gel-filtration as well as in ion exchange and no dissociation of any subunit was observed (*Figure 3.6b, 3.6c*). This is consistent with the published data that native Nup62-58-54 complex isolated from nuclei is highly stable and can tolerate conditions such as 2M urea and detergent treatment indicating tight protein-protein interactions (Finlay et al., 1991). Thus, combining co-expression with on-column protease elution, we could reconstitute pure recombinant Δ FG-xlNup62-58-54 complex in *E.coli* in one-step purification. In cases where the complex was used for crystallographic analysis, gel filtration and ion exchange chromatography steps were included to obtain highest level of purity.

3. Results

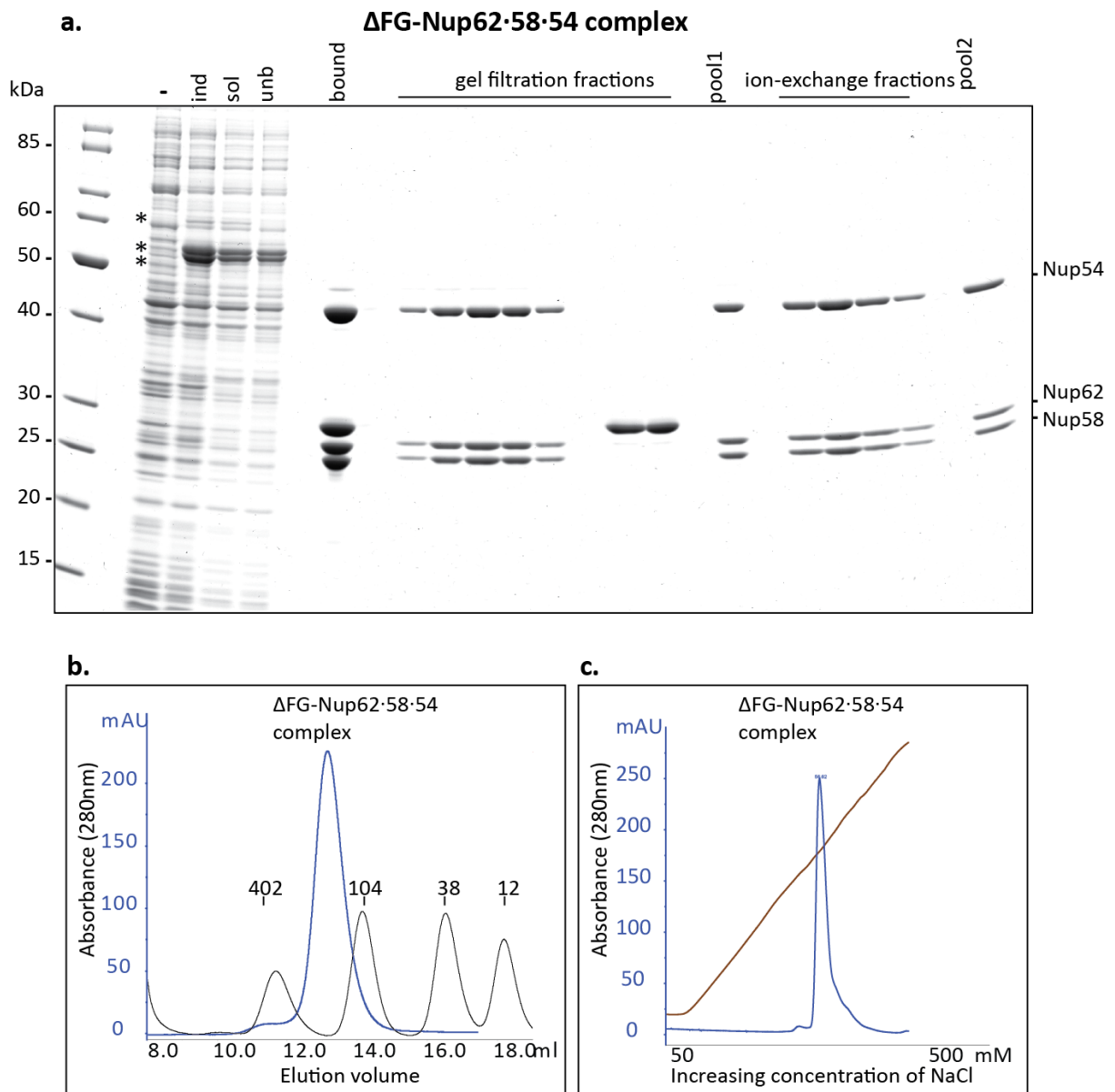


Figure 3.6: Purification of Δ FG-xlNup62·58·54 complex

a) Nickel pull down of Δ FG-xlNup62·58·54 complex from *E. coli* lysate using affinity tag on Δ FG-Nup54. The complex eluted after on-column protease cleavage shows stoichiometric amounts of Nup62, Nup54 and Nup58. The cleaved off ZZ-SUMO tag was separated from Δ FG-xlNup62·58·54 complex by gel-filtration. The complex also migrated as a single peak on ion-exchange column. **b)** Gel-filtration profile of Δ FG-xlNup62·58·54 complex showing the peak fraction wrt molecular weight standard. The complex migrates at a position expected for 230kDa protein complex, which is more than double its theoretical MW 94.7kDa, indicating a highly elongated complex. **c)** Ion-exchange profile of Δ FG-xlNup62·58·54 complex showing single peak elution of the complex.

3.1.6. Δ FG-xlNup62·58·54 is an elongated complex with 1:1:1 stoichiometry

During gel filtration chromatography, Δ FG-xlNup62·58·54 complex eluted much prior to its theoretical molecular weight. The complex migrates at a size of roughly 230kDa compared to its theoretical MW of 95kDa (assuming single copy of each subunit is present in the complex). This migration behavior indicated that either the native state of the complex is non-monomeric

3. Results

where more than one copy of each subunit is present or the complex folds into a highly elongated non-spherical structure.

In order to resolve this issue, we carried out size exclusion chromatography coupled Multi-angle light scattering and dynamic light scattering (SEC-MALS-DLS) analysis of Δ FG-xINup62·58·54 complex. The Δ FG-xINup62·58·54 complex eluate from a gel filtration column was directly analyzed for its light scattering properties at different angles using different LS detectors in an online mode. This led to determination of absolute molar mass, root-mean square radius and hydrodynamic radius of the complex (*Figure 3.7*). The Δ FG-xINup62·58·54 complex eluted as a single monodisperse peak (polydispersity of 1) from the gel filtration column indicating a homogenous species in terms of the MW. Mn (number average molar mass), Mw (weight average molar mass) and Mz (Z-average molar mass) for this peak are listed in the Table 2. The average molecular weight (M_{avg}) of this peak was estimated to be 95.2 ± 4.7 kDa. The value fits the scenario where a single copy of each subunit (sum of MW of subunits of Δ FG-xINup62·58·54 complex - $24.2(\text{Nup62}) + 25.8(\text{Nup58}) + 44.7(\text{Nup54})$ equals 94.7 kDa) is present. The error limits in the predicted MW, thus allow only one copy of each subunit (i.e. 1:1:1 subunit stoichiometry) in the complex and any other stoichiometry would result in a MW far outside the error range of the measurement. These results also confirm the stoichiometry of 1:1:1 seen on SDS PAGE during purification of Δ FG-xINup62·58·54 complex.

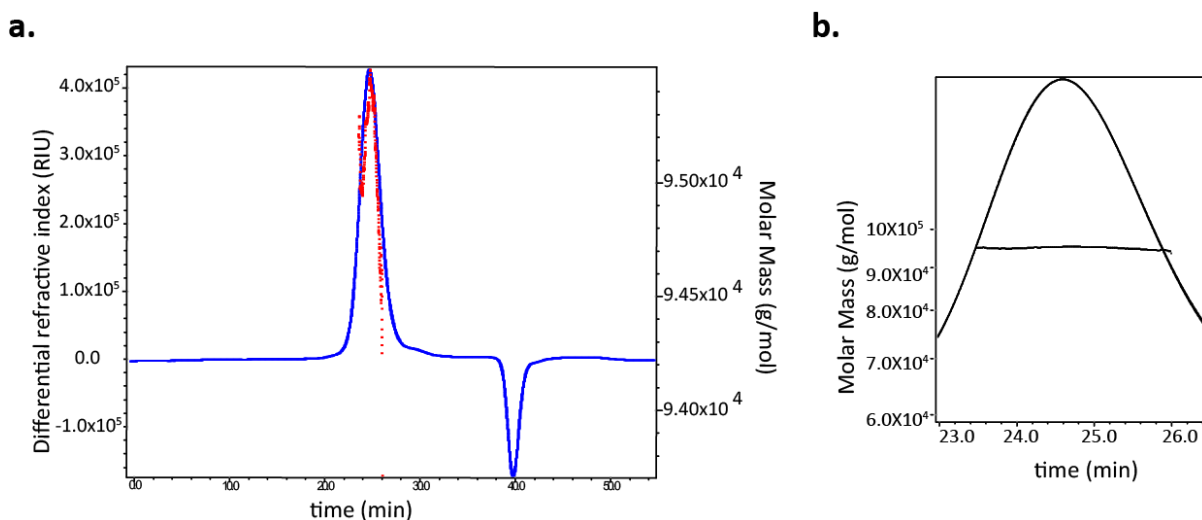


Figure 3.7: Stoichiometry of Δ FG-xINup62·58·54 complex determined by MALS

a) Light-scattering signal from gel-filtration fractions of Δ FG-xINup62·58·54 complex. The complex migrated as a single peak at an estimated MW of 95.2 ± 4.7 kDa. The blue line indicates the UV signal (280nm) and in red is the LS trace recorded from LS detector at 90° angle **b)** Zoomed in version of the peak fraction showing Molar mass distribution. Solid line indicates trace from the RI detector while dashed line in the middle is the MW determined for each slice.

3. Results

Molar mass moments (g/mol)	Root mean square radius (Rms) moments (nm)	Hydrodynamic radius (Rh) moments (nm)
Mn 9.514×10^4 ($\pm 0.134\%$)	Rn 4.6 ($\pm 28.1\%$)	Rh(n) 5.1 ($\pm 4.3\%$)
Mw 9.514×10^4 ($\pm 0.014\%$)	Rw 4.6 ($\pm 28.1\%$)	Rh(w) 5.1 ($\pm 4.3\%$)
Mz 9.514×10^4 ($\pm 0.299\%$)	Rz 4.6 ($\pm 28.1\%$)	Rh(z) 5.1 ($\pm 4.3\%$)
M (avg) 9.523×10^4 ($\pm 0.005\%$)	R (avg) 4.2 ($\pm 1.2\%$)	Rh(avg) 5.0 ($\pm 0.5\%$)

Table 2: Table showing properties of Δ FG-xlNup62-58-54 complex as determined by MALS analysis

Table showing the estimated molecular weight, root mean square radius and hydrodynamic radius for Δ FG-xlNup62-58-54 complex. “n”, number average; “w”, weight-average; “z”, Z-average distribution.

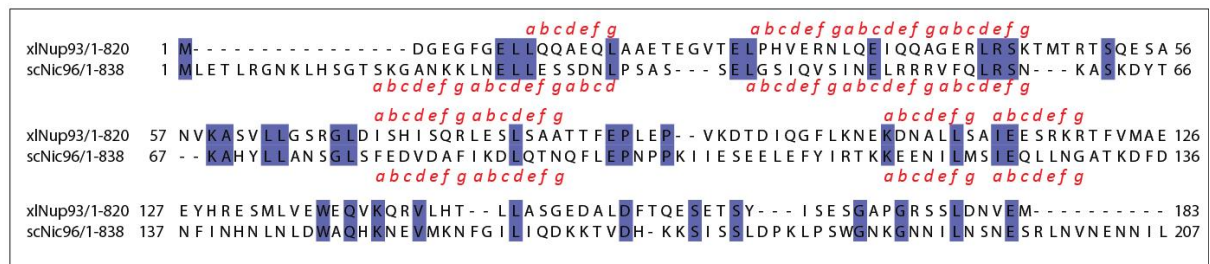
3.1.7. Nup93 interaction with Δ FG-xlNup62-58-54 complex

It has been reported that the N-terminus of scNic96 (the yeast homolog of mammalian Nup93) is directly involved in anchoring Nsp1-49-57 (yeast homolog of mammalian Nup62-58-54) complex to the NPC by coiled-coil interaction (Grandi et al., 1995b) (Bailer et al., 2001). In yeast deletion studies, N-terminal heptad repeats of Nic96 were shown to be responsible for interaction with Nsp1-49-57 complex while its central coiled-coil domains were found to be non essential for this interaction (Grandi et al., 1995b). In order to determine whether a direct interaction between xlNup93 and xlNup62-58-54 complex can be reconstituted in recombinant purified subunits; we co-expressed all four subunits of Δ FG-xlNup62-58-54-93 complex in *E.coli*. His-tag on Nup93 was used to pull out ZZ-tagged Δ FG-xlNup62-58-54 complex on Ni⁺² matrix. When expressed alone, full-length xlNup93 was largely insoluble. Co-expression with Nup62-58-54 complex greatly increased the solubility of Nup93. The quaternary Nup62-58-54-93 complex could be stably expressed in *E.coli* and we could pull out the complex containing all four subunits together. Thus, we could reconstitute the quaternary Δ FG-xlNup62-58-54-93 complex by co-expression in *E.coli* in a similar manner as Δ FG-xlNup62-58-54 complex.

On aligning the sequences of yeast and *Xenopus* Nup93 N-terminus, it was seen that the essential hydrophobic residues forming the heptad repeats are conserved in both organisms (Figure 3.8). In order to map the minimal domain in the N-terminus of xlNup93 (1-180) sufficient for interaction with Nup62-58-54 complex, we created ZZ-tagged constructs of different lengths in Nup93 N-terminus- Nup93N1 (1-82), Nup93N2 (1-95), Nup93N3 (1-157) and Nup93N4 (1-180). We tested these constructs in the co-expression assay with His-tagged Δ FG-xlNup62-58-54 complex and found that similar amounts of Nup93N-terminus were pulled out for Nup93N2 (1-95), Nup93N3 (1-157) and Nup93N4 (1-180) whereas the amount of associated Nup93 N-terminus slightly decreased for Nup93N1 (1-82). Therefore, we used xlNup93N2 (1-95) as the minimal interaction domain for purification of Δ FG-xlNup62-58-54-93n complex (Figure 3.8b).

3. Results

a.



b.

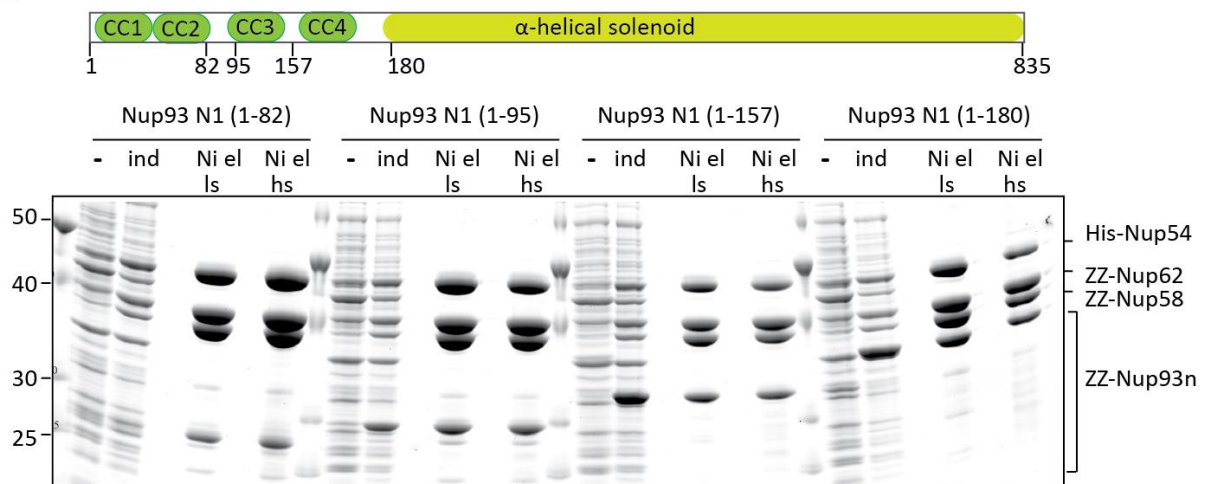


Figure 3.8: Mapping N-terminus of xlNup93 for interaction with ΔFG-xlNup62·58·54 complex

a. Sequence alignment of *X. laevis* and *S. cerevisiae* Nup93 N-terminus. The conserved amino acids are shown in blue. The residues forming the heptad repeats of coiled-coil domains in both species are indicated from a-g (red). **b.** Deletion mapping of xlNup93 N-terminus for interaction with ΔFG-xlNup62·58·54 complex. The constructs were created according to coiled-coil domain predictions (green) in xlNup93 N-terminus and co-expressed with ΔFG-xlNup62·58·54 complex in *E. coli*. The complex containing ZZ-tagged Nup93 N-terminus was pulled out using His-tag on Nup54. Based on this, the length of Nup93 N-terminus sufficient for ΔFG-xlNup62·58·54 complex interaction was determined to be residues 1-95. “-”, uninduced; “ind”, induced; “Ni el”, Imidazole eluate after binding Ni⁺² binding; “ls”- low salt (50mM NaCl) buffer, “hs”- high salt buffer (500mM NaCl).

3.1.8. xlNup93 N-terminus has weak affinity for ΔFG-xlNup62·58·54 complex

During purification of ΔFG-xlNup62·58·54·93n complex, Nup93n was often found to be in a sub-stoichiometric amount than the rest of ΔFG-xlNup62·58·54 complex. When a complex containing excess of Nup93n over Nup62·58·54 complex was loaded on a gel filtration column, we observed that majority of Nup93n dissociated and ran as a separate peak and only a minor fraction migrated along with the ΔFG-xlNup62·58·54 complex (Figure 3.9a). In order to determine if this was the actual stoichiometry of the Nup62·58·54·93n complex or a result of weak affinity of Nup93n for ΔFG-xlNup62·58·54 complex, we loaded the peak fraction from the gel-filtration chromatography on a second gel-filtration column.

3. Results

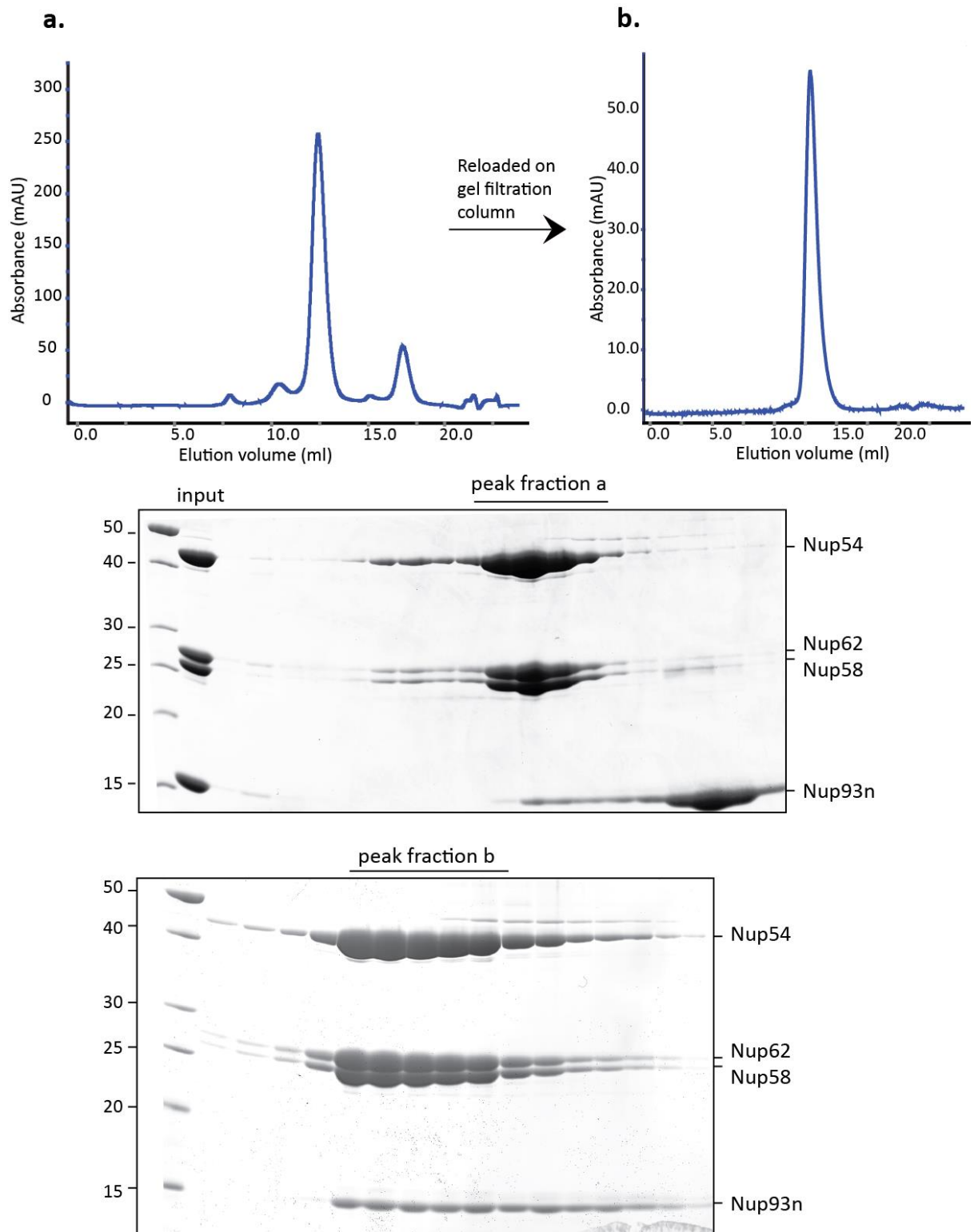


Figure 3.9: xlNup93N-terminus has weak affinity for Δ FG-xlNup62-58-54 complex

a. Gel filtration profile of Δ FG-xlNup62-58-54-93n complex. Δ FG-xlNup62-58-54-93n complex containing excess of Nup93n (1-95) was loaded on the gel filtration column. As seen in the corresponding SDS gel picture, Nup93n dissociated and migrated separately from the rest of the complex. **b.** Reloading the peak fraction from (a) on gel filtration column. xlNup93n, though sub-stoichiometric in amount, still dissociated from the rest of the complex, indicating a weak interaction.

3. Results

Again we observed a similar behavior that the remaining Nup93N also spontaneously dissociated from Δ FG-xlNup62·58·54 complex (Figure 3.9b). Therefore, we could conclude that xlNup93 N-terminus has a weak affinity for Δ FG-xlNup62·58·54 complex. When we compared the coiled-coil forming probability of vertebrate Nup93 (xlNup93 and hsNup93) with scNic96 and ctNic96, we observed that Nup93 N-terminus in fungi has stronger coiled-coil domain formation propensity compared to xlNup93. Therefore, we predict that there may be additional interaction sites on vertebrate Nup93 besides the N-terminus, responsible for anchoring the Δ FG-xlNup62·58·54 complex in the NPC.

The weak affinity of xlNup93n for Δ FG-xlNup62·58·54 complex was also analyzed by surface plasmon resonance (SPR) where xlNup93n was immobilized on a solid surface and Δ FG-xlNup62·58·54 complex was allowed to bind it at different protein concentrations. The K_d value for this interaction was found $1.77\mu\text{M}$ (Figure 3.10). Thus, we could confirm that low values of affinity of Nup93 N-terminus for Nup62·58·54 complex accounted for losing Nup93n during purification of Nup62·58·54·93n complex. However, so far we could not purify full-length xlNup93 due to its poor solubility in *E.coli*. It would be interesting to compare the affinity of full-length Nup93 with the coiled-coil domains at its N-terminus to determine if additional high affinity sites for Nup62 complex binding exist on Nup93.

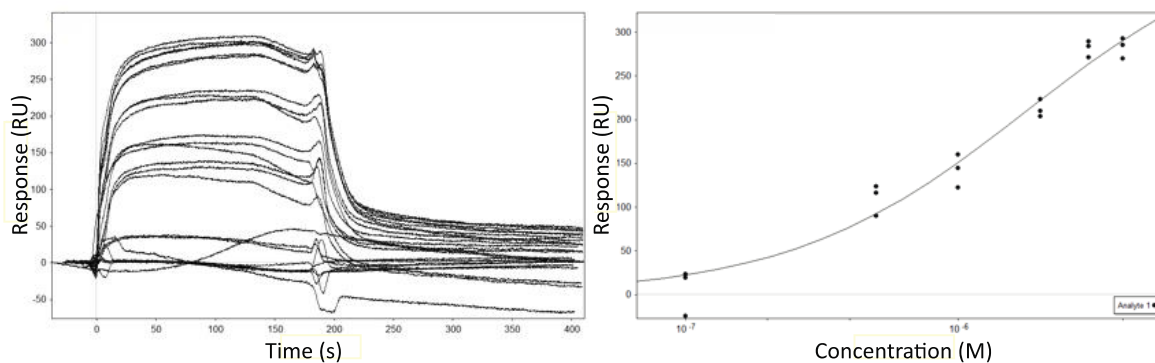


Figure 3.10: Affinity measurement for Nup93 N-terminus interaction with Nup62·58·54 complex

Surface plasmon resonance analysis was performed, and the sensograms were obtained from injection of Nup62·58·54 complex at five concentrations (each in triplicate) over surface immobilized Nup93 N-terminus. The complex was injected for 180 s and dissociation was monitored for 300 s. The K_d value for interaction was determined to be $1.77\mu\text{M}$.

3.1.9. Crystallization attempts for Δ FG-xlNup62·58·54 complex

Crystallization trials for Δ FG-xlNup62·58·54 and Δ FG-xlNup62·58·54·93n complex did not yield any crystals despite broad range of conditions tested (~ 4000 conditions for each). The failure in crystallization could be attributed to inherent flexibility of these complexes caused by internal linker regions or disordered ends of the subunits (Figure 3.2). To get Nup62·58·54 complex in a

3. Results

more rigid form, we immunized camelids with the Δ FG-xINup62-58-54-93n complex and obtained nanobodies against it. The sections below describe the expression, isolation and screening of nanobodies against Δ FG-xINup62-58-54-93n and crystallization trials of nanobody conjugated Δ FG-xINup62-58-54 complex.

3.1.10. Nanobodies (Nbs) against Δ FG-xINup62-58-54 complex

Nanobodies (Nbs) are 'heavy chain only antibodies' occurring in the family Camelidae that includes camels, llamas, guanacos, alpacas and vicuñas. The antigen-binding fragment in these single-domain antibodies comprises of only the variable domain of the heavy chain (referred to as VHH or variable domain of the heavy chain of heavy-chain antibody) instead of two domains (VH, for variable domain of heavy chain and VL, for variable domain of light chain) found in conventional antibodies (*Figure 3.11a*). These single domain fragments have been reported to express really well in bacteria and have been used in successful crystallization and structure determination of difficult proteins (Chung et al., 2011) (Lam et al., 2009).

Since attempts to crystallize Δ FG-xINup62-58-54-93n on its own failed, we raised nanobodies against it in order to find small high affinity binding partners that could stabilize the complex and aid in crystallization. High flexibility in the linker regions connecting the coiled-coil modules might be one of the reasons for not being able to crystallize the complex. Since nanobodies have been shown to make protein complexes rigid by binding to flexible conformations (Lam et al., 2009), they might be good candidates for co-crystallization of Nup62-58-54-93n complex.

We obtained 29 nanobody sequences against Δ FG-xINup62-58-54-93n complex, which were cloned into bacterial expression vectors. Since nanobodies contain disulfide bonds as a part of their core framework, periplasmic expression in *E.coli* was used as a method of choice for their isolation. The purified nanobodies were then analyzed for the binding sites on Nup62-58-54-93n complex.

3. Results

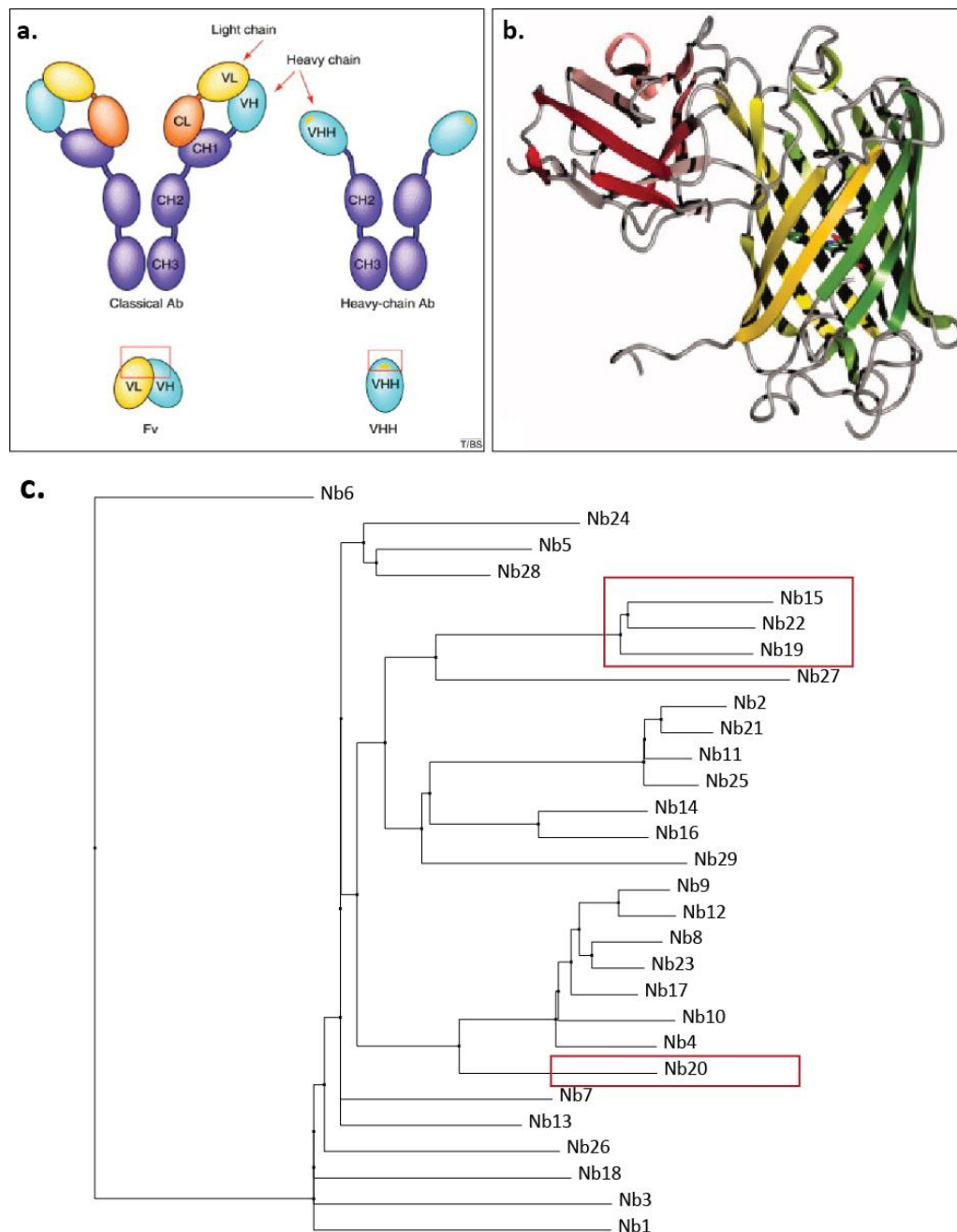


Figure 3.11: Nanobodies against Δ FG-xINup62·58·54 complex

a. Schematic comparison of domain structure in classical antibodies with heavy chain antibodies found in camelids. Nanobodies are variable domains of heavy chain antibodies (VHH) harboring the entire antigen-binding site (Adapted from Muyldermans, 2001) **b.** Crystal structure of GFP-GFP nanobody complex showing nanobody (red) binding to GFP (green) via CDR loops (Kubala *et al*, 2010) **c.** Phylogenetic tree from multiple sequence alignment of 29 nanobodies obtained against xINup62·58·54-93n complex. The nanobodies recognizing the trimeric xINup62·58·54 complex only are shown in red boxes (explained in section 3.1.13).

3.1.11. Δ FG-xINup62·58·54 complex recognition by the nanobodies

28 out of 29 nanobodies that could be successfully purified after bacterial periplasmic expression were tested for in-vitro binding to the purified Δ FG-xINup62·58·54 and Δ FG-xINup62·58·54-93n

3. Results

complexes. 24 nanobodies were found to recognize and pull out the Δ FG-xl Nup62·58·54 complex (Figure 3.12). All nanobodies that bound Δ FG-xlNup62·58·54 complex pulled out similar amounts of Δ FG-xlNup62·58·54 complex (except Nb24 and Nb27, which bound weakly). We did not find any nanobody that could stabilize the interaction between xlNup93n and Δ FG-xlNup62·58·54 complex in this analysis. In order to find suitable candidates for crystallization, the positive nanobodies were screened for their binding sites on Δ FG-xlNup62·58·54 complex. This was done in the pull-down experiments by incubating nanobodies with individual subunits of the Δ FG-xlNup62·58·54 complex (Nup54 α/β domain, Nup62) or with the binary complexes (Nup62·58 and Nup62·54)

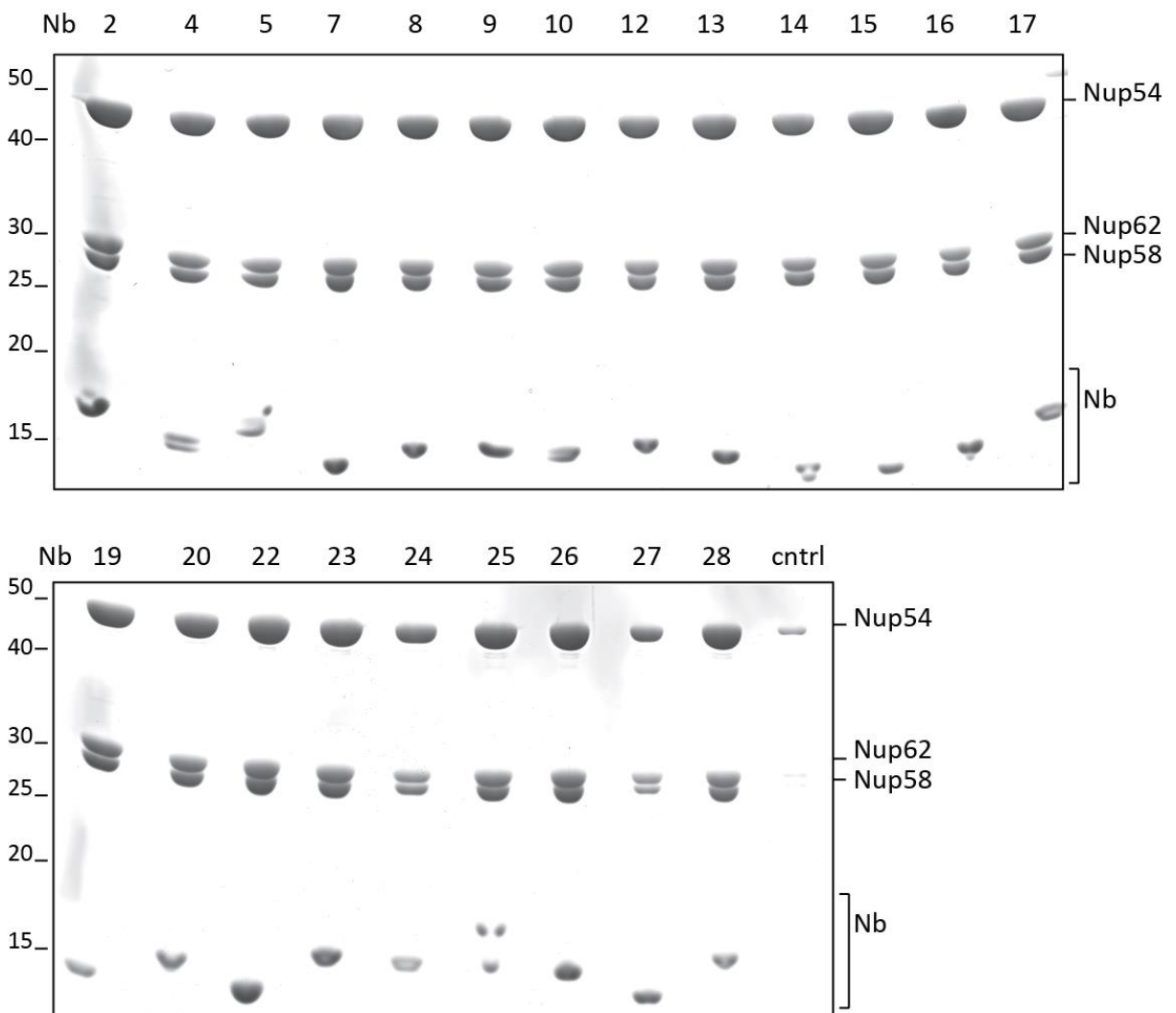


Figure 3.12: All nanobodies bind Δ FG-xlNup62·58·54 complex with high affinity

Binding assays showing pull-down of Δ FG-Nup62·58·54 complex with individual nanobodies. Nanobodies were mixed with equimolar amounts of Δ FG-Nup62·58·54 complex and incubated for 1h at room temperature. The complex was pulled down on Ni^{+2} beads using C-terminal His-tag on the nanobodies. Equal amounts of nickel eluate from each pull down were loaded on SDS gel. Each Nb, except Nb24 and Nb27, pulled out similar amount of Δ FG-xlNup62·58·54 complex.

3. Results

3.1.12. Most nanobodies recognize Nup54 α/β domain in the Δ FG-xlNup62·58·54 complex

In the screening for binding sites of Nbs on Δ FG-xlNup62·58·54 complex, 18 out of 24 nanobodies recognized and pulled out α/β domain in Nup54 (Nup54 146-372). Thus, most of the nanobodies we obtained from immunization of Nup62·58·54 complex were directed against Nup54 α/β domain (Figure 3.13).

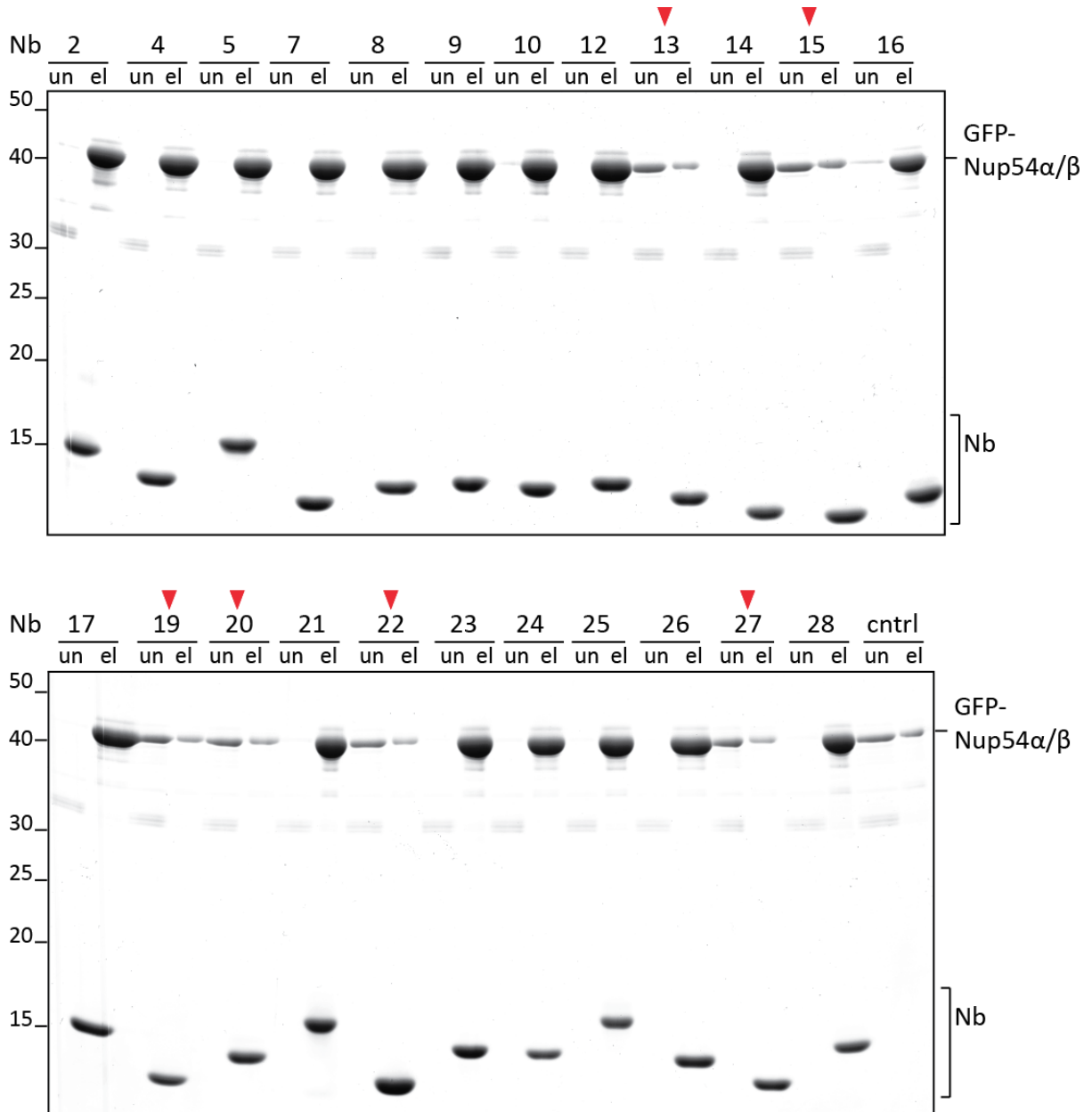


Figure 3.13: Most nanobodies recognize xlNup54 α/β domain in the Δ FG-xlNup62·58·54 complex

Binding assays showing recognition of Nup54 alpha-beta domain by most nanobodies. The nanobodies were incubated with GFP-Nup54 α/β domain and the complex was pulled out using His-tag on nanobodies on nickel beads. Unbound fractions (un) and nickel eluent (el) were loaded adjacent to each other on the SDS gel. For most nanobodies GFP-Nup54 α/β domain was found in the eluent fractions. Lanes marked with arrows show nanobodies that do not recognize Nup54 α/β domain (Nb 13, 15, 19, 20, 22 and 27) and are interesting candidates for crystallization. (Elate is 5X more concentrated than the unbound fraction).

3. Results

The reason for this can be higher immunogenicity of xINup54 α/β domain compared to the coiled-coil domains of Δ FG-xINup62·58·54 complex or selection bias for Nup54 α/β domain during phage display. We tested whether Nup54 α/β domain can be separated from the coiled coil regions of Δ FG-xINup62·58·54 complex and purified on its own. Although highly soluble, Nup54 α/β domain heavily associates and co-purifies with *E.coli* chaperones (data not shown).

This indicated presence of exposed hydrophobic sites on Nup54 α/β domain that might be shielded during its association with the coiled-coil domains of Nup62·58·54 complex. We could separate the associated *E.coli* chaperones from the Nup54 α/β domain on an anion-exchange column, but the resulting protein aggregated heavily, thus confirming that Nup54 α/β domain physically interacts with the rest of the Nup62·58·54 complex.

3.1.13. Nanobody 15 recognizes the trimeric coiled-coil interface of Δ FG-xINup62·58·54 complex

Four nanobodies (Nb15, 19, 20 and 22) were found to recognize exclusively the ternary coiled-coil interface of Δ FG-xINup62·58·54 complex. These nanobodies failed to bind the individual subunits (Nup62, Nup54 α/β domain) or the binary complexes (Nup62·58 or Nup62·54). Thus, they were specific for their binding to a coiled-coil interface that is formed only when all three subunits Nup62, Nup58 and Nup54 are present.

Such conformation specific nanobodies are ideal candidates for co-crystallization because they stabilized the trimeric complex state of the Nup62·58·54 complex over the individual subunits. Also, since they themselves have a rigid fold, these nanobodies will further reduce the flexibility of coiled-coil domains improving chances of its crystallization.

Sequence analysis of these trimeric coiled-coil conformation-specific nanobodies revealed that they belonged to two distinct classes. Nb15, 19 and 22 (trimeric coiled-coil interface class I) were very similar in their sequence and bound to the same site on the Δ FG-xINup62·58·54 complex. This was observed in a binding assay where these nanobodies competed against each other for binding to Δ FG-xINup62·58·54 complex. Nb20 on the other hand, had a distinct sequence from Class I nanobodies and recognized a second site on Δ FG-xINup62·58·54 complex. These two classes of nanobodies were also compatible for simultaneous binding to Nup62·58·54 complex and could form a pentameric Nb15-Nb20-Nup62·58·54 complex (*Figure 3.16e*). *Figure 3.14* shows a comparison between recognition sites of Nb15 (representative of trimeric coiled-coil interface binding Nbs) and Nb2 (representative of Nup54 α/β domain binding Nbs). While Nb15 solely recognizes trimeric state of Nup62·58·54 complex, Nb2 binds Nup62·58·54 as well as Nup62·54 and Nup54 α/β domain.

3. Results

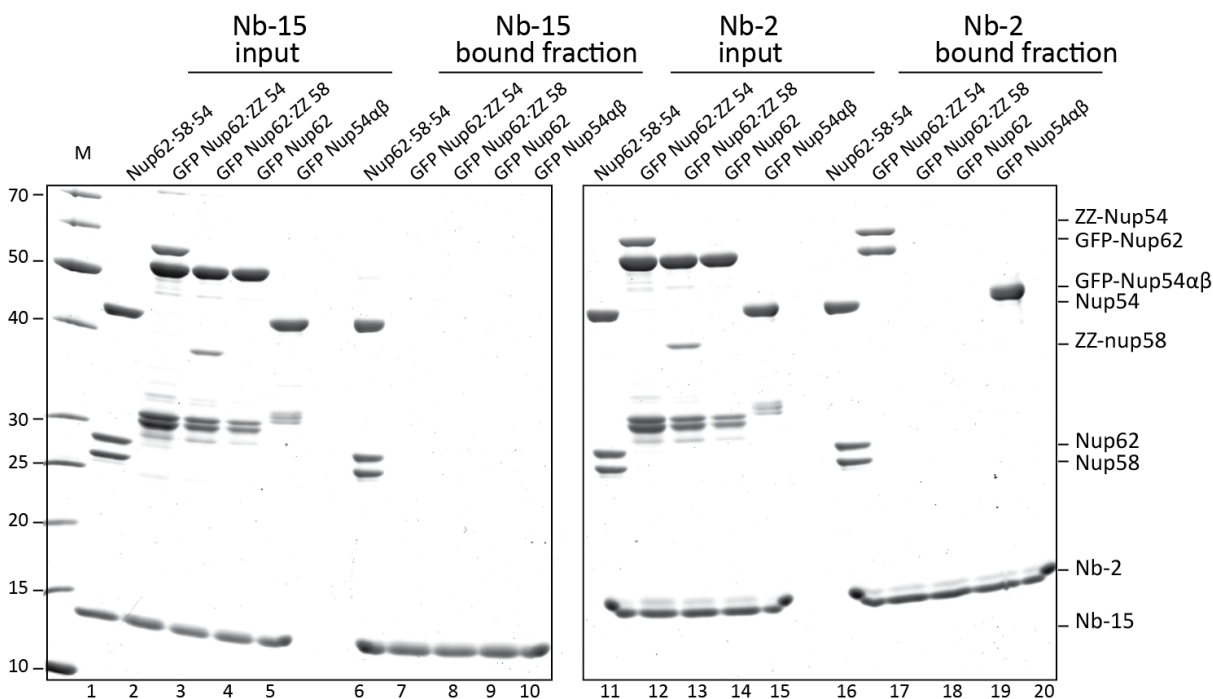


Figure 3.14: Nb15 exclusively binds trimeric coiled-coil interface in ΔFG -xlNup62-58-54 complex

Binding assays of Nb15 (representative of trimeric coiled-coil domain binding Nbs) and Nb2 (representative of Nup54 α/β domain binding Nbs) with trimeric Nup62-58-54 complex, binary Nup62-58 and Nup62-54 complexes, monomeric Nup62 and Nup54 α/β domain. Input fractions are loaded on the left while corresponding nickel eluent fractions bound to nanobodies are loaded on the right in each gel. Nb15 binds and pulls out only the trimeric Nup62-58-54 complex (lane 6), while Nb2 like all other alpha-beta domain binding nanobodies, binds Nup54 α/β domain (lane 20) and therefore, also pulls out trimeric Nup62-58-54 complex (lane 16) and binary Nup62-54 complex (lane 17) but not Nup62-58 (lane 18) or Nup62 (lane 19).

3.1.14. Crystallization of Nanobody conjugated ΔFG -xlNup62-58-54 complex

We purified nanobody-conjugated ΔFG -xlNup62-58-54 complexes in large-scale preparations for crystallization trials. In brief, ΔFG -xlNup62-58-54 was purified as before by coexpression followed by on-column protease elution. Chromatographic purification with gel-filtration and ion exchange columns further polished the complex. *E. coli* periplasmically expressed Nbs containing C-terminal His₆ tag were purified via Nickel chromatography and gel-filtration to remove aggregates and impurities. Purified Nup62-58-54 complex and Nbs were mixed in 1:1.5 molar ratio and incubated for 1h at RT. The complex between Nb and Nup62-58-54 complex was pulled out by Ni⁺² affinity chromatography. Excess of nanobody was separated from complex by gel-filtration, where Nup62-58-54 complex bound to Nb migrated as a single peak while excess of the free Nb came out in later elution fractions.

3. Results

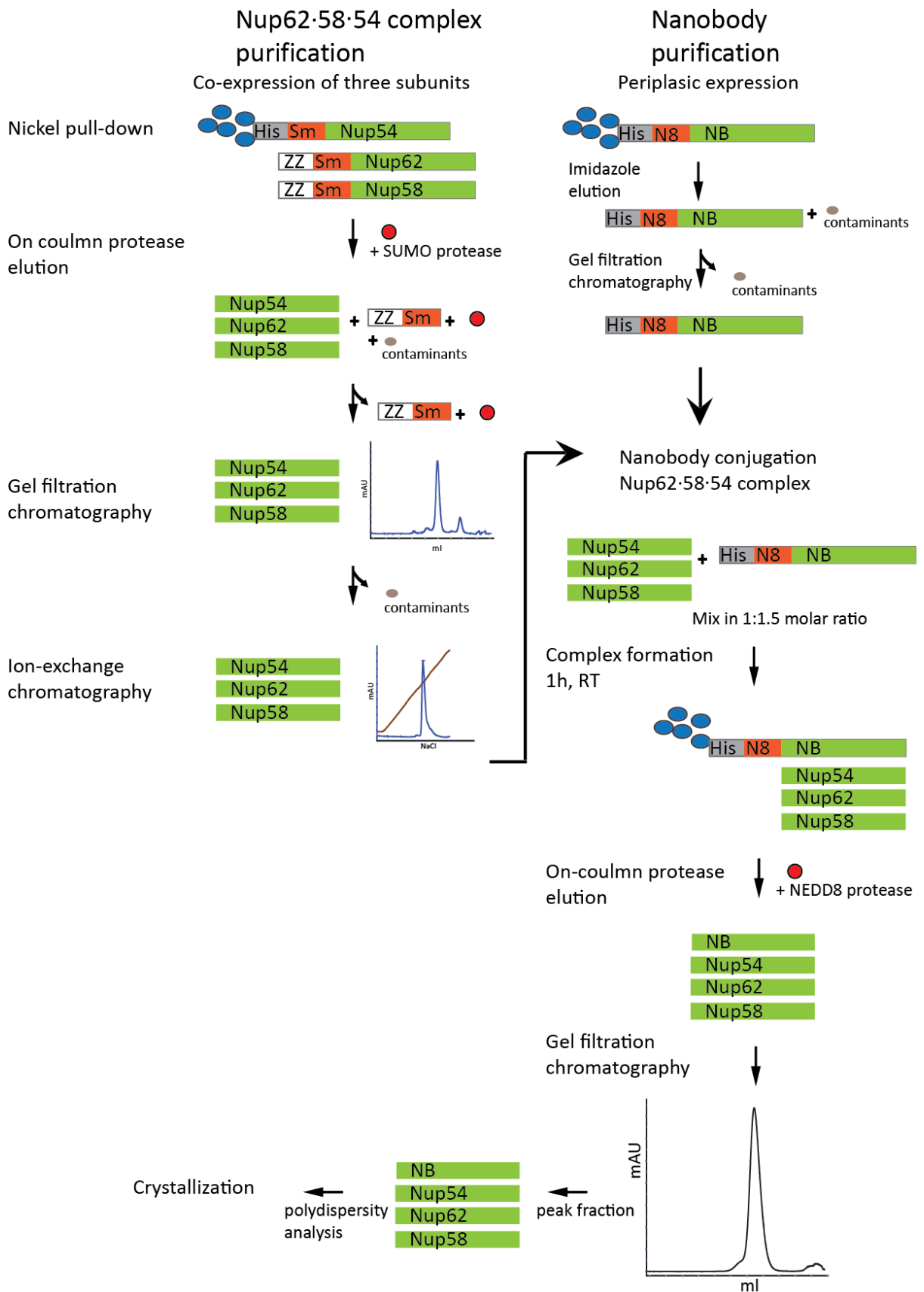


Figure 3.15: Scheme for large-scale purification of nanobody containing Δ FG-xlNup62·58·54 complex

See text for details

3. Results

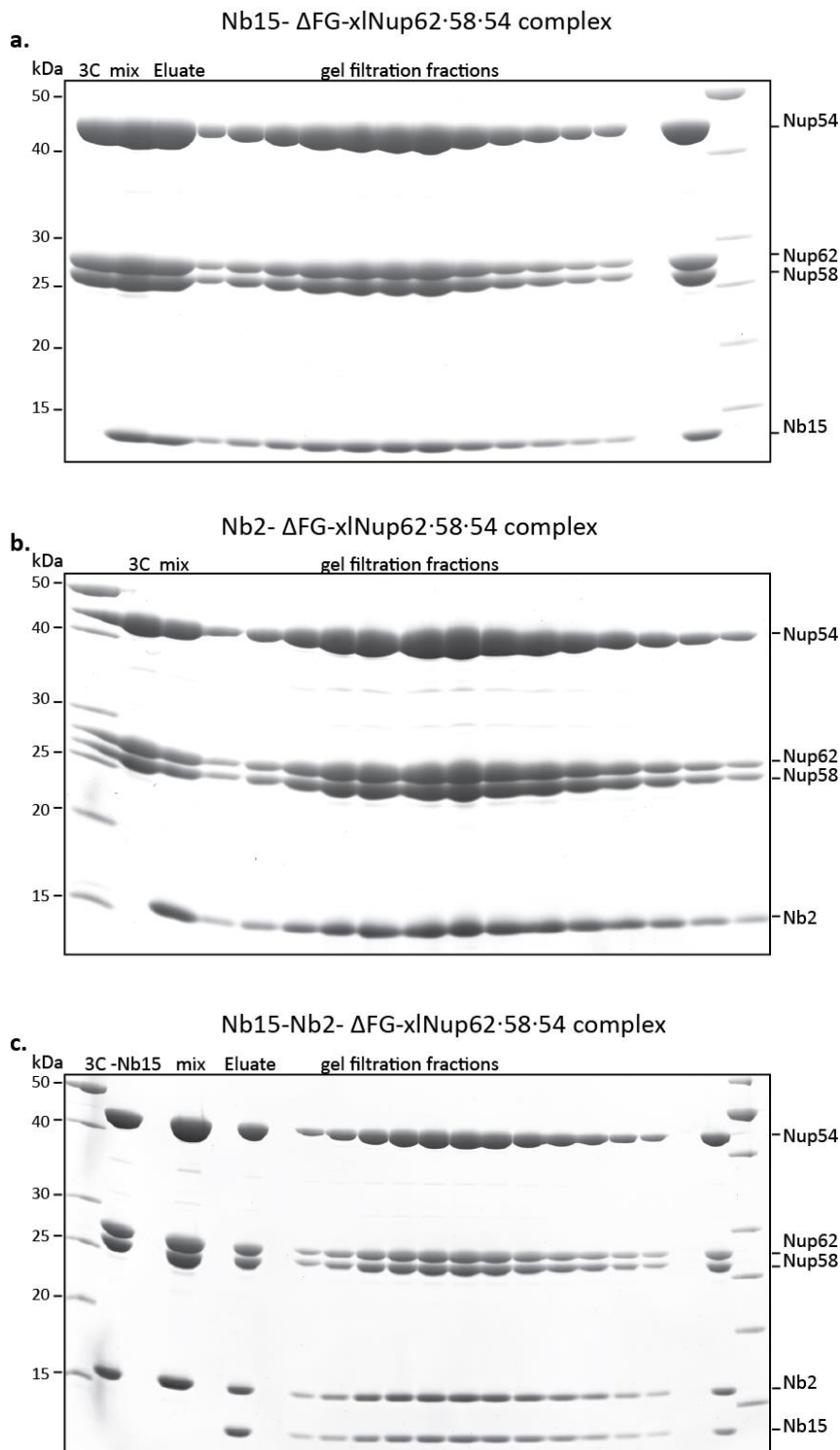
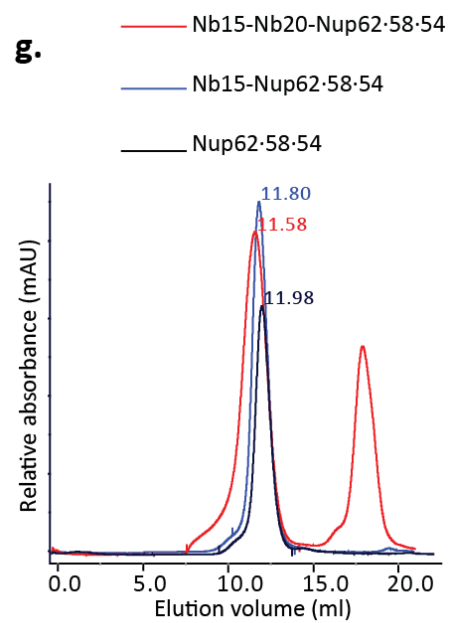
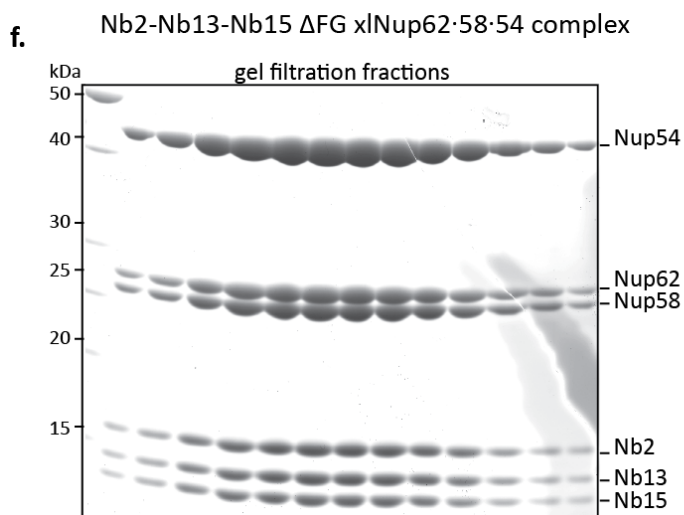
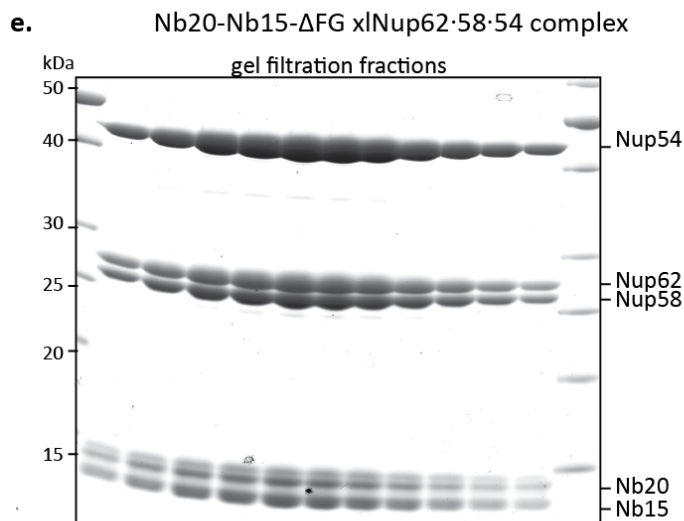
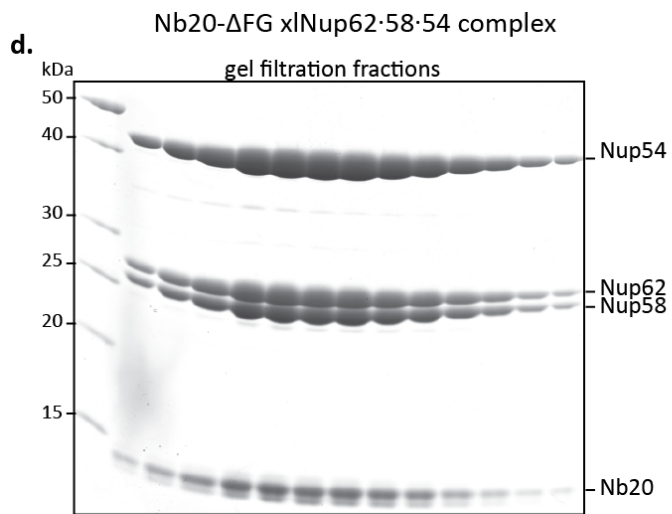


Figure 3.16: Various nanobody - Δ FG-xINup62·58·54 complexes screened in crystallization trials

Δ FG-xINup62·58·54 complex conjugated to **a.** Nb15 (binds trimeric coiled-coil interface Class I) **b.** Nb2 (binds Nup54 α/β domain) **c.** Nb15 and Nb2 **d.** Nb20 (binds trimeric coiled-coil interface Class II). **e.** Nb15 and Nb20 **f.** Nb15, Nb2 and Nb13 (binds Nup62·54 interface). Gel filtration fractions after preparatory scale complex formation were loaded on SDS gel. 3C- input Nup62·58·54 complex, mix- Stoichiometric amounts of nanobody and Nup62·58·54 complex incubated for conjugation, El- Nup62·58·54 complex pulled out by nanobody on Nickel column **g.** Overlay of gel filtration profiles of Nup62·58·54 complex, Nup62·58·54 complex conjugated to single nanobody and Nup62·58·54 complex conjugated to two nanobodies showing corresponding shifts in the elution volume with increasing size of the overall complex.

3. Results



Continued from previous figure

3. Results

The quality of Nb conjugated Nup62·58·54 complexes were determined by polydispersity measurement in QELS (Quasi-elastic light scattering). The fractions with polydispersity <15% were used for crystallization trials (*Figure 3.15*). For conjugation of two-or more Nbs to the Nup62·58·54 complex, the gel-filtration fractions from first conjugation were incubated again with the second nanobody and complex was directly applied to gel-filtration.

The nanobodies screened in the crystallization trials included Nb15 (trimeric coiled-coil interface, Class1), Nb20 (trimeric coiled-coil interface, Class2), Nb2 (Nup54 α/β domain binder) and Nb13 (Nup54·62 binder). In addition, the pentameric complexes Nb15-Nb20 Nup62·58·54, Nb15-Nb2 Nup62·58·54, and Nb15-Nb13 Nup62·58·54 and hexameric Nb15-Nb13-Nb2 Nup62·58·54 were also tested (*Figure 3.16*).

Nb20 that belonged to the second class of ternary coiled-coil interface binding nanobodies had a longer CDR3 loop compared to Class 1 nanobodies and contained an inter-loop disulfide bond between CDR2 and CDR3. This disulfide bond was found to be extremely reduction-sensitive and could be easily reduced in Nup62·58·54 complex bound Nb20. Thus, Nb20 was not a good candidate for crystallization even though it stabilized the ternary coiled-coil interface of Nup62·58·54 complex on account of heterogeneity introduced by its reduction-sensitive behavior.

A full crystallization sparse-matrix screen comprising of ~4000 crystallization conditions was tested for crystal formation. For Nb15-Nup62·58·54 complex, two hits obtained under two different conditions showed rod-shaped crystals (*Figure 3.17*). The crystals from one condition (37% Polypropylene-glycol 400, 15% ethanol) could be optimized in manual drop format. The diffraction from these crystals was however limited to 7.5Å resolution.

3.1.15. Attempts to improve crystal diffraction

Various standard methods used in protein crystallography were tested to improve crystal diffraction from Nb15- Δ FG-xINup62·58·54 complex crystals. Additive screening (addition of small compounds in the crystallization drop), crystal dehydration (by transferring crystals into wells containing increasing concentrations of precipitants) (Heras and Martin, 2005) etc. was tried. However, no improvement in diffraction was observed with these methods.

3. Results

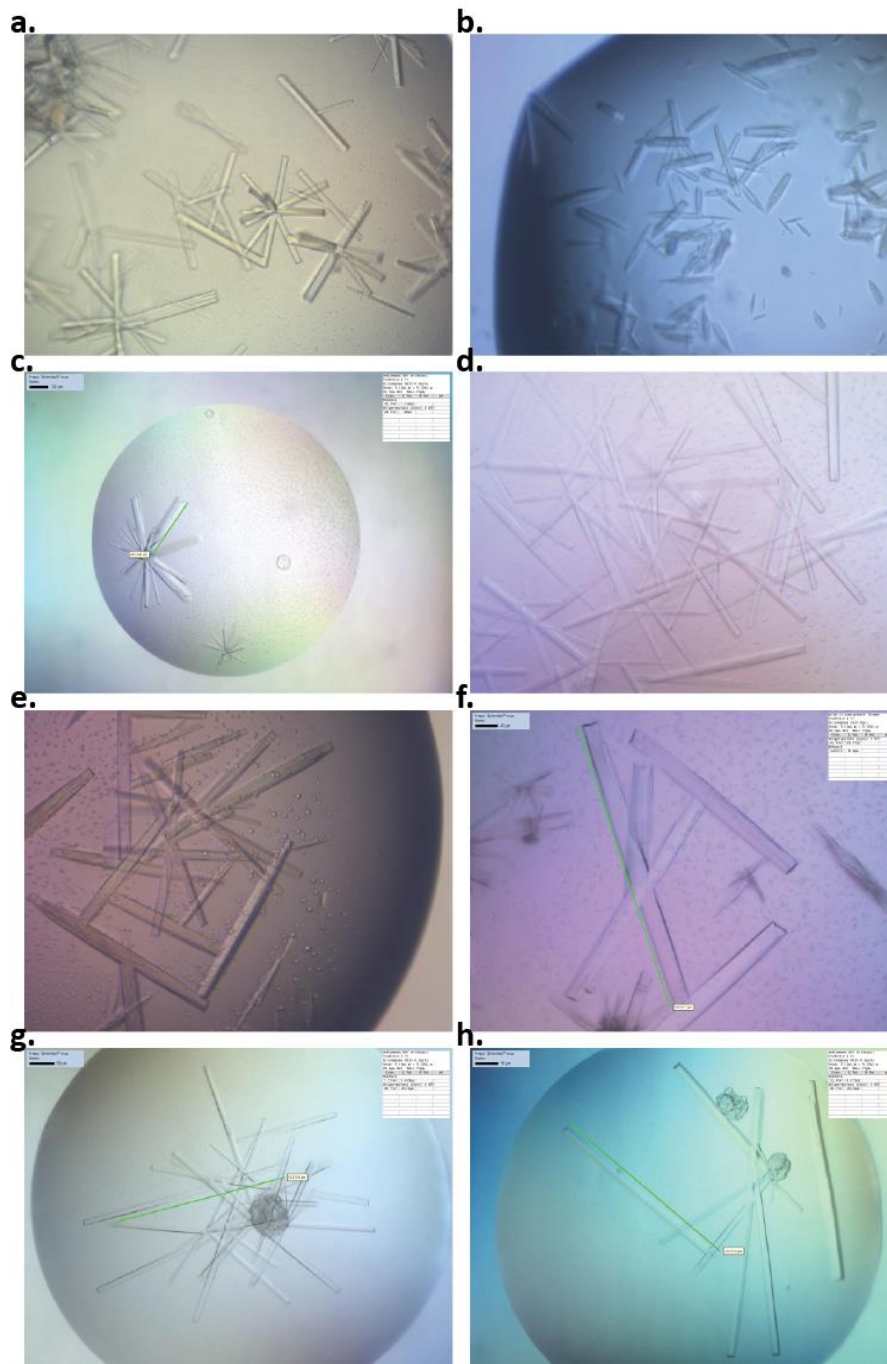


Figure 3.17: Rod-shaped crystals of Nb15- Nup62-58-54 complex

Rod-shaped crystals of Nb15- Δ FG-xNup62-58-54 complex grown under different conditions. The crystals appeared after 2 days and grew to their maximum size in 1 week. The diffraction from these crystals, however, was limited at 7.5-Å resolution.

3.1.16. Limited proteolysis of Δ FG-xNup62-58-54 complex

Limited proteolysis followed by mass-spectrometry is widely used approach in protein crystallization that is used to identify stable protein domains. Stable protein domains resistant to proteases tend to result in better-diffracting crystals. It is known that conformationally flexible or solvent exposed regions are favored sites of protease action leaving the rigid domain of

3. Results

protein intact (Cohen et al., 1995). In order to define the boundaries of interacting coiled-coil domains in Δ FG-xlNup62·58·54 complex assuming that the interface regions remain intact during protease digestion, we performed limited proteolysis of Δ FG-xlNup62·58·54 and Δ FG-xlNb15-Nup62·58·54 complex with chymotrypsin. Chymotrypsin was chosen since it has a broad-range of specificity compared to other proteases such as GluC. It cleaves at the C-terminal end of aromatic residues Tyr, Trp, Phe and Leu but also shows some cleavage at the C-terminal side of hydrophobic amino acids such as Met and Ile. Different time-points of chymotrypsin digestion over different substrate to protease amounts (1:100, 1:1000, 1:10,000 w/w) were screened to obtain a stable protease digestion pattern (*Figure 3.18a*). After optimizing the conditions for protease digestion, the pattern of stable bands that appeared is shown. Mass spectrometry analysis of digested fragments revealed that the C-terminal domains in all three subunits were sites of protease digestion (*Figure 3.18c*). These ends might be the regions where Nup93 interacts Nup62·58·54 complex. In the absence of interaction partner (as Nup93n dissociated from Nup62·58·54 complex during purification), the flexible domains become available for protease action. The Nb15 conjugated Nup62·58·54 complex showed slightly different pattern for Nup54 digestion compared to the native complex (*Figure 3.18b*). Nanobody 15 binding might shield conformationally flexible regions on Nup54, protecting it from protease action.

Next, large-scale preparation of Nb15-Nup62·58·54 complex was subjected to chymotrypsin digestion under optimized conditions. The digested complex was analyzed on gel-filtration to separate the protease-resistant complex. We observed a main peak comprising of Nb15-Nup62·58·54 fragments running at size slightly smaller than the native complex. This was preceded by a pre-peak fraction, which were essentially the oligomers of Nup62·58·54 complex in a non-stoichiometric amount. Also, the two peaks could not be completely separated on gel-filtration column (*Figure 3.19*). Thus, cleaving Nb15-Nup62·58·54 complex by the proteases induced oligomerization of the digested fragments indicating that the complex gets unstable on protease cleavage. Mass spectrometric analysis of digested fragments in the main peak revealed that mainly the C-terminal regions of all three subunits were cleaved off indicating that these might be the weak affinity sites in the complex, whose cleavage induces oligomerization of the remaining complex.

3. Results

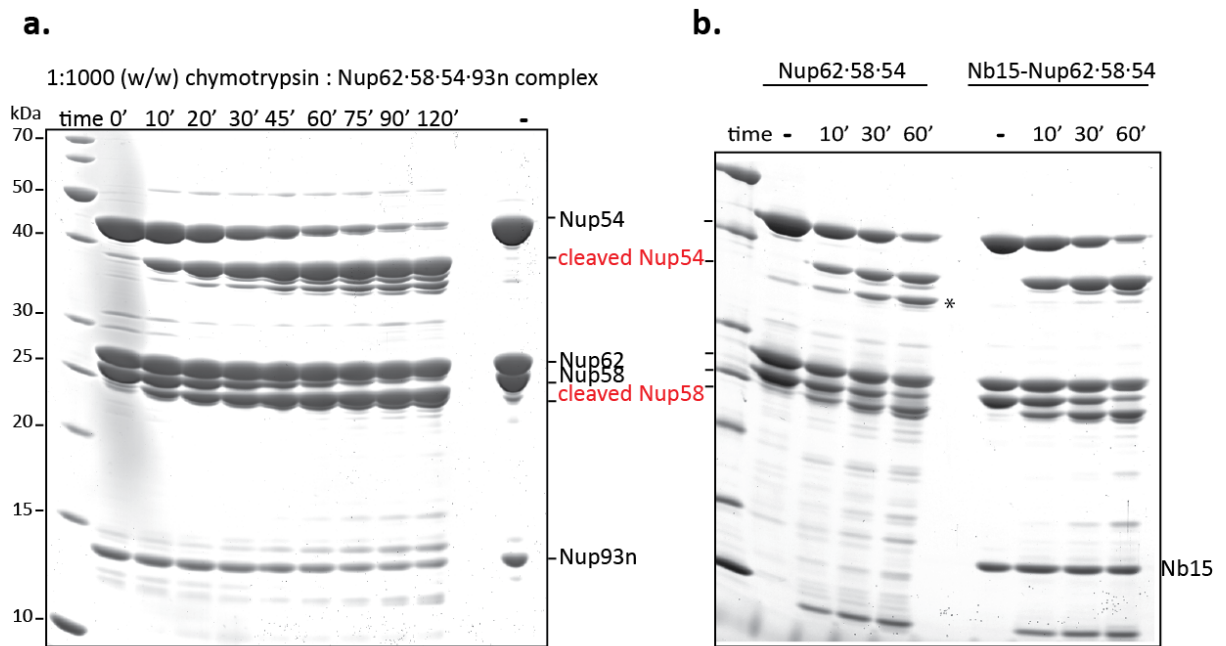


Figure 3.18: Limited proteolysis of Δ FG-xl Nup62·58·54 complex with chymotrypsin

a. Δ FG-Nup62·58·54·93n complex incubated with α -chymotrypsin in 1:1000 (w/w) protease: protein ratio. At indicated time intervals, samples were taken out and boiled in 1X SDS-sample buffer to terminate the reaction. Nup54 and Nup58 showed degradation into smaller fragments **b.** Comparison of chymotrypsin digestion pattern of xlNup62·58·54 complex with Nb15- xlNup62·58·54. Band marked with asterisk show the digestion product of Nup54 that might be shielded from protease action due to nanobody binding. **c.** Mass-spectrometry analysis of smaller fragments obtained by limited proteolysis. The bands were sliced out of the SDS-gel and analyzed by MALDI-TOF. Sequence in black and red indicate the missing and present regions respectively. Peptides for C-terminal domains of Nup54 and Nup58 were found to be mainly missing indicating that they are the major sites of protease digestion and might be exposed in the complex due to sub-stoichiometric amounts of Nup93.

3. Results

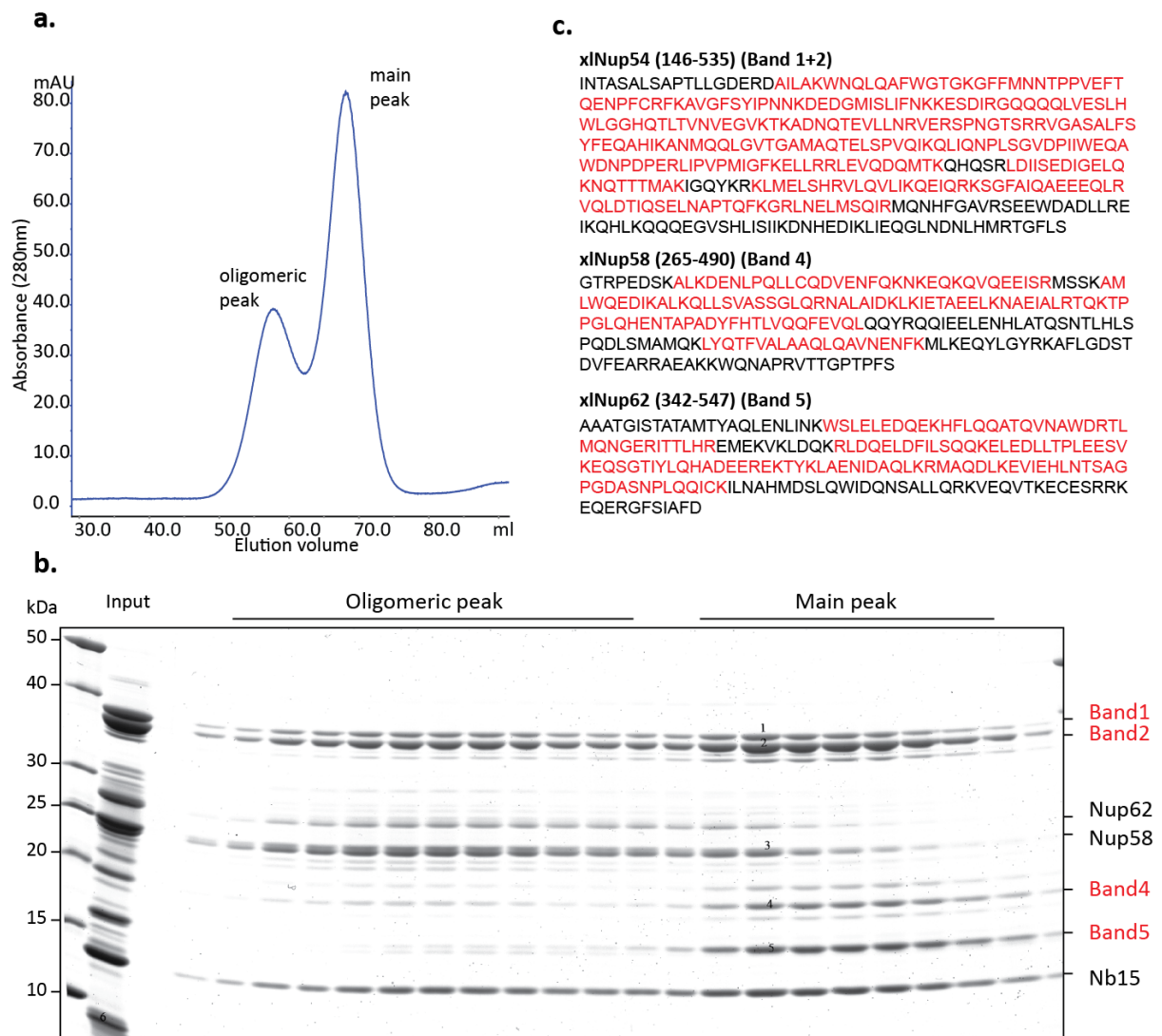


Figure 3.19: Gel-filtration of Nb15-ΔFG-Nup62-58-54 complex after limited proteolysis with chymotrypsin

a. Gel filtration profile of chymotrypsin digested Nb15-ΔFG-Nup62-58-54 complex, showing the mixed population of partially digested Nb15-Nup62-58-54 complex (oligomeric peak) and the main peak corresponding to the smaller complex which could not be separated from the oligomers by gel filtration **b.** Mass-spectrometry analysis of digested Nup54, Nup62 and Nup58 showing protease sensitive regions (black) and protease resistant (red) domains.

3.1.17. In-situ limited proteolysis

In-situ limited proteolysis has also been shown to improve crystal diffraction of many proteins where addition of trace amounts of protease to the crystallization drop was shown to remove flexible N and/or C-terminal ends resulting in better crystal packing (Dong et al., 2007). The method is especially applicable in problematic cases where obtaining the protease-digested fragments in homogeneous form is difficult. We added chymotrypsin in trace amounts to the crystallization drop of Nb15-Nup62-58-54 complex (chymotrypsin: protein concentration ranging from 1:2000- 1:10,000) and performed the extensive sparse matrix screening for Nb15-

3. Results

Nup62·58·54 complex in the presence of the protease. This time we obtained a greater number of crystal hits than the previous screen (>10 out of 3000 conditions tested).

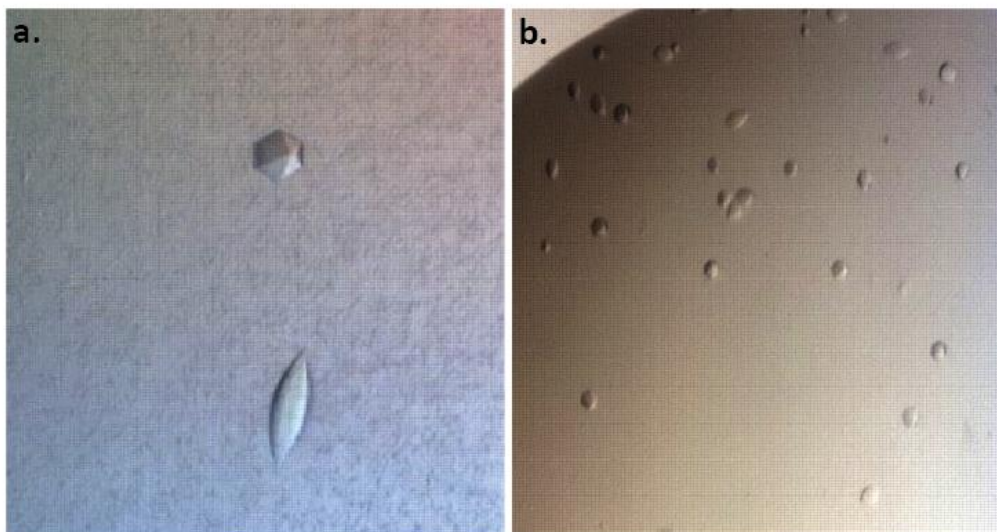


Figure 3.20 In-situ limited proteolysis of Nb15 Δ FG-xlNup62·58·54 complex

(a-b) Crystals obtained after in-situ limited proteolysis of Nb15-Nup62·58·54 complex. Trace amounts of chymotrypsin were added to the protein prior to setting crystallization drops. Multiple hits were obtained with different crystal morphology than obtained before for the non-proteolyzed complex.

The morphology of crystals was different than before, as they were single rods or small cubes/plates compared to multiples needles observed with the non-proteolyzed Nb15-Nup62·58·54 complex before (Figure 3.20). These crystals however, were very small in size and stopped growing after 30-40 μ m and hence, could not be used for crystallographic analysis.

3.1.18. Cytosolic expression of nanobodies and Nup93n-Nb15 fusion

The presence of disulfide bonds in the nanobodies poses a problem for their expression in bacterial cytoplasm, as it possesses an oxidizing environment. For this reason, periplasmic secretion signals have to be fused at the N-terminus of Nbs to transport them to the bacterial periplasm where disulfide bonds formation and protein folding occurs. This strategy though successful for obtaining large amounts of nanobodies, does not allow fusion of N-terminal bulky affinity tags or protease cleavage sites (e.g. ZZ-sumo tags). We therefore explore a bacterial strain (NEB SHuffle) for nanobody expression, which is capable of forming disulfide bonds in its cytoplasm, thus overcoming the need of fusing periplasmic secretion signals. The strain lacks major cytoplasmic reductive pathways (*trxB* and *gor* pathways) and has a chromosomal insertion of the enzyme disulfide bond isomerase (DsbC). We found that nanobodies containing the N-terminal fusion partners were extremely well expressed in this strain and did not show any degradation, indicating proper folding and correct disulfide bond formation (Figure 3.21). Thus,

3. Results

we could circumvent the need for periplasmic expression for correct folding of nanobodies by expressing them in the SHuffle strain. This was an important outcome since we could now fuse these nanobodies with specific proteins such as GFP, protease recognition folds, ZZ or MBP solubility enhancer tags and other proteins to express them in bacterial cytoplasm without any major degradation problems.

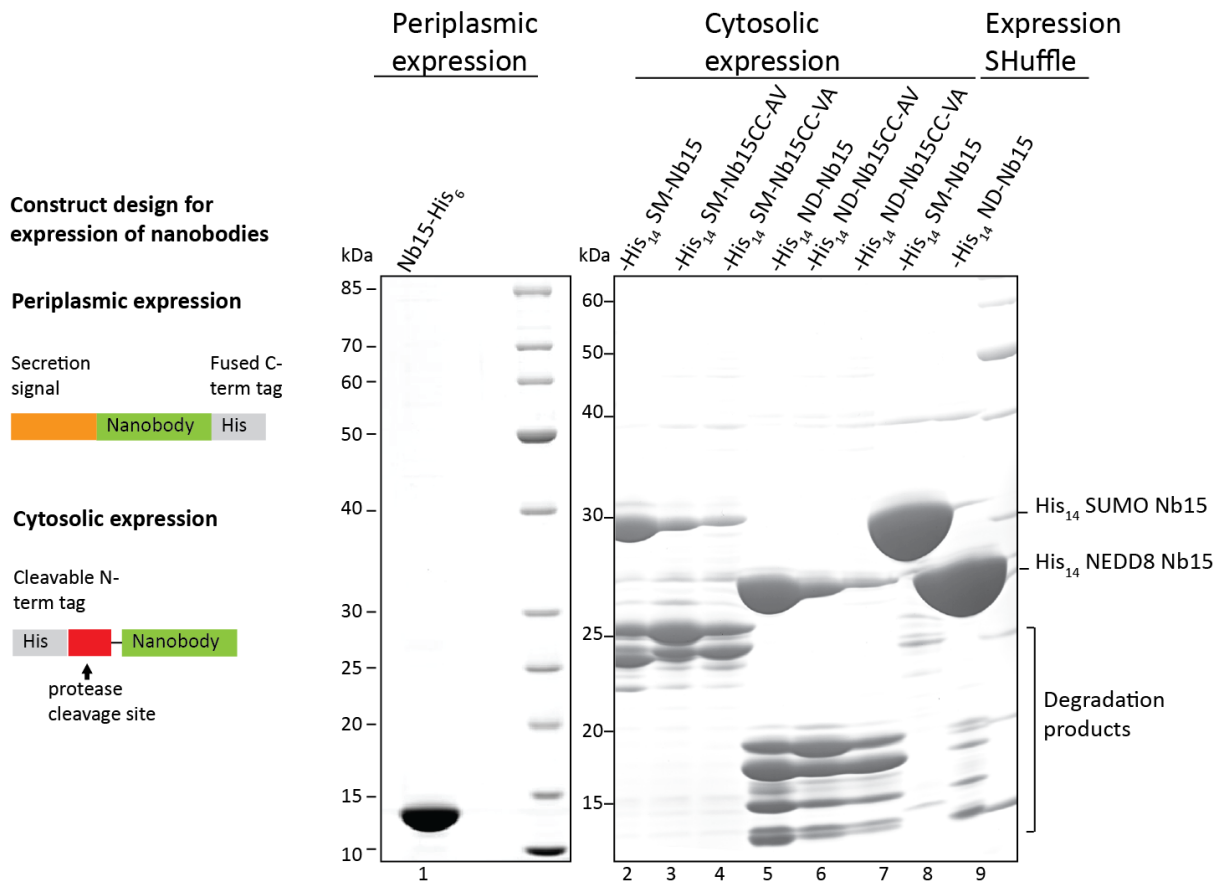


Figure 3.21: Bacterial expression of disulfide-bond containing nanobodies

Comparison of periplasmic and cytosolic expression of Nb15 (in NEB express and NEB SHuffle strains). Periplasmic expression of nanobodies relies on N-terminal secretion signal and non-cleavable C-terminal His-tag (Scheme on the left) and in general could not be combined with bulky expression enhancing tags and protease cleavage sites (left gel panel). In the cytosol of conventional bacterial expression strains (NEB express for e.g.), nanobodies undergo degradation (lane 2-7) due to lack of disulfide bond formation. We identified that NEB SHuffle strain can be successfully employed for the disulfide bond formation in nanobodies, leading to stable final product (lane 8-9) (right gel panel). “His₁₄ SM”, His₁₄-brSUMO; “His₁₄ ND”, His-NEDD8; “CC-AV”, Disulfide bond forming Cys22 and Cys96 in the nanobody framework mutated to Alanine and Valine; “CC-VA”, Disulfide bond forming Cys22 and Cys96 in the nanobody framework mutated to Valine and Alanine respectively.

3. Results

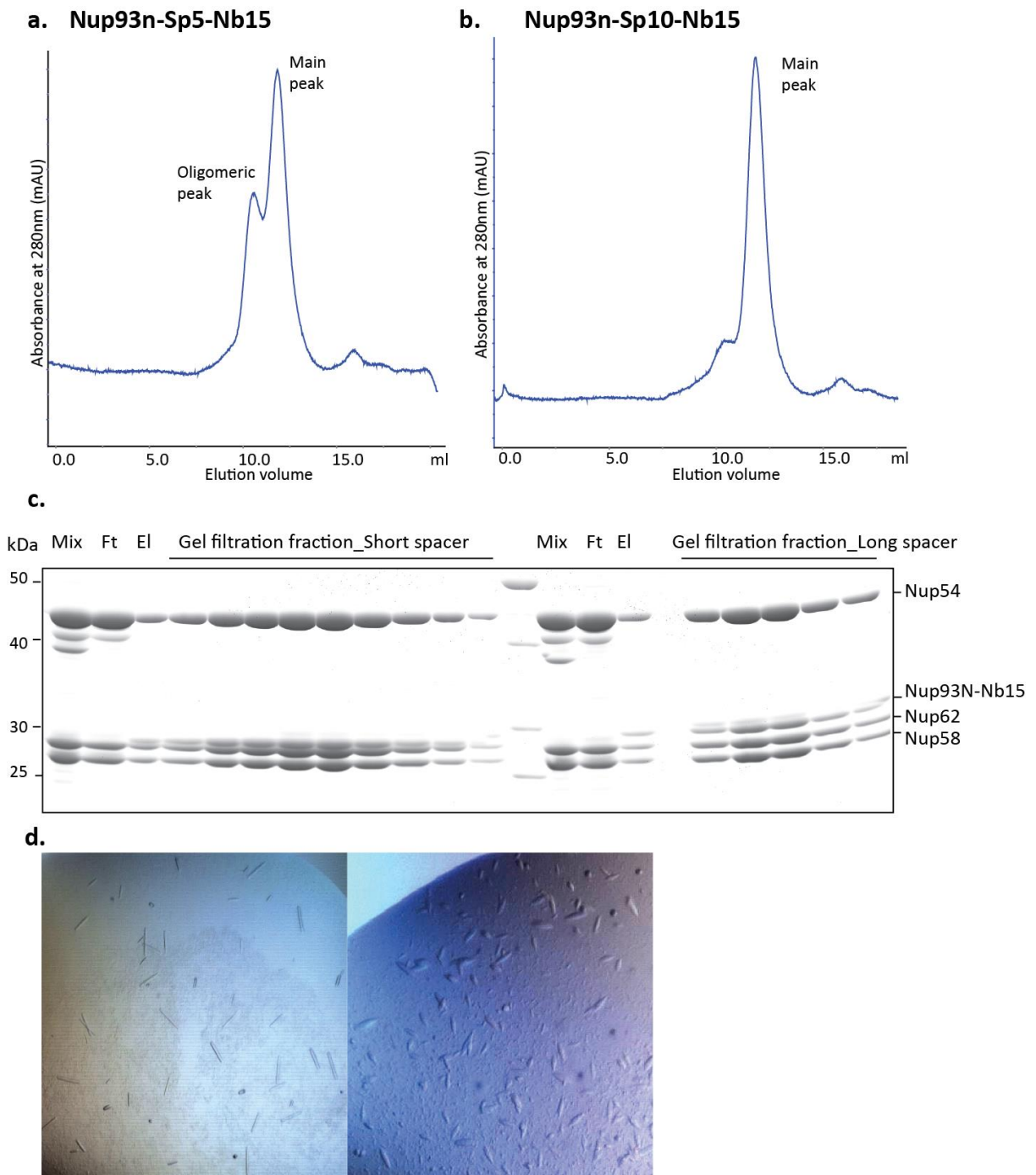


Figure 3.21: Nanobody 15 fusion to Nup93n (1-95) and optimization of spacer lengths

(a-c) Gel-filtration profiles of Nup62·58·54 complex conjugated with Nup93-Sp-Nb15. Effect of different spacer lengths between Nup93n and Nb15 on oligomerization of Nup62·58·54 complex were observed. Excess of Nup62·58·54 complex was incubated with Nup93n-sp-Nb15 (mix) to promote intra-complex incorporation of Nup93n into the Nup62·58·54 complex. With a short spacer of 5 residues (Nup93n-Sp5-Nb15), Nup93n is not incorporated into the Nup62·58·54 complex and leads to inter-complex interactions causing oligomerization (left). With a longer spacer length of 10 residues (Nup93-Sp10-Nb15), however, no oligomerization was observed indicating that Nup93n is incorporated in to the Nup62·58·54 complex. The corresponding gel filtration fractions are shown in the SDS gel below. **d.** Crystals obtained for Nup93N-Sp10-Nb15 conjugated Nup62·58·54 complex using in-situ limited proteolysis with chymotrypsin.

3. Results

Nup62·58·54 complex has interaction sites for Nup93n but owing to their weak affinity towards each other (as shown in Section 3.1.8), the interaction is not stable. Crystallization of Nup93N-terminus with Nup62·58·54 complex however is important to understand its anchorage in the NPC. In order to achieve this aim, we tried to stabilize Nup93n in the Nup62·58·54 complex by fusing it in-frame to the Nb15 with a short flexible linker in between. This fusion would prevent its spontaneous dissociation from the Nup62·58·54 complex and might enhance the chances of its crystallization.

We fused Nup93n (1-95) to the N and C-terminus of Nb15 and used this fusion to incorporate Nup93n into the Nup62·58·54 complex. We tested the effect of two different linker lengths (5 and 10 residues) between Nb15 and Nup93n. We also tested the effect of different orientations of Nup93 N-terminus with respect to Nb15 by fusing it to either at the N- or the C-terminus of Nb15. Interestingly; we found that fusion of Nup93 at the N-terminus of Nb15 (Nup93-Sp-Nb15) were better behaved than the C-terminal fusions (Nb15-Sp-Nup93) in single expression. Therefore we proceeded with complex formation between Nup93n-Sp-Nb15 (LSp-Long spacer 10 residues, SSp- short spacer 5 residues) and Nup62·58·54 complex.

Conjugation of Nup93n-Sp-Nb15 to Nup62·58·54 complex revealed that a long spacer length (10 residues) between Nup93n and Nb15 (Nup93n-Sp10-Nb15 conjugated to Nup62·58·54 complex) migrated as a single peak on the gel filtration column (*Figure 3.21b*). With a short spacer length between Nup93n and Nb15 (Nup93n-Sp5-Nb15) two peaks were observed after Nup62·58·54 complex binding (*Figure 3.21a*). One migrated at the native complex position preceded by pre-peak indicating soluble oligomers. Such oligomerization can be caused by failure of Nup93n to correctly incorporate into Nup62·58·54 complex due to smaller linker length. Free Nup93n thus, might engage in inter-complex interactions with another Nup62·58·54 complex unit causing oligomerization. The Nup93n-Sp10-Nb15 Nup62·58·54 complex containing incorporated Nup93N was then purified in large-scale for crystallization trials. Crystals were obtained for this complex after in-situ limited proteolysis and are currently limited in their size for X-ray diffraction.

Thus, we could crystallize nanobody conjugated Nup62·58·54 complex and found a way to stably incorporate Nup93 N-terminus into the complex by fusing it to the nanobodies. Obtaining better diffracting crystals for both these complexes is a part of the ongoing work. For this, we aim to combine various approaches based on limited proteolysis, construct optimization and routinely used crystallographic methods such as micro seeding. We are also exploring more nanobodies that could stabilize trimeric Nup62·58·54 complex conformation further and aid in better crystal packing. The structure of Nup62·58·54 complex will bring us very close to completing the inventory of crystallized NPC components.

3.2. Mutational analysis of the exportin CRM1

3.2.1. Strong evolutionary conservation of 20 residues in CRM1

As described in the introduction, CRM1 belongs to the Importin- β superfamily of transport receptors, which are characterized by the presence of HEAT repeats as a constituting structural element. A single HEAT repeat unit comprises of two antiparallel α -helices (of 10-20 residues each) connected together by a short linker loop. The α -helices on the convex outer surface are termed as 'A-helices' while the ones forming the concave inner surface are termed as 'B-helices'. In yeast as well as mammals, CRM1 structure is made up of 21 HEAT repeat elements stacked in a superhelical arrangement. This overall structural conservation of CRM1 is also true at the sequence level with 96% of residues being identical in human and fish (*D.rario*) CRM1.

We analyzed CRM1 sequence from 58 diverse organisms starting from lower eukaryotes such as *Trichomonas* till humans including diverse eukaryotic clades such as plants and fungi (Figure 4.1). Any bias due to overrepresentation of any particular clade was avoided by selecting only single orthologues from similar sequences (by Thomas Güttler). The analysis revealed 20 CRM1 residues to be strictly conserved throughout the evolution. No substitutions in these residues was observed anywhere in different classes of organisms. Since, CRM1 is overall a very strongly conserved NTR and is essential in both the lower as well as higher eukaryotes, conservation of these 20 residues throughout evolution suggests their importance in CRM1 function. Therefore, we separately mutated each of these 21 residues into closely related amino acids to study the effect of such subtle mutations on CRM1 function.

3.2.2. Analysis of CRM1 point mutants

Figure 4.2 shows the distribution of the absolutely conserved CRM1 residues in its structure. 10 out of 20 residues are present in the hydrophobic cleft. The remaining residues are distributed in the other parts of CRM1 structure with P427 being present in the hinge region of the acidic loop. The mutations in all these residues were carefully designed in pymol, to prevent any drastic changes in the overall CRM1 structure. All mutants were expressed in *E.coli* and purified under similar conditions. For some residues in the hydrophobic cleft, more than a single mutation was designed to identify the residues that can be tolerated. Inside the nucleus, CRM1 binds RanGTP and its cargoes in a cooperative manner.

3. Results

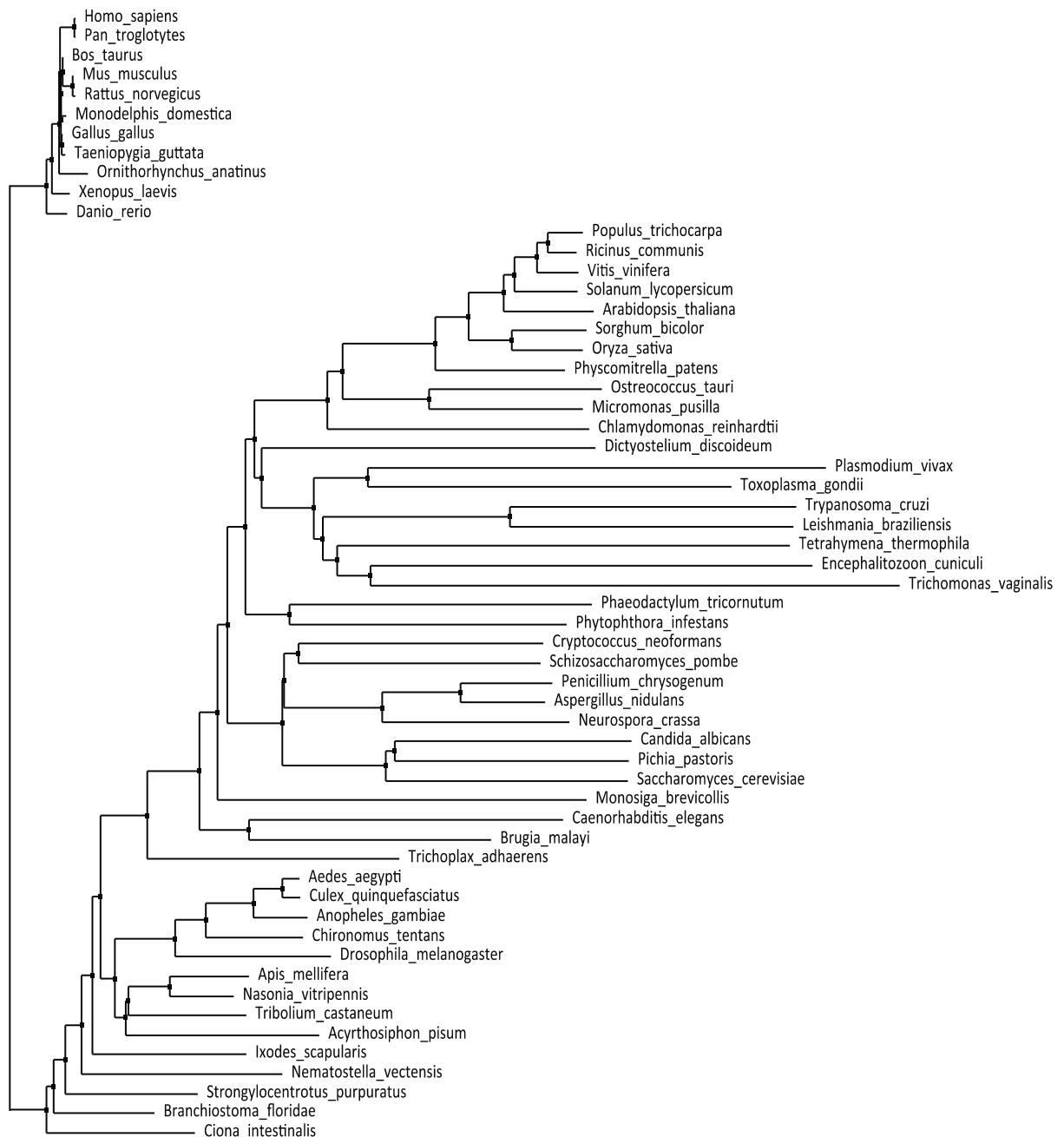


Figure 4.1: Sequence analysis of CRM1 from diverse eukaryotic species

Phylogenetic tree showing CRM1 sequence identity in different eukaryotic species. CRM1 sequences from 58 indicated species were analyzed for sequence identity. The tree was obtained by neighborhood joining method using percentage identity after ClustalW sequence alignment.

Therefore, in order to analyze the functional significance of these mutations binding assays were performed in two sets where a) Binding to PKI-NES (Protein kinase inhibitor-nuclear export signal) in the presence and absence of RanGTP Q69L (5-180) was observed and b) Binding to RanGTP Q69L (5-180) in the presence and absence of PKI-NES was analyzed. (RanGTP Q69L 5-180 is C-terminally truncated GTPase deficient form of Ran).

3. Results

PKI-NES (LALKLAGLDI), a classical leucine rich NES was one of the first export signals recognized for CRM1-mediated export. We used N-terminal His-ZZ tag fusions of PKI-NES to immobilize it on IgG sepharose. Binding assays were then performed for the wild type CRM1 and CRM1 mutants in two parallel binding reactions. One set of binding mix contained RanGTP Q69L 5-180 in the binding mix, whereas the other did not contain it. In this manner, the functional role of these mutants for RanGTP/cargo cooperativity was studied. In an alternative scheme, ZZ-tagged RanGTP Q69L 5-180 was immobilized on IgG matrix and binding was performed in the absence or presence of PKI-NES. RanGTP Q69L.

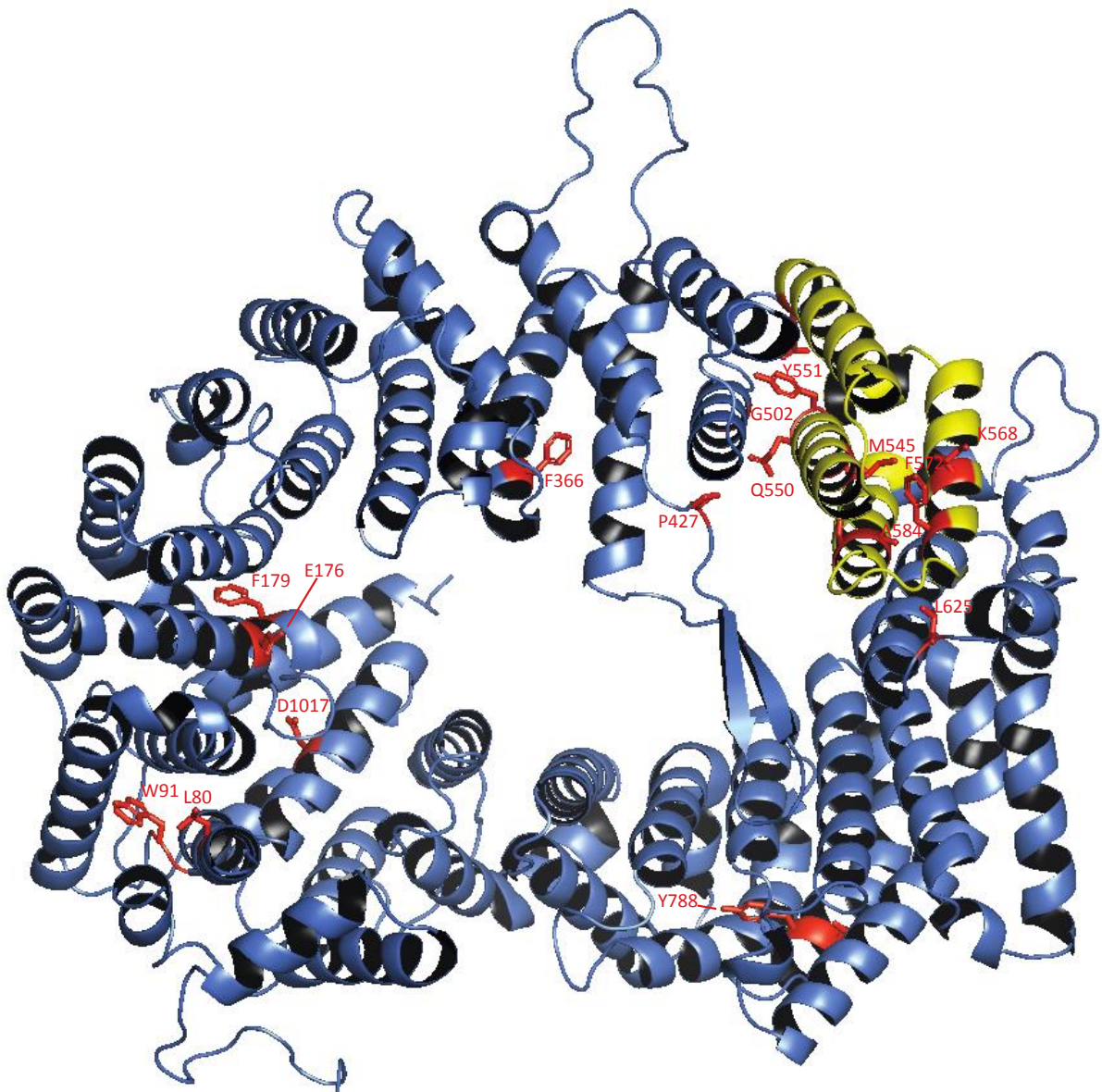


Figure 4.2: Distribution of evolutionary conserved residues on mmCRM1

Structure of mmCRM1 in the export complex (blue) showing distribution of evolutionary conserved residues (red). Hydrophobic cleft is colored in yellow. SPN1 and RanGTP have been omitted for clarity.

3. Results

3.2.3. CRM1 mutants stabilized in high affinity, cargo-binding state

Certain conserved residues in CRM1 when mutated resulted in a stronger cargo binding. These mutants can be grouped under two classes based on RanGTP and cargo binding cooperativity. The first class comprised of mutants where cargo binding was strong and the RanGTP-cargo binding cooperativity was lost (G502A/S, P427G). These mutants bound PKI-NES much stronger than wild type both in the absence or presence of RanGTP (Figure 4.3). G502 is located in the HEAT repeat 10B surrounding the HC whereas P427 is present in the hinge region of the acidic loop 9. Both these mutants probably have a permanently open hydrophobic cleft that can bind the cargo but is incapable of closing thereby preventing cargo dissociation. If this was true, such mutants with open HC should also be compatible with Ran-binding.

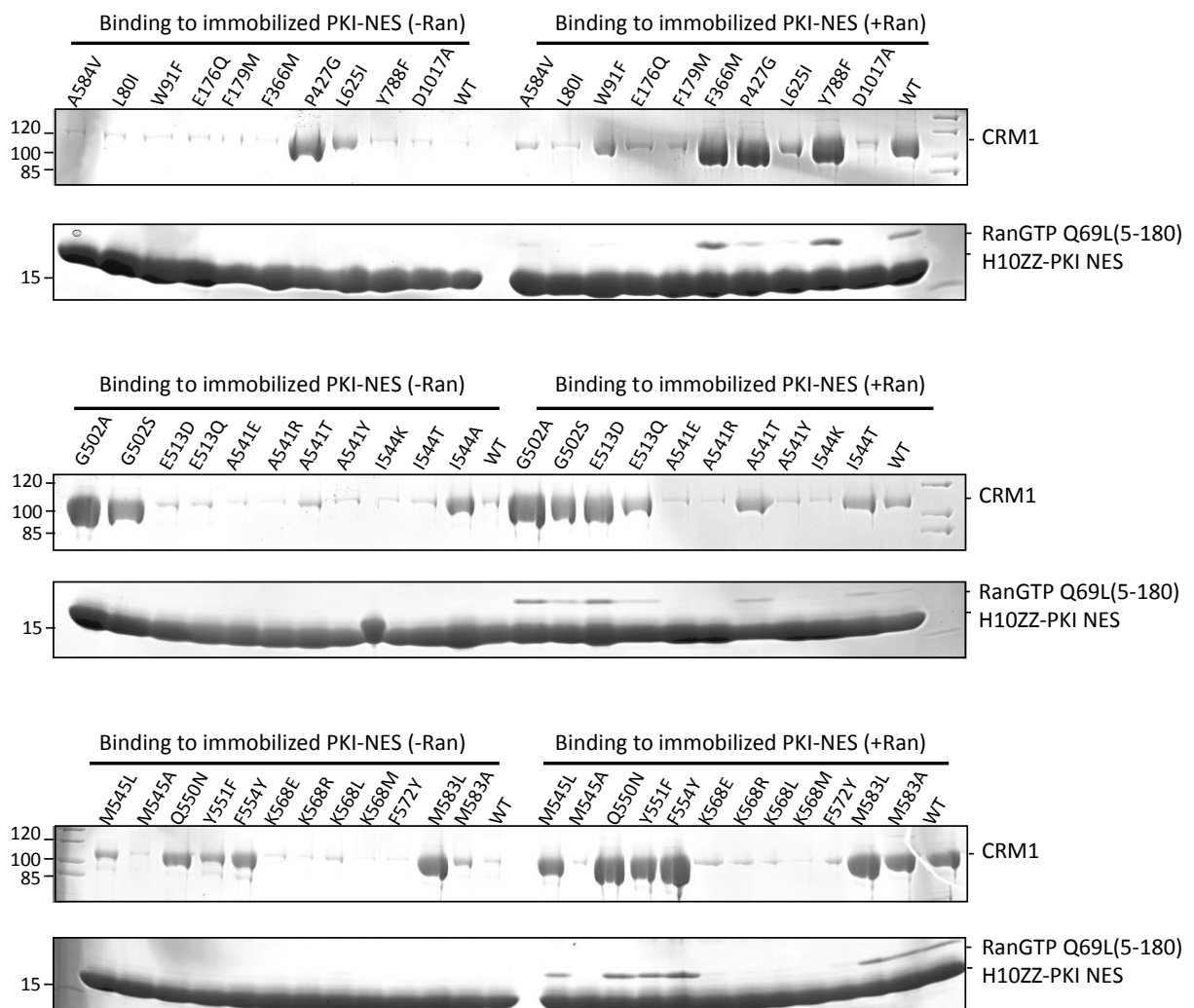


Figure 4.3: Cargo-binding of CRM1 mutants in the absence and presence of RanGTP

20 μ l IgG beads saturated with immobilized ZZ-tagged PKI-NES were incubated with 1 μ M CRM1 in binding buffer (50mM Tris-HCl pH 7.5, 300mM NaCl) in the presence or absence of 1 μ M RanGTP Q69L (5-180) and 1X E-mix. Complex formation was allowed for 2-3h at 4°C under shaking conditions. The proteins bound to IgG beads were eluted after thorough washing in 1X SDS sample buffer and analyzed on SDS-gel.

3. Results

Indeed we observed that RanGTP binding was stronger in these mutants and they bound Ran both in the absence or presence of cargo, again indicating a permanently open hydrophobic cleft or nuclear conformation (*Figure 4.4*).

The second class of mutants bound cargo stronger than the wild-type CRM1 but in a RanGTP dependent fashion (E513D, Q550N, Y551F, F554Y, M583L). These mutants bound PKI-NES much stronger than the wild-type CRM1 in the presence of RanGTP and also showed a weak binding in absence of RanGTP (*Figure 4.3*). Thus, these mutants probably stabilized CRM1 in cargo-binding conformation than its cargo free conformation. This was also observed when they bound RanGTP in the absence of PKI-NES (indicating a partially open HC) and showed increased RanGTP binding in the presence of the cargo (*Figure 4.4*). Another set of mutants that bound cargo strongly than the wtCRM1 included F366M and Y788F. These residues lie far away from the hydrophobic cleft in CRM1 structure and their effect cannot be explained based on the changes in the acidic loop and C-terminal helix. Therefore, these residues might be crucial in stabilizing the distinct conformations of CRM1.

In HeLa nuclei live export assay, the mutants that bound PKI-NES strongly and showed no RanGTP-cargo cooperativity (G502A/S, P427G) failed to exclude cargo (GFP-tagged PKI) from the nuclei (*Figure 4.4*). It is quite possible that as these mutants bind cargo permanently without dissociation, they fail to carry out multi-round export across NPCs leading to cargo accumulation in the nucleus. On the other hand, the mutants that bound PKI-NES strongly and also exhibited RanGTP-cargo cooperativity (E513D, Q550N, Y551F, F554Y, M583L, F366M and Y788F) were functional in the export assay since they excluded fluorescently labeled cargo from the nucleus at a level comparable to wild-type CRM1 (*Figure 4.5*). Whether these mutants exhibit any export defect or not can be more stringently analyzed by studying live nuclei export kinetics or affinity measurements to determine cargo association and dissociation rates.

3. Results

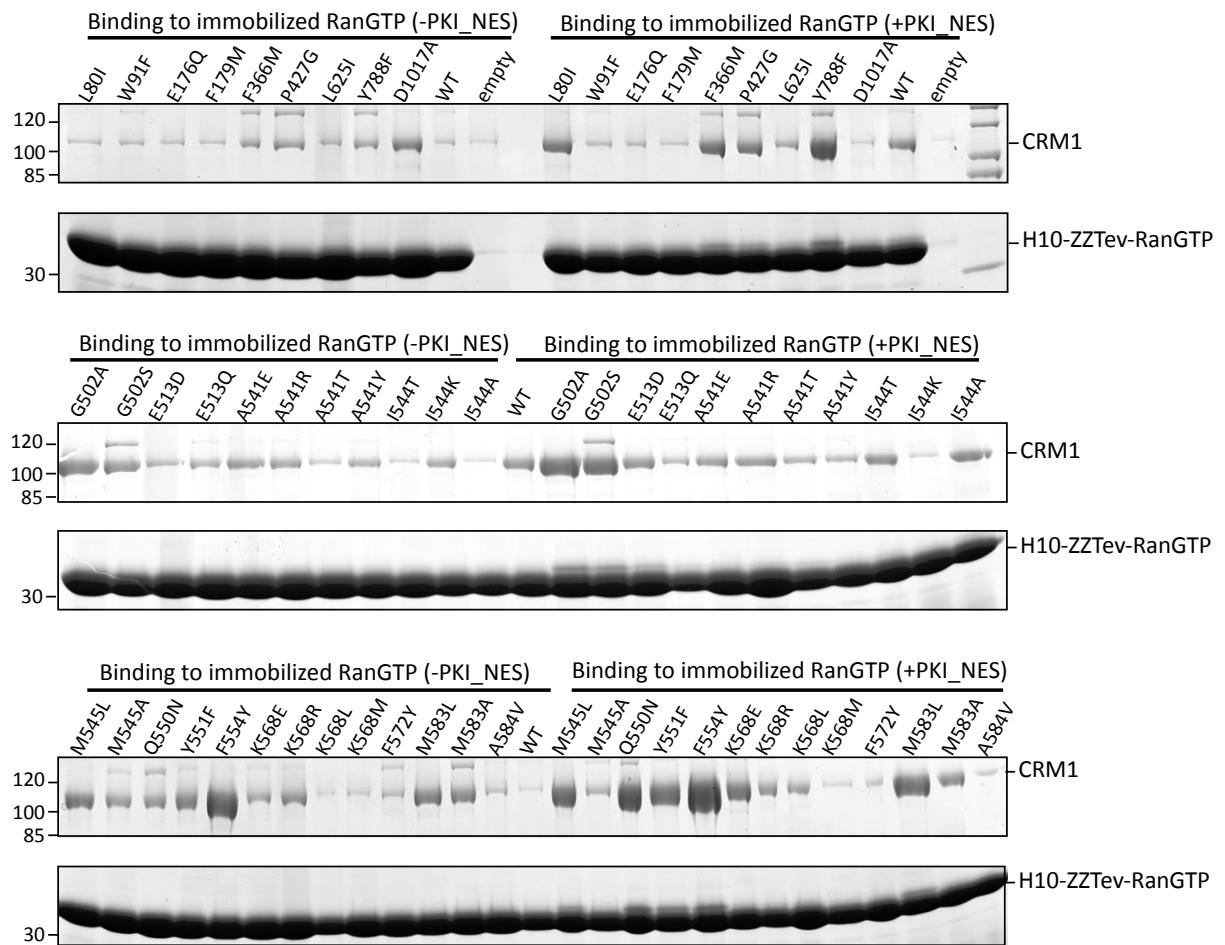


Figure 4.4: RanGTP binding of CRM1 mutants in the presence or absence of cargo

20 μ l IgG beads saturated with immobilized ZZ-tagged RanGTP (5-180) were incubated with 1 μ M CRM1 in binding buffer (50mM Tris-HCl pH 7.5, 300mM NaCl) in the presence or absence of 1 μ M PKI-NES and 1X E-mix. Complex formation was allowed for 2-3h at 4 $^{\circ}$ C under shaking conditions. The proteins bound to IgG beads were eluted after thorough washing in 1X SDS sample buffer and analyzed on SDS-gel.

3.2.4. CRM1 mutants stabilized in low affinity, cargo-free state

Interestingly, mutations in the conserved residues that lie outside the hydrophobic cleft (L80I, W91F, E176Q, F179M, L625I and D1071A) were not tolerated. These mutants failed to bind PKI-NES or RanGTP in the binding assay indicating a loss in function (Figure 4.3, 4.4). This clearly indicates the functional importance of the conserved residues in stabilizing various CRM1 conformations. However, to establish that these mutants are structurally intact and are indeed functional, we performed binding with a very strong cargo (super-PKI, or S-PKI). S-PKI has been previously shown to bind so strongly to the HC that it could force CRM1 into a nuclear conformation leading to loss of RanGTP-cargo binding cooperativity (Güttler et al., 2010). These mutants bound super-PKI in a manner similar to the wild type, indicating that indeed the structure of these mutants is intact, however, though due to subtle mutations, they were incapable of binding weak cargoes such as PKI-NES (Figure 4.5). In the HeLa nuclei export assays,

3. Results

these mutants were capable of exporting GFP-tagged PKI-NES cargo from the nucleus (Figure 4.8).

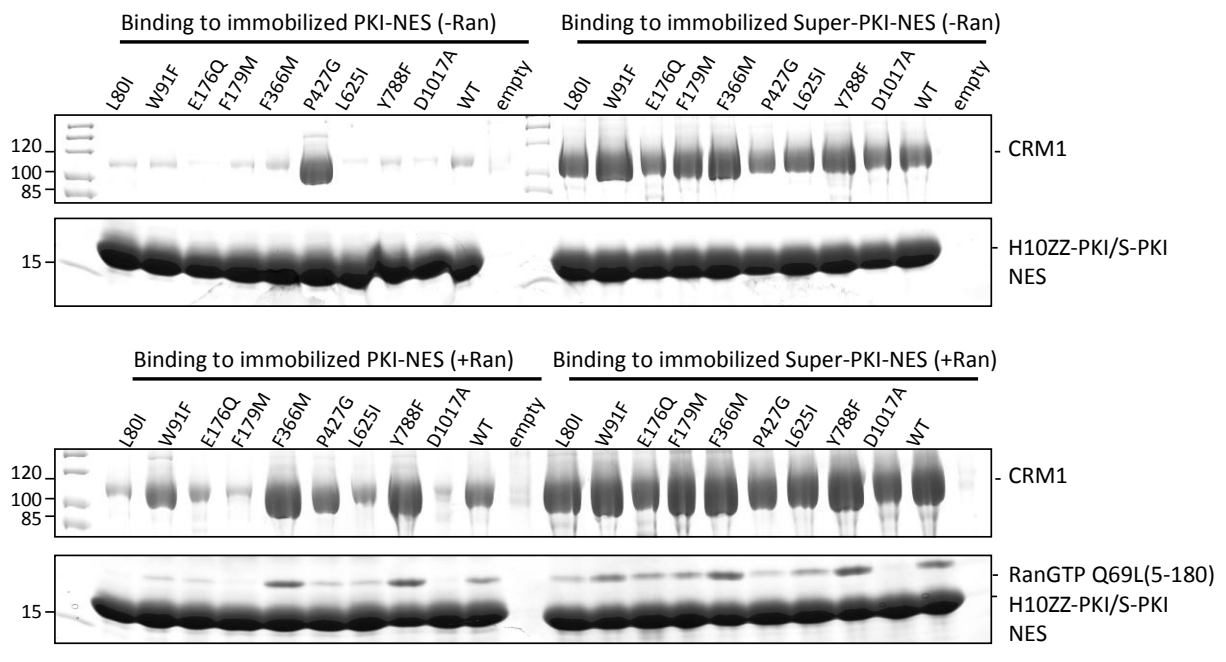


Figure 4.5: Comparison of Super-PKI and PKI binding of low affinity CRM1 mutants

CRM1 mutants with low binding affinity for PKI-NES were tested for binding to a stronger cargo (Super-PKI II NES) in the absence and presence of RanGTP. All mutants bound super PKI NES (upper and lower right panel) with similar affinity both in presence and absence of RanGTP indicating that they can be forced to adopt nuclear conformation in the presence of a strong cargo.

3.2.5. CRM1 mutants with a blocked hydrophobic cleft

Hydrophobic cleft is the most conserved site in overall CRM1 structure mainly because it is the site where NES-containing export cargo-recognition and binding occurs. Any mutation in the HC that prevents cargo binding either by steric hindrance or loss of free binding energy is bound to be deleterious to CRM1 function. Many of the conserved residues were located in the HC (A541, I544, M545, K568, F572, A584, M583) and exchanging them to other residues was not tolerated. The only exception was M583L, which is present at the base of the HC where an increase in cargo binding compared wild type CRM1 was observed. Strong mutations such as A541E/R/Y, I544K, K568E/R/L/M did not show any PKI or RanGTP binding in the binding assays as well as there was no export in HeLa nuclei. Subtle mutations at these sites such as A541T, I544A/T, M545A, A584V, F572Y, M583A, however, were export competent.

3. Results

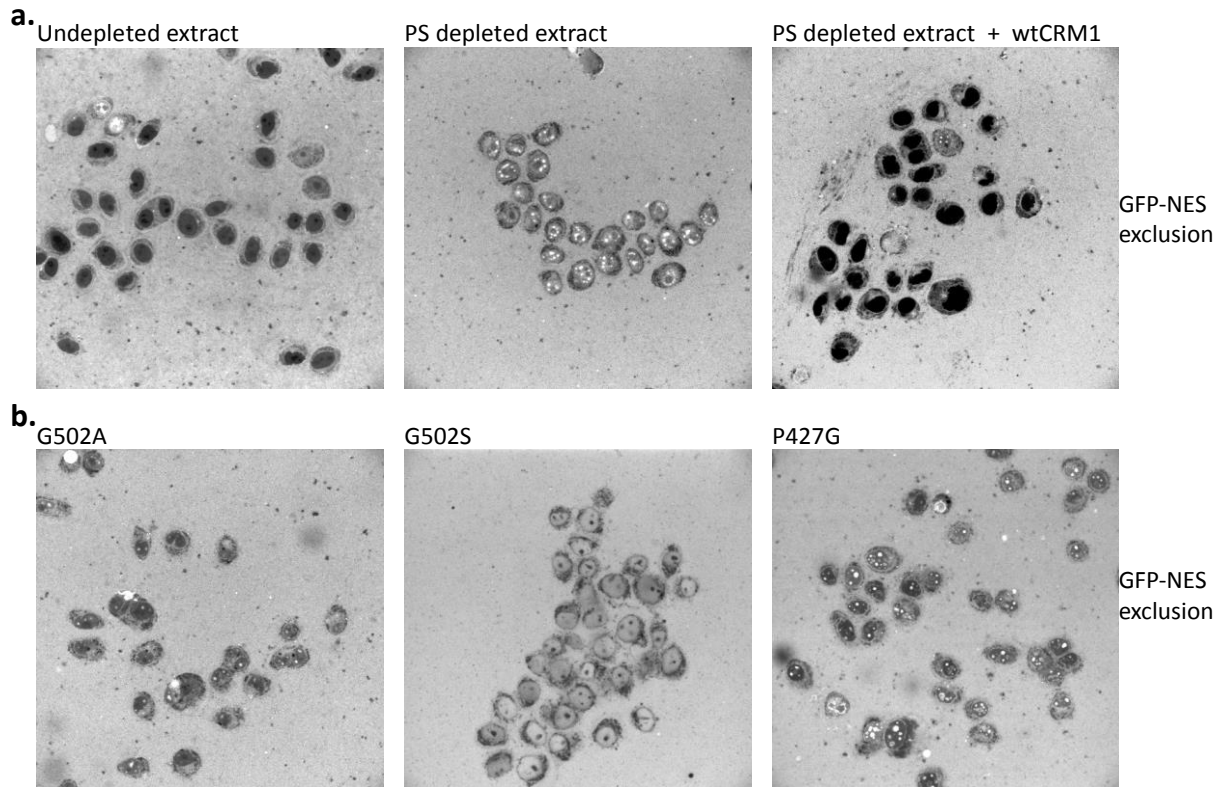


Figure 4.6: Export assay in HeLa nuclei for high affinity, RanGTP independent CRM1 mutants

a. HeLa cell nuclei were incubated with high speed *Xenopus* egg extract (undepleted or Phenyl-sepharose depleted) in the presence of Ran-regenerating system and $1\mu\text{M}$ wild-type CRM1/CRM1 mutants for 20 minutes before addition of $4\mu\text{M}$ GFP-tagged PKI-NES. The nuclei were further incubated for 10 minutes at room temperature to allow active export. Confocal sections from imaging in the GFP-channel are shown. GFP tagged PKI_NES exclusion was observed in HeLa nuclei incubated in undepleted *Xenopus* egg extract due to export by endogenous CRM1, while nuclear accumulation was observed in case of phenyl sepharose depleted extract. Addition of wild-type CRM1 led to complete exclusion of cargo from the nucleus. **b.** Export assay in HeLa nuclei incubated in Phenyl-sepharose depleted *Xenopus* egg extract with addition of indicated CRM1 mutants. Cargo accumulation in nucleus was observed for all three mutants- G502A, G502S and P427G indicating export incompetence of these mutants.

3. Results

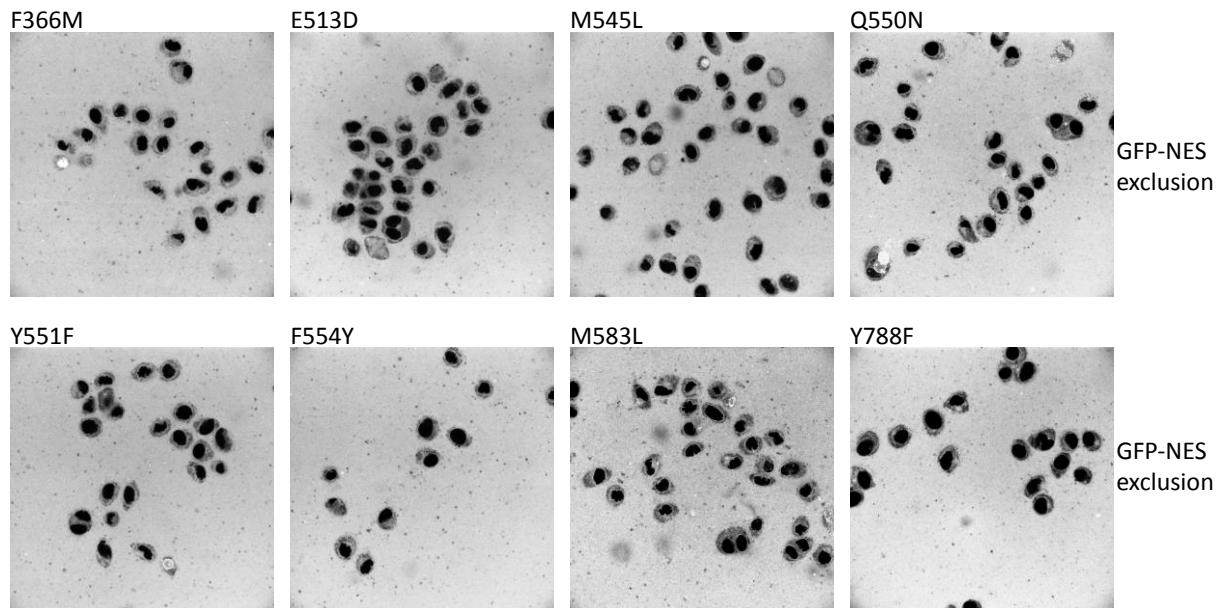


Figure 4.7: Export assay in HeLa nuclei for high affinity, RanGTP dependent CRM1 mutants

Export assay in HeLa nuclei pre incubated in Phenyl-sepharose depleted Xenopus egg extract in the presence of Ran-regenerating system and $1\mu\text{M}$ indicated CRM1 mutants for 20 minutes. $4\mu\text{M}$ GFP-tagged PKI-NES were added and further incubated for 10 minutes. All the mutants above efficiently exported cargo out of the nucleus.

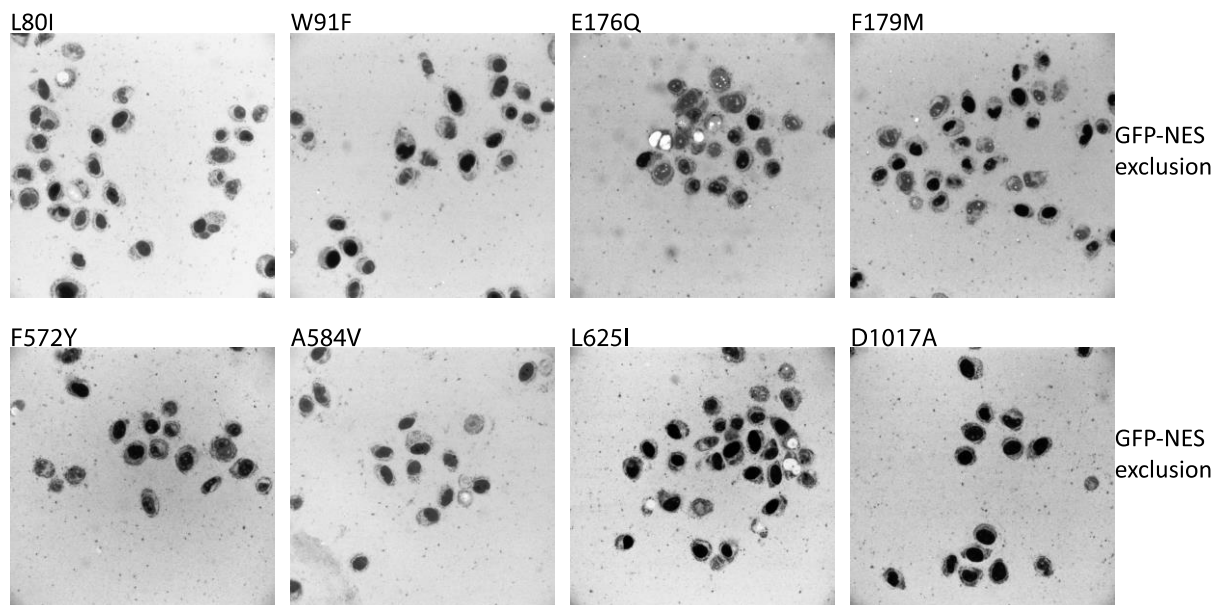


Figure 4.8: Export assay in HeLa nuclei for low affinity CRM1 mutants

Export assay in HeLa nuclei pre incubated in Phenyl-sepharose depleted Xenopus egg extract in the presence of Ran regenerating system and $1\mu\text{M}$ indicated CRM1 mutants for 20 minutes. $4\mu\text{M}$ GFP-tagged PKI-NES were added and further incubated for 10 minutes. These mutants were also found to export cargo out of the nucleus.

3. Results

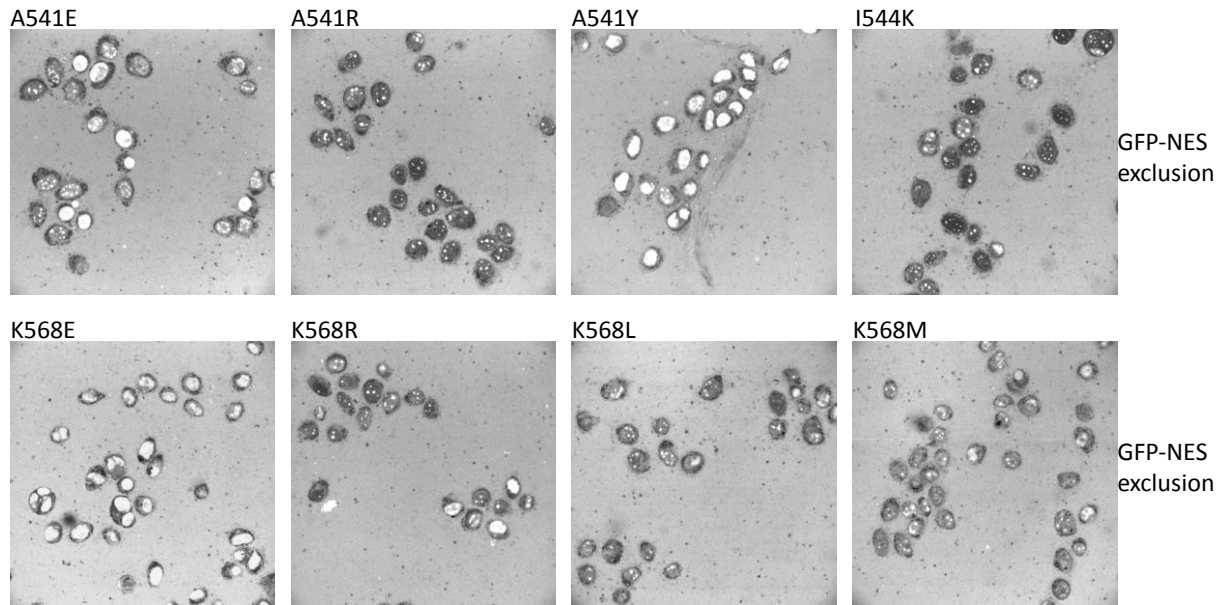


Figure 4.9: Export assay in HeLa nuclei for CRM1 mutants with blocked hydrophobic cleft

Export assay in HeLa nuclei pre incubated in Phenyl-sepharose depleted Xenopus egg extract in the presence of Ran regenerating system and 1 μ M indicated CRM1 mutants for 20 minutes. 4 μ M GFP-tagged PKI-NES were added and further incubated for 10 minutes. Cargo-accumulation in the nucleus was observed for these mutants indicating their inability to bind PKI-NES in the hydrophobic cleft.

3. Results

Mutation	Location	PKI binding compared to WT	Ran binding compared to WT	Behavior in export assay	
P427G	Acidic loop	****	*	Dead	CRM1 mutants with strong cargo binding
G502A	Surrounding HC	****	***	Dead	
G502S	Surrounding HC	****	***	Dead	
E513D	Surrounding HC	****	*	Active	
E513Q	Surrounding HC	**	*	Active	
Q550N	Surrounding HC	***	***	Active	
Y551F	Surrounding HC	***	**	Active	
F554Y	Surrounding HC	***	****	Active	
M583L	Base of HC	***	*	Active	
F366M	Ran-binding domain	***	***	Active	
Y788F	Between Ran and cargo binding domain	***	***	Active	
L80I	Ran-binding domain	-	-	Active	CRM1 mutants with weak cargo binding
W91F	Ran-binding domain	<	-	Active	
E176Q	Ran-binding domain	-	-	Active	
F179M	Ran-binding domain	-	-	Active	
L625I	Between Ran and cargo binding domain	<	<	Active	
D1017A	Between Ran and cargo binding domain	-	-	Active	
A541E	Base of HC	-	*	Dead	CRM1 mutants with blocked hydrophobic cleft
A541T	Base of HC	<	*	Active	
A541R	Base of HC	-	*	Dead	
A541Y	Base of HC	-	*	Dead	
I544K	Base of HC	-	-	Dead	
I544T	Base of HC	<	*	Active	
I544A	Base of HC	<	*	Active	
M545A	Base of HC	-	-	Active	
K568E	Wall of HC	-	*	Dead	
K568R	Wall of HC	-	*	Dead	
K568L	Wall of HC	-	*	Dead	
K568M	Wall of HC	-	-	Dead	
F572Y	Wall of HC	-	-	Active	
A584V	Base of HC	-	-	Active	

Table 3 Table summarising the properties of CRM1 mutants

“*”, Binding compared to the wild-type CRM1; one star indicates similar level of binding as the wild type, two, three and four stars indicate higher levels of binding than the wild-type with four stars indicating the highest level. “-”, No binding; “<”, weaker binding than the wild-type CRM1.

4. Discussion

Eukaryotic nuclear pore complex is a fascinating nanostructure that has been functionally and structurally conserved throughout the evolution and acts as the sole mediator of macromolecular exchange across the nuclear envelope. The giant machinery with a MW ~120MDa in humans and ~60MDa in yeast is made up from 30 different subunits (nucleoporins), which assembles together into a highly symmetrical supramolecular structure, with octahedral symmetry in the plane of the nuclear envelope. In the last decade, molecular structures of a large number of nucleoporins and nuclear transport receptors have been described. Most of the nucleoporins exhibit a combination of super α -helical solenoid folds and β -propeller domains, which deviate strongly from other protein structures in the PDB. Not just the architectural elements of the NPC are structurally conserved; the key players involved in nucleo-cytoplasmic transport process also exhibit a distinct morphology that has been conserved throughout the evolution in different organisms. A large number of nuclear transport receptors belonging to the Importin- β superfamily (such as Importin- β and CRM1) have an alpha-helical solenoidal domain organization. All these NTRs shift between a cargo loading and cargo release state in response to RanGTP binding and traverse through the NPC by interacting with the FG-repeat domain containing Nups. With the high-resolution structures emerging for both the NPC structural elements (i.e. scaffold forming nucleoporins) and the functional players involved in nucleo-cytoplasmic transport (i.e. soluble NTRs), it is becoming evidently clearer that the two might share a common evolutionary origin. An inherent structural flexibility is also emerging to be a common feature of both, with it being more pronounced in the soluble NTRs that undergo rigid body conformational changes during their role as macromolecule shuttles across the NPC. Many NTRs including CRM1 have been crystallized in complex with RanGTP and/or cargo as the binding of the interaction partners imparts the structural rigidity to the overall exportin structure, favoring crystallization. Nucleoporins on the other hand exist as a part of large stable subcomplexes and as a result have been crystallized with their interaction partners in order to achieve homogeneity and structural rigidity required for crystallization. Nup62·58·54 nucleoporin complex forms part of the NPC central channel that contributes considerable mass of FG-repeats required for the formation of NPC's permeability barrier. The exportin CRM1 on the other hand, overcomes this transport barrier by interacting with the FG-repeats and adopts different conformations in cargo bound or cargo free form during the translocation events. The crystal structure of CRM1 in different functional complexes is known whereas that of Nup62·58·54 nucleoporin complex is not available so far.

4. Discussion

4.1.1. Insights into the molecular behavior of the Nup62·58·54 complex

The Nup62·58·54 nucleoporin complex contributes a considerable share of the NPCs FG repeat mass and it is undisputed that it is crucial for the transport function of NPCs in a variety of organisms including *S.cerevisiae* and *Xenopus laevis*. Considering this, the existing literature is, however, surprisingly controversial in describing this complex. The controversies concern the number of Nup62·58·54 complexes per NPC, the stoichiometric ratio between the three subunits (Nup62, 58, 54 in metazoans; Nsp1p, Nup57, Nup49 in *S.cerevisiae*), the question as to which of the three subunits make direct contacts with each other, the question as to whether the Nup62·58·54 complex oligomerizes and if such oligomerization is physiologically relevant or an experimental artifact. The last question is also related to a mechanistic debate, namely if the Nup62 complex mediates facilitated translocation by dilating and contracting the NPC's central channel in response to shifts in homooligomeric contact sites, as suggested by Melcak and coworkers (Solmaz et al., 2013; Solmaz et al., 2011; Melcak et al., 2007) or if it supports the permeability barrier by contributing barrier-forming FG domains. In this project, we could answer many of these questions in an unambiguous manner especially those concerning Nup62·58·54 complex stoichiometry and its oligomerization behavior. Our experimental data directly contradicts the Nup62·58·54 complex behavior in the NPC argued by (Solmaz et al., 2011), which is solely based on the alternative interactions observed for the fragments of Nup62, Nup58 and Nup54 in the crystal structures.

4.1.1.1. A clear 1:1:1 stoichiometry of the *xl*Nup62·58·54 complex

In various studies, different modes of isolation of Nup62·58·54 complex from different nuclear envelope preparations have resulted in contradictory observed stoichiometries of Nup62·58·54 complex. The picture is further complicated by the fact that in these preparations a fourth protein at 45 kDa co-purifies with the rest of the complex. In earlier studies it was reported to be the proteolytic product of Nup58 while later on suggested to be an alternatively spliced isoform of Nup58. The isoform is exclusive to rats and does not exist in other organisms.

To highlight the observed discrepancies in vertebrate Nup62·58·54 complex stoichiometry, a ~550kDa Nup62·58·54 complex in 4:1:4 ratio (Finlay et al., 1991) and a ~234kDa Nup62·58·54·45 complex with 1:1:1:1 subunit ratio (Guan et al., 1995) were isolated after detergent treatment whereas a 1:1:2 subunit ratio of ~231kDa (Kita et al., 1993) and ~400-500kDa (Buss and Stewart, 1995) Nup62·58·54 complex was isolated on high salt treatment of the nuclear envelopes.

The picture is more clearer in yeast where genetic tagging and biochemical fractionation has revealed 1:1:1 stoichiometry of the Nsp1·Nup49·57 complex (Grandi et al., 1995b) (Bailer et al., 2001). Considering the overall architectural conservation of NPCs in yeast and higher eukaryotes,

4. Discussion

it is difficult to imagine the reasons behind such divergence in the nucleoporin ratios in different organisms and is most likely a consequence of different technical procedures.

In our experiments with recombinant co-expression and isolation, a 1:1:1 stoichiometry of Nup62·58·54 complex from *X.laevis* was found to be highly reproducible. The fact that the complex was assembled inside *E.coli* in the presence of thousands of other *E.coli* proteins and an identical subunit ratio was pulled out (in the SDS gel pattern) each time indicates that this indeed is the real stoichiometry for the Nup62·58·54 complex. The ratio is consistent with the proteomic studies (Cronshaw et al., 2002; Ori et al., 2013) where a 2:1:1 stoichiometry of Nup62, Nup58 and Nup54 has been reported. Considering Nup62 is part of another nucleoporins complex (Nup62·88·214) as well, a 1:1:1 stoichiometry of Nup62·58·54 complex fits the scenario. Multi-angle light scattering experiments confirmed this stoichiometry where the observed molecular mass of the complex perfectly matched the theoretical mass when single copy of each subunit is present.

4.1.1.2. *Oligomerization of Nup62·58·54 complex- real or artifacts?*

A 'dynamic' Nup62·58·54 complex assembled by Melcak *et al* (Melcak et al., 2007) showed protein concentration dependent oligomeric assemblies in solution. These assemblies were postulated to be physiologically relevant in context of the nuclear pores. It was argued that macromolecular transport driven large-scale conformational changes could be induced in the NPCs due to dynamic association and dissociation of the Nup62·58·54 complex. Even when one was to assume this was true, the copy number of the subunits in their model with 128 copies of Nup62, 32 copies of Nup58 and 64 copies of Nup54 (4:1:2 stoichiometry) is extremely large and cannot be reconciled with any other study nor with the arrangement of other nucleoporin complexes in the NPC. The main reason of the dynamic behavior of Nup62·58·54 complex observed by Melcak *et al* could be the partial nature of the complex studied by them. The oligomer forming Nup62·58·54 complex studied by them did not contain Nup54 α/β domain and the long coiled-coil domain of Nup58 (Nup58 CC1). Understandably so, this dynamic complex failed to crystallize.

The xINup62·58·54 complex assembled by us is significantly larger than the previously studied complex. The fact that it contains all the structural elements of Nup62·58·54 complex including Nup54 (Nup54 α/β domain) and the first coiled-coil domain of Nup58 (Nup58 CC1) makes this complex as close as it can get to the native FG-repeat containing Nup62·58·54 complex. In fact, all these structural domains were found to be absolutely essential during in-vitro Nup62·58·54 complex assembly and its incorporation into the *Xenopus* NPCs (Bastian Hülsmann, unpublished

4. Discussion

data). Small deletions of 5-10 residues from the N-terminus of any subunit in the assembled xINup62·58·54 complex resulted in loss of complex formation.

We did not observe any oligomeric tendencies in xINup62·58·54 complex. In fact, the complex could be diluted significantly and concentrated to upto 20mg/ml for crystallization without observing any dynamic oligomers (as judged by its gel-migration behavior and polydispersity analysis). The complex was found to be monomeric in MALS analysis.

Thus, we propose that the interactions observed by Melcak *et al* are most likely physiologically irrelevant or are an experimental artifact owing to the severely 'chopped off' nature of the studied complex. The behavior of Nup62·58·54 complex has great implications on our understanding of NPC architecture especially when a ring contraction and dilation model has been proposed for NPC channel based on these interactions.

4.1.1.3. Interaction amongst the subunits of Nup62·58·54 complex

While a direct interaction between Nup62·54 and corresponding scNsp1·Nup57 is well established (Bailer *et al.*, 2001) (Buss and Stewart, 1995) no direct experimental evidence exists for Nup62 interaction with Nup58. The current model of vertebrate Nup62 complex arrangement (Buss and Stewart, 1995) (Solmaz *et al.*, 2011) advocates Nup54 as platform providing separate binding sites for Nup62 and Nup58 interaction. In our co-expression system, we could address this question and to our surprise discovered that a direct interaction between Nup58 and Nup62 indeed exists. The two interacted strongly and pulled each other out from *E.coli* lysate during co-expression. However, the binary Nup62·58 complex was aggregation-prone in the absence of Nup54. The observation is further strengthened by the fact that deleting the proposed interaction region on Nup54 for Nup58 (Solmaz *et al.*, 2011), a trimeric Nup62·58·54 complex still assembled (data not shown), possibly due to a direct interaction between Nup62 and Nup58. Furthermore, we also obtained nanobodies that recognized a trimeric xINup62·58·54 complex exclusively and failed to bind the binary Nup62·54 or Nup62·Nup58 complexes. Since, nanobodies are known to bind specific conformations in small regions of their target proteins, this suggests a direct Nup62·58·4 interaction or at least a conformational proximity of Nup62 and Nup58 which can not be explained by separate binding sites of Nup62 and Nup58 on Nup54.

4.1.2. Nup62·58·54 complex anchorage to the NPC

Most physical features of NPCs are conserved in yeast and higher eukaryotes (except smaller size of yeast NPCs ~66kDa compared to human NPCs ~120kDa). However, a basic cell-cycle difference exists between the two in terms of their behavior during mitosis. *S.cerevisiae* undergoes closed mitosis where both the nuclear envelopes as well as the NPCs stay intact

4. Discussion

throughout the mitotic process and no disassembly of NPCs into individual components occurs. In higher eukaryotes such as *Xenopus* and humans, nuclear envelope breakdown is an essential process required for chromosomal segregation and includes NPC's disassembly into soluble nucleoporin subcomplexes. In higher eukaryotes, Nup62·58·54 complex exists as a separate entity in the mitotic cytosol physically disassociated from Nup93 (Macaulay et al., 1995), which interacts strongly with another nucleoporins and exists as Nup93·205 subcomplex. The yeast Nic96-Nsp1 subcomplex in contrast, never disassembles. In *S.cerevisiae*, it is well established that N-terminus of Nup93 containing coiled-coil domains is responsible for Nsp1 complex integration into the NPC scaffold. Deletion of 200 residues at the N-terminus of Nup93, which harbors coiled-coil domain forming heptad-repeats, results in loss of interaction between scNic96 and scNsp1 complex and associated growth defects in the strains carrying Nup93N-terminal deletion. However, surprisingly limited amount of data exists for vertebrate Nup62·58·54 complex anchorage to the NPC. Biochemical extraction of Nup93 and Nup62 from HeLa extracts showed that only a fraction of the two proteins associate together (Grandi et al., 1997). Whether this could be attributed to a loss of protein interaction during isolation procedure owing to weak interactions or the actual stoichiometry of the Nup93-Nup62 complex in vertebrates has never been studied. Furthermore, whether Nup93 coiled-coil domains are the only interaction sites for Nup62 complex in the NPC, as described for yeast, is not known.

We looked into the interaction between N-terminal coiled-coil containing region of recombinant xINup93 and xINup62·58·54 complex in-vitro and discovered that an inherently weak interaction between the two exists. In our experiments, we could specifically pull out Nup62 complex using affinity tag on Nup93 N-terminus and *vice-versa*. The fact that this interaction occurs in *E.coli* cytosol suggests it to be highly specific. However, spontaneous dissociation of Nup93 from the rest of the complex during successive purification steps indicated a weak interaction, different from the coiled-coil interaction within Nup62 complex that was found to be stable in all tested conditions. Our experiments suggested the K_d values of affinity of Nup93 N-terminus for Nup62 complex to be in the order of $2\mu\text{M}$. Such low affinity for interaction might be relevant in the context of NPC where the concentration of the constituents is in the range of $200\mu\text{M}$. Alternatively, additional anchorage sites for Nup62·58·54 complex in the NPC might be present that have not been discovered yet.

Based on domain topology, xINup62·58·54 complex differs from scNsp1·Nup49·57 complex in that it contains an additional region of ~200 residues at the N-terminus of Nup54 coiled-coil domains that is predicted to fold into α helices and β -sheets (Nup54 α/β domain). This domain, as mentioned above, was missing in the complex assembled by (Melcak et al., 2007) that showed

4. Discussion

high oligomeric tendencies. Also, we obtained many nanobodies directed against this domain, indicating that it might be more exposed or hydrophilic than the rest of the coiled-coil domains. We propose that this domain might be an additional anchorage site for the vertebrate Nup62 complex in the NPC.

The dissociation of nucleoporins subcomplexes from each other during mitosis, in general, is poorly understood at a molecular level. Although phosphorylation of Nups is known to be involved in this process, presence of multiple phosphorylation sites on each Nup complicate the study for determining the target sites for phosphorylation and resulting dissociation. The process can only be studied in higher eukaryotes and not in yeast. Such a direct weak interaction for vertebrate nucleoporins belonging to two separate subcomplexes has not been described before. One possibility to easily disassemble NPC is to evolve multiple sites of weak interaction between the nucleoporin subcomplexes that can be easily disrupted during mitosis resulting in formation of stable entities that reside in cytosol until NPC reassembles.

4.1.3. Nanobodies as tools for co-crystallization of Nup62·58·54 complex

With the crystal structures available for many nucleoporins, internal flexible loops or hinge regions in the Nups, appear to be a recurring theme. These have been observed in the crystal structures of subunits of scNup84 subcomplex, large inner ring Nups (scNup192, scNup188, scNic96, scNup170) and even smaller Nups (scNup53). Flexibility in the domains hampers crystallization, and as a result, many of these structures (scNup192, scNup188, scNup170, subunits of scNup84 complex) have been solved partially by crystallizing the stable domains only. We explored the potential of small, high affinity binders called nanobodies that have a rigid scaffold, for stabilizing potentially flexible Nup62·58·54 complex and aiding in its crystallization. Nanobodies have the potential for co-crystallizing entire nucleoporins because of their rigid scaffold where binding in the flexible domains of Nups could reduce their flexibility. But this strategy has not been tried for Nups so far although it has proven to be successful for co-crystallization of some of flexible parasite domains.

Other advantage of using nanobodies is that the conformation specific nanobodies can be easily obtained by customizing the panning strategy during phage-display. Since, we screened for the entire Nup62 complex binders during phage display and not for its individual subunits, the nanobodies obtained were reflective of the conformation its subunits adopt in the context of the complex. Also, different nanobodies recognize different binding sites on the Nup62·58·54 complex and were screened for in-vitro binding to the Nup62·58·54 complex to find the most stabilizing ones. 18 out of 24 nanobodies were directed against Nup54 α/β domain, indicating that this domain is not shielded in the Nup62 complex. Out of the remaining nanobodies, four

4. Discussion

nanobodies bound exclusively to trimeric Nup62 complex and did not recognize its individual subunits or any combination of binary complexes. This indicated that the nanobodies bind a site in the Nup62·58·54 complex where all three subunits interact closely with each other. Thus, these nanobodies stabilize the conformation of trimeric Nup62·58·54 complex and also probably rigidify the flexible nature of the complex.

4.1.3.1. Nanobody mediated crystallization of Nup62·58·54 complex

Our attempts to crystallize Nup62·58·54 complex did not succeed after exhaustive trials. However, crystallization trials of Nup62·58·54 complex conjugated to Nb15 led to immediate crystallization of this complex. The crystals appeared after 2 days and reached their maximum size in 1 week. Thus, conjugation of Nup62·58·54 complex with Nb15 clearly helped in its crystallization. Moreover, this positive effect on crystallization of Nup62·58·54 complex was observed only for Nb15 (that stabilizes the trimeric Nup62·58·54 complex state) and not with other nanobodies (e.g. the ones that bind the Nup54 α/β domain or Nup62·54 binary complex). This indicates that it is indeed the flexibility in the coiled-coil regions of Nup62·58·54 complex that hampered its crystallization. Nb15 binding to a flexible region in the coiled-coil domains might rigidify it allowing crystal packing. Rod-shaped crystals obtained for this complex are currently limited in their diffraction potential to 7.5Å. However, this is the first successful crystallization of the entire structured region of Nup62·58·54 complex so far. Attempts to improve the crystal diffraction potential are ongoing.

In a similar manner, nanobody conjugation approach can also be applied to other NPC subcomplexes, whose crystal structures are not available yet. One such example is Nup62·88·214 complex, which also contains coiled-coil domains as their constitutive elements and probably have similar flexible domains. Other examples are large Nups such as scNup192, which have been crystallized partially and the structures of their rigid domains are available whereas information for the hinge regions connecting these domains is completely missing.

4.1.3.2. Exploiting nanobodies for stabilizing Nup93 interaction with Nup62·58·54 complex

Inherent weak affinity of Nup93N-terminus coiled-coil domains for the Nup62 complex (as described in Section 3.1.8) posed a problem for the structural determination of Nup62·58·54·93n complex. Spontaneous dissociation of Nup93n during purification and in the crystallization drops led to sample heterogeneity and failure to understand this complex at a molecular resolution. In the absence of any published biochemical data on the subunits of Nup62·58·54 complex interacting with Nup93n, we adopted an unbiased approach for integrating Nup93n into the Nup62·58·54 complex structure by fusing it to Nb15. However, to obtain this, we first had to establish a successful expression system for nanobodies in the

4. Discussion

bacterial cytoplasm. Traditional approach for nanobody expression in *E.coli* has relied on their periplasmic expression. The reason for this is the presence of disulfide bonds essential for maintaining the core structural framework of the nanobody scaffold. The periplasmic expression however, does not allow too much flexibility in the construct designs. We achieved a high yield of correctly folded nanobodies by exploiting NEB SHuffle strain (Section 3.1.18). Using this approach, we managed to successfully obtain high yields of cytoplasmically expressed nanobodies fused to partner of choice. In our case, this fusion product involved an expression enhancing tag with a bulky protease cleavage site and in frame fusion of Nup93 with varying linker lengths in between. By optimizing linker lengths and orientation of fusion partners in the final products, at the end, we were able to successfully incorporate Nup93n into the Nup62·58·54 complex (Section 3.1.18).

4.1.4. Mechanism of cargo-RanGTP cooperativity on CRM1

CRM1 (or XPO1) is an essential nuclear transport receptor involved in exporting hundreds of structurally unrelated cargoes out of the nucleus. The versatility in CRM1 function is rooted in the unique manner in which it binds its cargoes compared to other commonly known transport receptors. In comparison to the other known exportins such as CAS, Exportin-t and Exportin-5, where the export cargo interacts directly with both the exportin and RanGTP, CRM1 export complexes show no direct interaction between the cargo and RanGTP (Güttler and Görlich, 2011). Cargo-recognition site in CRM1 has been shown to be located in a narrow pocket formed by HEAT repeats 11A and 12A, on the outer surface of CRM1 ring, that is lined mainly by the hydrophobic (Φ) residues and is thus referred to as the hydrophobic cleft (or HC) (Monecke et al., 2009). A majority of export substrates bind CRM1 through a 10-15 residue long peptide stretch that fits into the rigid Φ pockets formed by the Φ residues of its HC (Güttler et al., 2010). RanGTP binding occurs on the inner surface of CRM1 near the N and C-terminal helices. CRM1 binds RanGTP and cargo in a cooperative manner as it was observed that CRM1 binding to RanGTP increases its affinity for cargo binding by ≈ 1000 fold and *vice-versa*. In the absence of a direct interaction between RanGTP and cargo in CRM1 export complex, how is this cooperative binding achieved has been an intriguing question in the field. To explain this nature, an allosteric mechanism for RanGTP-cargo cooperativity on CRM1 was proposed where binding of RanGTP on the inner surface of CRM1 induces changes in its overall conformation bringing it to a conformational state compatible with cargo binding. Recent structural characterizations of CRM1 in different functional complexes, that include structures of the ternary CRM1·RanGTP·SPN1 export complex (Monecke et al., 2009), intermediate CRM1·SPN1 (Dong et al., 2009) and CRM1·RanGTP (Güttler et al., 2010) complexes, cytoplasmic CRM1·RanGTP·RanBP1

4. Discussion

disassembly complex (Koyama and Matsuura, 2010) and unliganded scCRM1 (Saito and Matsuura, 2013) have revealed many features of this allosteric mechanism of RanGTP-cargo cooperativity.

An 'open' hydrophobic cleft as observed in the CRM1-RanGTP-SPN1 export complex and the intermediate CRM1-SPN1 and CRM1-RanGTP complexes is a conformation CRM1 adopts in order to accommodate and bind the cargo-NES in the HC. In contrast, in the crystal structures of the CRM1-RanGTP-RanBP1 disassembly complex (a complex that forms during cargo-unloading stage of CRM1 when Y-phosphate in RanGTP is hydrolyzed by concerted actions of RanBP1/BP2 and RanGAP), a relatively 'closed' HC conformation was observed. Similar closed HC conformation was also present in the crystal structure of free or unliganded scCRM1 and ctCRM1. From these structures, it has become clear that the 'opening' and 'closing' of the HC during cargo and RanGTP binding/unbinding events is a consequence of overall conformational changes in CRM1. One of the structural rearrangements in CRM1 that is readily visible is the conformational change in the beta-hairpin loop present at the base of HEAT repeat 9 (termed as acidic loop due to the acidic residues present at its tip that interact with RanGTP in the export complex). The acidic loop adopts an extended conformation in the cargo-bound CRM1 and a compact, 'flipped-back' conformation in cargo-free CRM1. Another local structural change observed in these crystal structures is the arrangement of the last α -helix in CRM1, the HEAT repeat 21B. In the cargo-bound state, this C-terminal helix is stacked parallel to other HEAT repeat helices in the CRM1 ring (as observed in the export complexes) whereas in the cargo-free state, it is oriented away from the toroid, extending to the center of the ring and touching the base of the HC.

Overall, it is clear from above studies that CRM1 is stabilized into two distinct conformational states during the transport process, a cargo-bound form and a cargo-free form. In order to identify the residues that are critical in adopting and/or stabilizing these distinct CRM1 conformations, we adopted an evolution-directed mutagenesis approach. CRM1 is an essential exportin in both lower eukaryotes (yeast) as well as mammals and is remarkably well conserved in the animal kingdom with human CRM1 sharing 98% identity with mouse (*Mus musculus*) and 96% identity with fish (*Danio rerio*). The conservation of CRM1's overall structure in the entire eukaryotic clade (from yeast to humans) indicates that the functional properties of the transport receptor have been essentially conserved across evolution.

Here, we report that CRM1 sequence analyzed from 58 diverse eukaryotic clades (*Figure 5.1*) starting from parasitic protozoans such as *Trichomonas*, *Toxoplasma* till higher eukaryotes including plants, fungi, amphibians and mammals revealed 20 absolutely identical residues that have not undergone any change throughout evolution. Such strong evolutionary conservation of

4. Discussion

these residues in the CRM1 sequence from the bottom of the eukaryotic life to its top suggests that these residues must have been under strong evolutionary pressure and any substitutions in them has not been tolerated. Since, macromolecular conservation heavily correlates to its functional roles, we asked how changes in these residues would affect CRM1 function. To achieve this, subtle point mutations in these 20 residues in mmCRM1 were designed carefully to prevent any drastic changes in the overall CRM1 architecture.

The effects for these mutations were readily visible in the binding assays, where binding of CRM1 point mutants to immobilized cargo (PKI-NES) and RanGTP was analyzed in comparison to the wild-type CRM1. In addition, the live export assays in HeLa nuclei analyzed the ability of these mutants to carry out multiple rounds of cargo-export through the NPCs in the context of *in-vivo* situation.

4.1.5. Various conformational states of CRM1 mutants

Mutations in the strictly conserved CRM1 residues resulted in generation of CRM1 mutants stabilized in different conformational states. Some of these mutations favored cargo-binding conformation while others stabilized cargo-free CRM1 conformation relative to the wild-type CRM1.

G502A/S, P427G mutants bound both PKI-NES and RanGTP much stronger compared to the wild-type CRM1. The binding to the cargo and RanGTP was high even when the other binding partner was absent. Thus, in these mutants a) CRM1 is stabilized in its 'cargo-bound' conformation and b) the binding cooperativity between RanGTP and cargo is essentially lost. These mutants did not carry out export of GFP-tagged cargo in the export assays as the cargo was predominantly localized in the nucleus even after prolonged incubation with mutant CRM1. In CRM1 export complex, both the residues G502 and P427 are located near the HC, with G502 being present towards the end of HEAT repeat 10B and P427 being part of the hinge region in the acidic loop formed by HEAT repeat 9.

As described above, HEAT repeat 9 undergoes large structural changes in response to cargo and RanGTP. It appears from the structural studies that an extended acidic loop is probably required to maintain an open HC for cargo and RanGTP binding. In our experiments, exchanging a rigid proline located in the hinge region of this loop to a flexible glycine resulted in an enhanced cargo and RanGTP binding indicating an open HC state and probably an extended AL conformation. Thus, it appears that the rigidity imparted by P427 in the hinge region of AC is required to stabilize the 'flipped back' conformation of the AL.

Comparing the structures of CRM1 in the export complex and unliganded CRM1 also shows rearrangement of side chains in the helices surrounding the HC. In our experiments, substituting

4. Discussion

G502 in the HEAT repeat 10 with alanine (or serine) that differ from it in a single alkyl group, strongly stabilized CRM1 in a cargo-binding conformation. Since, an alanine or serine moiety is more rigid than glycine, it indicates that the flexibility in this region is of critical importance in stabilizing open and closed conformation of the hydrophobic cleft.

It will be interesting to study both these mutants by structural analysis to observe how these conformations are indeed stabilized. Furthermore, since both these mutants bind cargo in a very stable manner, even in the absence of RanGTP, they are interesting candidates for crystallization of weak cargoes that have low affinity for CRM1 and hence as such cannot be crystallized.

CRM1 mutations E513D, Q550N, Y551F, F554Y, M583L, F366M and Y788F resulted in a much stronger binding to the cargo than the wild-type CRM1. However, RanGTP cooperativity was maintained in this case. Both E513 and Q550 are located in the HEAT repeat 11 and are engaged in salt bridge interactions with R553 and K590 respectively, located in the adjacent HEAT repeats, to stabilize an open HC conformation in the export complex. A mutation in these residues, E513D and Q550N, led to an increased cargo and RanGTP binding thereby indicating that the cargo binding conformation is further stabilized. Substituted residues in both cases were highly similar in their nature to the original residues, yet such substitutions are not observed in nature. This suggests that in the cellular milieu, CRM1 sequence is highly optimized for binding to both high affinity and low affinity cargoes. The cargo used in our assays, PKI-NES, is an inherently weak cargo (Güttler et al., 2010). Although, we did not test these mutants for binding to strong cargoes such as supra-physiological NESs, it is clear that the affinities for weak cargoes can be enhanced drastically with these point mutations.

Y551 is present in the linker region connecting helix 11B and 12A and hydrogen bonding with the residues in the adjacent HEAT repeats in the export complex. F554 on the other hand, is located at the far end of the HC between helices 11B and 12A whereas M583 is located at the base of HC. Mutation F554Y results in an additional hydroxyl group in the HC preventing its closure. M583 interacts directly with the cargo and participates in the formation of Φ pockets. The homologous residue in yeast M592 was recently also reported to adopt a different conformation in the unliganded scCRM1 compared to its conformation on CRM1 export complex (Saito and Matsuura, 2013). Thus, mutations in these conserved residues Y551F, F554Y and M583L might stabilize cargo-binding state of CRM1 by forcing an open HC conformation due to steric clashes with the neighboring residues. F66M and Y788F CRM1 mutants, on the other hand, harbor point substitutions in the conserved residues that are located far away from the HC. Still, these mutants resulted in much higher cargo and RanGTP binding compared to the wild-type CRM1.

4. Discussion

This further confirms that the evolutionary conserved residues indeed are crucial in stabilizing one or the other CRM1 conformation.

CRM1 mutations L80I, W91F, E176Q, F179M, L625A and D1017A resulted in extremely low affinities for PKI-NES and RanGTP binding. These mutants were predominantly stabilized in the cargo-free conformation. L80I (HEAT 2B), W91F (HEAT 2B), F179M (HEAT 4B), L625A (HEAT 13A) and D1017A (HEAT 21A) are present far away from the NES-binding site (HEAT 11 and 12) and hence do not participate in cargo-interaction or a direct RanGTP binding. E176 is however, involved in electrostatic interaction with RanGTP. Interestingly, all these mutants could bind a stronger NES (Super-PKI) (*Figure 5.5*). Super-PKI has very high affinity for CRM1 binding due to its rational design (Güttler et al., 2010) and is capable of forcing CRM1 into cargo-binding conformation even in the absence of RanGTP. Strong binding of these mutants to Super-PKI NES suggested that these mutants are indeed capable of binding cargoes, however, for weaker cargoes such as PKI-NES, the binding equilibrium is shifted towards the cargo-free state.

4.1.6. Inherent structural flexibility forms the basis of NPC's architectural and functional machinery

In this project, we analyzed the structural features of the most versatile nuclear transport receptor CRM1 and the least understood NPC architectural element, the Nup62·58·54 nucleoporin complex. Despite the diverse functional roles of CRM1 and Nup62·58·54 complex in forming the overall NPC functional and architectural machinery, inherent structural flexibility is a common theme and represents the dynamic nature of this machinery. While the structural flexibility in case of CRM1 is more pronounced and like in case of all other known NTRs, involves large-scale rigid body movements of its α -helical solenoidal fold in response to cargo binding and unbinding, the structural flexibility in case of Nup62·58·54 complex is more likely to be confined to the flexible hinge regions connecting the rigid α -helical, β -propeller or coiled-coil domains. This inherent structural flexibility poses serious challenges to its crystallization and structure determination. We overcame this problem for crystallization of Nup62·58·54 complex by selecting nanobodies that bind the flexible interfaces of Nup62·58·54 complex and in-turn rigidify it. CRM1 has already been crystallized in functional complexes with RanGTP and cargo, which also serve to rigidify its structure. For understanding the structural basis behind various conformational states of CRM1, we identified residues critical in imparting CRM1 its structural flexibility through mutating its evolutionary conserved residues.

Our experiments reveal that nanobodies can act as powerful tools in crystallization of flexible molecules and most likely can also be used for crystallization of large NTR-cargo complexes, and other nucleoporin subcomplexes that have not been crystallized so far due to their structural

4. Discussion

flexibility. They can also be used to stabilize weak interacting interfaces, such as those between nucleoporin subcomplexes, to gain information about nature of these interface regions. In our experiments, interface between Nup93 and Nup62·58·54 complex was surprisingly found to be very weak and might reflect a general nature of interaction between subcomplexes of vertebrate NPCs. Our approach of crystallizing subunits from different subcomplexes with the help of stabilizing nanobodies will result in a clearer picture of nucleoporin interaction and arrangement in the NPC and can be extended to other nucleoporin assemblies and NTR-cargo complexes. Moreover, we could also establish that single residues play a crucial role in stabilizing conformations of CRM1 and such evolutionary based mutagenesis approach can be applied to lock NTRs in one or other functional conformation, resulting in a rigid structure.

As of now, we are far from understanding structural and functional properties of NPCs at an atomic scale resolution. Integrative approaches involving X-ray crystallography, NMR, Cryo-EM, single-particle techniques, super-resolution microscopy along with development of novel biochemical tools will help us understand the nature of this macromolecular machine. Studying differences in the nature of nucleoporin interaction in the NPCs from lower and higher eukaryotes will further highlight the evolution undergone by this giant machinery.

5. Materials and Methods

5.1. Materials

Chemicals

Chemicals of highest grade purity were purchased from the following suppliers-Applichem GmbH (Darmstadt, Germany), Calbiochem (Part of Merck Millipore, Darmstadt, Germany), Carl Roth GmbH (Karlsruhe, Germany), GE Healthcare Europe GmbH (Friedberg, Germany) Life Technologies GmbH (Darmstadt, Germany), Merck Co. (Darmstadt, Germany), Pharmacia (Uppsala, Sweden), Roche Diagnostics (Mannheim, Germany), Serva Electrophoresis GmbH (Heidelberg, Germany), Sigma-Aldrich Co. (St. Louis, MO, USA) or Thermo Fisher Scientific (Waltham, MA, USA).

Instruments

Thermocycler	(SensoQuest, Göttingen, Germany)
DNA Gel Imager	(Vilber Lourmat, Eberhardzell, Germany)
MicroPulser electroporator	(Biorad, München, Germany)
Nanodrop 2000c	(PEQLAB Biotechnologie GmbH, Erlangen, Germany)
ISF1-X Climo-Shaker Incubator	(Kühner AG, Basel, Switzerland)
Äkta purifier	(Pharmacia, Uppsala, Sweden)
Sonifier	(Branson Ultrasonics, CT, USA)
Biophotometer plus	(Eppendorf AG, Hamburg, Germany)
Labfors3 bioreactor	(Infors, Bottmingen, Switzerland)
Pipettman pipettes	(Gilson, WI, USA)
5424 tabletop centrifuge	(Eppendorf AG, Hamburg, Germany)
5424 R tabletop centrifuge	(Eppendorf AG, Hamburg, Germany)
Haraeus multifuge 3SR	(DJB Labcare, Buckinghamshire, United Kingdom)
Sorvall RC6 plus centrifuge	(Thermo Fisher Scientific, Waltham, MA, USA)
Sorvall WX Ultra90 centrifuge	(Thermo Fisher Scientific, Waltham, MA, USA)
Sorvall Discovery M120 centrifuge	(Thermo Fisher Scientific, Waltham, MA, USA)
MALS miniDAWN TREOS	(Wyatt Technology Europe GmbH, Dernbach, Germany)
QELS Dynapro Nanostar	(Wyatt Technology Europe GmbH, Dernbach, Germany)
TCS SP5 confocal microscope	(Leica Microsystems, Mannheim, Germany)
Cartesian 8-channel nanodispenser	(Genomic solutions/Digilab Inc, MA, USA)
Freedom Evo liquid	(Tecan Group Ltd., Männedorf, Switzerland)

5. Material and Methods

Handling robot

Rock Imager automatic (Formulatrix Inc, Waltham, MA, USA)

Imaging system

Buffers and standard solutions

10x PfuS polymerase buffer	20mM Tris-HCl pH9.0, 250mM KCl, 15mM MgSO ₄ , 100mM (NH ₄) ₂ SO ₄ , 1% Tween20, 1mg/ml BSA.
10x DNA ligase buffer	500mM Tris pH7.5, 100mM MgCl ₂ , 100mM DTT, 10mM ATP, 250g/ml BSA.
50x TAE (Tris-Acetate-EDTA) 242g	Tris-base, 57.1ml Acetic acid, 100ml 0.5M EDTA buffer pH8.0, ddH ₂ O to 1l.
2x SDS sample buffer	125mM Tris pH6.8, 3% SDS, 50mM DTT, 1M Sucrose,
10x SDS-PAGE running buffer	300g Glycine, 60g Tris-base, 24g SDS, ddH ₂ O to 2l.
Orange-G DNA loading buffer	10mM Tris-HCl pH8.0, 10mM EDTA pH8.0, 50% Gly- cerol (w/v), 0.25% Orange-G (w/v)
Colloidal coomassie	0.08% (w/v) Coomassie brilliant blue G250, 1.6% (v/v) ortho-phosphoric acid, 8% (w/v) Ammonium sulfate, 20% (v/v) Methanol.
Coomassie stock solution	2% (w/v) Coomassie brilliant blue G250 in 50% ethanol
20X E-Mix (energy-regenerating System)	20mM HEPES pH 7.5, 250mM Sucrose, 200mM Creatine phosphate, 10mM ATP, 10mM GTP, 1mg/ml Creatine kinase
10X-Ran Mix regenerating system)	80μM RanGDP, 1μM RanBP1, 8μM RannGAP, 16μM (Ran- NTF2

Software

Lasergene 10	(DNA-Star, Madison, WI, USA)
GeneDesigner 2	(DNA2.0, Menlo Park, CA, USA)
Oligo 6.8	(Molecular Biology Insights, Cascade, CO, USA)
Photoshop CS5	(Adobe Systems, San Jose, CA, USA)
Papers 2	(Mekentosj, Aalsmeer, Netherlands)
PyMol	(Schrödinger, Portland, OR, USA)
Scaffold 3	(Proteome Software, Portland, OR, USA)
Illustrator CS5	(Adobe Systems, San Jose, CA, USA)

5. Material and Methods

Microsoft office 2011	(Microsoft Corporation, Redmond, WA, USA)
Bookends	(Sonny Software, Chevy Chase, MD, USA)
LASAF	(Leica, Manheim, Germany)
Mac OS X Vers.10.6.7	(Apple, Cupertino, CA, USA)
UNICORN 5.0	(Amersham Biosciences, Uppasala, Sweden)
Astra V	(Wyatt Technology, Santa Barbara, CA, USA)
Virtual Box	(Oracle Corporation, Redwood Shores, CA, USA)
Rockmaker	(Formulatrix Inc, Waltham, MA, USA)

Online tools

BLAST	http://blast.ncbi.nlm.nih.gov/Blast.cgi
ClustalW	http://ebi.ac.uk/Tools/msa/clustalw2
Uniprot	http://uniprot.org
COILS	http://embnet.vital-it.ch/software/COILS_form.html
Paircoils	http://gropus.csail.mit.edu/cb/paicoil/cgi-bin/paircoil.cgi
PSIPRED	http://bioinf.cs.ucl.ac.uk/psipred
ProtParam	http://web.expasy.org/protparam
Swiss Model	http://swissmodel.expasy.org/workspace
Jalview	http://jalview.org
Protein Data Bank	http://rcsb.org/pdb

E.coli growth media

LB liquid medium	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl
2YT liquid medium	1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl
TB liquid medium	1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 72mM K ₂ HPO ₄ , 17mM KH ₂ PO ₄
Culture plates	1.5% agar in LB or 2YT medium

All media were autoclaved before use. Both LB and 2YT media were supplemented with 2% glycerol and 30mM K₂HPO₄ for bacterial culture. Specific antibiotics at final concentrations of

5. Material and Methods

ampicillin (100 µg/ml), kanamycin (25 µg/ml) and spectinomycin (50µg/ml) were added in the media for plasmid selection and bacterial cultivation.

***E.coli* strains**

Cloning strains

TOP 10F' (Invitrogen)	F'(lacIqTn10(TetR)) <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU galK rpsL endA1 nupG</i>
NEB turbo (New England Biolabs)	F' <i>proA⁺B⁺ lacI^q</i> Δ <i>lacZM15</i> / <i>fhuA2</i> Δ(<i>lac-</i> <i>proAB</i>) <i>glnV galK16galE15R(zgb-210::Tn10)Tet^SendA1</i> <i>thi-1</i> Δ(<i>hsdS-mcrB</i>)
NEB 10-beta England Biolabs)	Δ(<i>ara-leu</i>) 7697 <i>araD139 fhuA</i> Δ <i>lacX74 galK16 galE15</i> (New e14- φ80 <i>dacZ</i> ΔM15 <i>recA1 relA1 endA1 nupG rpsL</i> (Str ^R) <i>rph spoT1</i> Δ(<i>mrr-hsdRMS-mcrBC</i>)

Protein expression strains

BLR (Invitrogen)	F' <i>ompT hsdS_B (r_B⁻ m_B⁻) gal dcm</i> Δ (<i>srl-recA</i>) 306::Tn10- <i>tetA</i>
NEB Express England Biolabs)	<i>fhuA2 [lon] ompT gal sulA11 R (mcr-73::miniTn10--(New</i> <i>Tet^S)2 [dcm] R(zgb-210::Tn10--Tet^S) endA1</i> Δ(<i>mcrC-</i> <i>mrr</i>)114::IS10
NEB Shuffle (New England Biolabs)	F' <i>lac pro lacIq</i> / Δ(<i>ara-leu</i>)7697 <i>araD13 fhuA2</i> Δ(<i>lac</i>)X74 Δ(<i>phoA</i>)PvuII <i>phoR ahpC* galE</i> (or U) <i>galK</i> Δ <i>λatt::pNEB3-r1-cDsbC</i> (SpecR, <i>lacIq</i>) Δ <i>trxB</i> <i>rpsL150(StrR) Δgor</i> Δ(<i>malF</i>)3

Primers

Highly pure oligonucleotides purified via desalting or HPLC, were synthesized from Sigma-Aldrich.

Plasmids

All plasmids generated and/or used in this study were based on pQE-80 series (Qiagen) of vectors. *ColE1ori* (Qiagen)- 30-40 copies/cell; p15A (kind gift from R.Lutz)- 10-12 copies/cell; *CloDF3 ori-* (Duet vector system, Novagen)- 30-40 copies/cell.

5. Material and Methods

Gene synthesis

Customized DNA sequences with optimized codon usage for protein expression in *E.coli* were obtained by *de-novo* synthesis from GenScript (Piscataway, NJ, USA).

DNA Sequencing

Plasmids were sequenced by DNA sequencing service from SeqLab (Göttingen).

Crystallization screens

96-tube format crystallization conditions screening suites – Classics, Classics II, Classics Lite, JSCG+, ProComplex, PEGs, PEGs II, PACT, CompAS were obtained from Qiagen (Hilden, Germany); MIDAS was obtained from Molecular Dimensions (Suffolk, UK); Index was obtained from Hampton Research (Aliso Viejo, CA, USA) and Wizard 1+2, Wizard 3+4 were obtained from Emerald Biosystems (Bainbridge island, WA, USA).

5.2. Methods

5.2.1. Standard techniques in Molecular Biology

Polymerase Chain Reaction (PCR)

To amplify target DNA, oligonucleotide primers complementary to the region to be amplified, were synthesized with desired restriction enzyme recognition sites at their 5' ends. The DNA was amplified using polymerase chain reaction (Mullis et al., 1992) on the template DNA. High-fidelity PfuS polymerase (produced in the lab by S.Frey) was used for all standard cloning reactions.

A PCR mix was prepared as follows:

50-100ng template DNA
1X dNTP mix (from 10X stock dNTP mix containing 2.5mM each dATP, dCTP, dGTP, dTTP)
1 μ M forward primer (from 100 μ M stock solution)
1 μ M reverse primer (from 100 μ M stock solution)
1X PfuS polymerase buffer (from 10X stock solution)
1ng/ μ l PfuS polymerase (from 100ng/ μ l stock solution)
ddH ₂ O to adjust the final volume

The protocol used for amplification included:

Initial denaturation at 98°C for 2 min
25-30X amplification cycles
Denaturation at 98°C for 30 sec
Primer annealing at $T_{\text{annealing}}$ for 30 sec
DNA elongation at 68°C for T_{amp}
Final extension at 68°C for 7 min

5. Material and Methods

$T_{\text{annealing}}$ or annealing temperature implies average annealing temperature specific for primer pair based on primer length and composition. It was calculated using Oligo 6.8.

T_{amp} or amplification time implies time duration for primer elongation on the given template. It depends on the amplification rate of the polymerase and length of template to be amplified. (For pfuS polymerase, amplification rate is $\sim 30\text{sec/kb}$)

Mutagenesis PCR

For deletions of small parts from the target DNA, additions of few base pairs to the target DNA or for single amino acid substitutions, mutagenesis PCR was carried out. HPLC-purified, 5'-phosphorylated primers were designed for amplifying the entire plasmid. PCR amplification was carried out as above. After amplification, the DNA was precipitated by addition of isopropanol (1:1 v/v) followed by centrifugation for 10 minutes (room temperature, 13,000 rpm). The supernatant was removed as completely as possible and the resulting pellet was suspended in Dpn1 digestion mix (25 μl H₂O, 3 μl NEB buffer 4 and 2 μl Dpn1 enzyme). The digestion was carried out at 37°C for 1h to digest the methylated non-mutated template DNA. The final product was purified by DNA gel electrophoresis and ligated (see below). 1-2 μl of the ligation mix was transformed into electrocompetent *E.coli* cloning strains and plated on selection medium. The plates were incubated overnight at 37° C.

Agarose gel electrophoresis

Agarose gel electrophoresis separates DNA fragments based on their size. The protocol used was employed according to (Sambrook et al., 1989) 0.8-2% (w/v) agarose gel was prepared by dissolving agarose in TAE buffer followed by heating and cooling down. DNA samples mixed with 1:10 (v/v) Orange-G DNA buffer were loaded on the gel and electric potential of $\sim 160\text{V}$ was applied. Electrophoresis was carried out for $\sim 30\text{min}$ and the DNA fragments were visualized under UV light. A small amount of DNA intercalating dye, Ethidium bromide (EtBr), was added to liquid agarose prior to gelation to promote visualization of DNA under UV light.

DNA gel extraction and purification

After agarose gel electrophoresis, size-separated DNA fragments were cut out from the gel. DNA was extracted from the agarose gel slices using Zymoclean Gel DNA recovery kit (Zymo Research, Freiburg, Germany) as per manufacturer's instructions.

DNA concentrations were determined with NanoDrop ND-2000 spectrophotometer.

DNA digestion with restriction enzymes

All restriction enzymes were purchased from NEB (New England Biolabs, Beverly, MA, USA). The enzymes were mixed with the DNA to be cleaved in the presence of buffers and reagents according to the manufacturer's instructions. At least a two-fold excess of restriction enzyme

5. Material and Methods

over DNA was used. The digestion reaction was performed at 37° C for 1h. For overnight digestion, the amount of enzyme used was reduced. After DNA cleavage, alkaline phosphatase treatment (FastAP, Fermentas) was carried out for 10min to remove phosphate groups from the ends of vector DNA to prevent religation during the subsequent ligation step. The cleaved DNA was purified on agarose gels as above and size-specific bands were used for ligation.

DNA Ligation

A ligation reaction mix was prepared as follows:

~20ng/kb vector DNA (~30fmol)
~40ng/kb insert DNA (~60fmol)
1X DNA ligation buffer (from 10X stock solution)
1X E-mix (from 10X stock solution)
10ng/μl T4 DNA ligase (produced in the lab by S.Frey)
ddH ₂ O to adjust the final volume to 10μl

The ligation reaction was typically carried out at 37°C for 30min for sticky-ends and 16°C overnight for blunt-end DNA. 1μl of ligation mix was transformed in *E.coli* and plated onto selection media for further analysis.

DNA transformation into *E.coli* by electroporation

Electro-competent *E.coli* cells were prepared in the lab by G. Kopp. For electroporation, 1μl of ligation mix was mixed with 45μl of electro-competent *E.coli* cells in an electroporation cuvette (BioRad, Burlington, USA). Electroporation was performed with GenePulser™ (BioRad) according to manufacturer's instructions. The electroporated cells were suspended in 1ml 2YT medium supplemented with 0.2% glucose and incubated at 37°C for 1h in a Thermomixer (Eppendorf, Germany) while shaking. Cells were subsequently plated on selective media containing appropriate antibiotics for plasmid selection and incubated overnight at 37°C.

Colony PCR

Colony PCR is an analytical method for identifying correct clones from bacterial colonies obtained on selection plates after ligation. In this method, bacterial cells are directly used as template for amplification rather than purified DNA. The technique relies on the fact that high temperatures used in the PCR reaction are sufficient to lyse the cells and expose the DNA for amplification. Therefore, a large number of clones can be screened simultaneously for identifying the gene of interest. For this, single bacterial colonies were picked up from plates with a toothpick or autoclaved pipette tips and thoroughly suspended in 50μl 2YT medium. 1μl of this mix was used in the PCR-mix (see above). Tth polymerase (produced in the lab by S.Frey)

5. Material and Methods

was used for colony PCR. The products were analyzed on by DNA gel electrophoresis (as above) and positive hits were identified.

Test protein expression

Another method for fast screening of clones is small-scale test expression of protein encoded by the gene of interest. For this method, bacterial colonies were suspended in 250µl 2YT medium as above and grown for 2-3h at 37°C in Thermomixer, shaking at 1200rpm. 50µl of cell suspension of non- induced sample was set aside prior to addition of IPTG. To the rest of cell suspension, 1mM Isopropyl-β-D-thiogalactopyranosid (IPTG) was added to induce expression of encoded protein and cells were grown further for 2h. Cells were spinned down in tabletop centrifuge and suspended in 100µl SDS sample buffer for analysis by SDS gel electrophoresis.

Plasmid isolation from E.coli

Small-scale plasmid purification was done using NucleoSpin[®] Plasmid Mini kit (Macherey Nagel, Germany) and a preparative-scale using NucleoBond[®] Xtra Midi kit (Macherey Nagel, Germany) according to the manufacturer's instructions. 250ml bacterial cultures containing the target plasmid were grown at 37°C in LB medium supplemented with appropriate antibiotic. G.Kopp provided the technical assistance for plasmid isolation.

DNA sequencing

Plasmid DNA was sent to Seqlab Sequence Laboratories (Göttingen, Germany) for DNA sequencing. The results were analyzed with the SeqMan program of the Lasergene10 software.

5.2.2. General protein expression and purification

Protein Expression

E.coli expression strains (Neb Express, BLR or NEB shuffle) were used for protein expression. Various parameters such as the nature of N-terminal affinity tags (expression enhancers, solubility enhancers), solubility of final product and yield of final product etc. determined the expression conditions like induction strength, expression time, and expression temperatures. Individual parameters were optimized individually for each protein, in an analytical-scale expression prior to large-scale expression.

In general, ~500ng plasmid DNA encoding the desired protein was transformed into electro-competent *E.coli* cells. Cells were recovered by incubating in 1ml 2YT medium at 37°C for 1h while gently shaking. For co-transformation of two or more plasmids, plasmids were mixed together and a final DNA concentration of not more than ~1000ng DNA was used for transformation. The transformed cells were plated on selective media containing appropriate antibiotics and incubated overnight at 37°C. A single colony was picked up and used for

5. Material and Methods

inoculation of 100-150ml 2YT or TB medium containing appropriate antibiotics. The primary culture obtained after overnight bacterial growth at 30°C was used to inoculate secondary or expression cultures in the same medium (containing antibiotics) used for primary cultures. The cultures were grown at 30°C in 5L flasks till an OD_{600nm} of 0.8-1.0 (for 2YT medium) or 2-3 (for TB medium) was obtained. Protein expression was induced by addition of Isopropyl-β-D-thiogalactopyranosid (IPTG) at a final concentration of 0.025-1 mM (depending upon expression temperature and time). The cultures were grown further for 3-6h (at 28-37°) or 14-16h (at 18°C). The cells were harvested by centrifugation (at 5000rpm 10-15min, 4°C) in centrifugation buckets. Prior to harvesting, 1mM phenylmethanesulfonylfluoride (PMSF) was added to the culture as protease inhibitor whereas 10mM ethylenediaminetetraaceticacid (EDTA) was added to the centrifugation buckets. The culture supernatant was removed as completely as possible and the pellets were suspended in Resuspension buffer (50mM Tris-HCl pH7.5, 200-300 mM NaCl, 2-10mM DTT) in a volume leading to an OD_{600nm} of ~200. Subsequently, the cells were either lysed directly or alternatively flash-frozen in liquid nitrogen and stored at -80°C till further processing.

Cell lysis

The cells were typically lysed by freeze/thawing followed by sonication (4X, 1min pulses with 50% duty cycle) on ice. Clarified solutions were obtained from the crude lysates by ultracentrifugation (38,000 rpm for 2-3h) at 4°C. The supernatant was carefully removed and used further as starting material for affinity purification.

Affinity chromatography with Ni⁺² matrix

Bivalent heavy metals such as Ni⁺², Co⁺² or Cu⁺² etc. have high affinities for binding to Histidine residues commonly found in proteins. These metal ions form a reversible complex with imidazole rings of the *His*-residues at physiological pH, which can be dissolved by addition of excess of imidazole at high concentrations. This principle forms the basis of Ni⁺² affinity chromatography used for purification of various poly-histidine tag containing recombinant proteins expressed in *E.coli*. The proteins were engineered to express a cleavable His₁₀ or His₁₄ tag at their N-terminus (or non-cleavable His₆ at C-terminus) for purification via Ni⁺² affinity chromatography.

In general, 24% Ni⁺² EDTA silica matrix (self-made by D.Görlich) was equilibrated with Resuspension buffer (as above) and mixed with clarified *E.coli* lysate containing recombinant protein. The mix was incubated at 4°C (cold room) for 1h while in constant rotation. 5-10mM imidazole was added to the binding mix to reduce non-specific binding of contaminant proteins. The mix was then applied to gravity-flow columns (Sigma-Aldrich) and the matrix was allowed to settle while the flow-through was collected. The affinity matrix was washed extensively (4-5

5. Material and Methods

Column Volume or CV) with Wash buffer (50 Tris-HCl pH7.5, 100-200mM NaCl, 2-5mM DTT, 5-10mM Imidazole). Bound protein was eluted off the matrix by addition of excess Imidazole (500mM Imidazole in wash buffer) in 500µl fraction volume. 1µl of each eluted fraction was blotted on nitro-cellulose membrane, which was stained with amido-black stain. The peak fractions containing highest amounts of protein were pooled together. Protein concentration was determined by measuring absorption at 260 and 280nm. Samples collected at each step of expression and purification were analyzed by SDS gel electrophoresis followed by Coomassie staining of protein bands. In some cases, buffer exchange was done using PD10 or Nap5 buffer exchange columns (GE Healthcare) to remove imidazole prior to freezing. Sucrose to a final concentration of 250mM was added to the purified proteins, which were flash-frozen in liquid nitrogen and stored at -80°C.

Affinity chromatography with IgG-sepharose matrix

ZZ-tag (derived from Protein A of *S.aureus*) is an expression (and solubility) enhancing affinity tag that binds strongly to the Fc region of IgG. Higher expression levels, greater solubility and reduced background levels of non-specific binders to IgG matrix are some advantages of using this tag. The ZZ affinity-tag was used in combination with His-tag to provide a complementary second round of purification in order to achieve higher purity or in analytical binding assays (see below) where low background binding of contaminant proteins was desired.

In general, IgG sepharose™ 6 Fast Flow (GE healthcare) was pre-washed 2-3X column volume with equilibration buffer (50 Tris-HCl pH7.5, 200-300mM NaCl) followed by washing with 0.2 M glycine buffer pH2.0 (3X column volume) and equilibrated again in the equilibration buffer (3X column volume). Recombinant protein containing N-terminal ZZ-tag was applied to the matrix and incubated for 2h at 4°C under rotating conditions. After binding, the matrix was washed thoroughly with equilibration buffer (3-4 CV) and the bound protein was eluted off the column by protease cleavage (see below). In case of small scale binding assays, SDS buffer was used for elution.

In-solution protease cleavage of affinity tags and reverse Ni⁺² chromatography

For most recombinant proteins expressed in this study, recognition sites for endoproteases followed the N-terminal affinity tag. The endoproteases such as TEV protease, SUMO (Small-ubiquitin-related modifier) protease and NEDD8 protease (S.Frey, manuscript in preparation) were expressed and purified in the lab (by S.Frey) and provided great tools for the design of orthogonal purification systems for engineered proteins. The recognition site for TEV protease is a small peptide motif (ENLYFQG) introduced at the C-terminus of the affinity tag, whereas SUMO and Nedd8 proteases recognize a tertiary-fold in the protein (SUMO or NEDD8) structure.

5. Material and Methods

Besides conferring protease cleavage sites, these proteases also served as expression and/or solubility enhancers.

For in-solution protease cleavage, *His*-tagged protease was added to the Nickel-eluate (imidazole exchanged using PD10 desalting column) in appropriate amount (100-fold excess of substrate over protease for TEV and NEDD8 protease, 1000-fold excess of substrate over protease for SUMO protease). The mix was incubated for 30min-1h for SUMO and NEDD8 proteases and overnight for TEV protease cleavage at 4°C.

The protease along with the cleaved-off *His*-tags, un-cleaved target protein and other non-specific contaminants were removed by rebinding of the mix to the Nickel matrix (described above). The flow-through was collected and used as purified protein for further analysis.

On-column protease elution

In order to achieve a higher level of purity, bound proteins can be eluted directly from the affinity matrix by protease cleavage. The method has advantage over other methods of protein elution (*eg.* Imidazole or SDS elution) that the non-specifically bound proteins are not eluted off the column. In addition, protein can be eluted directly in the storage buffer and therefore, it bypasses the need for another round of buffer exchange and removal of affinity tag by reverse Ni⁺² chromatography (as described above). Therefore, the method combines high specificity with time-efficiency to achieve large amounts of pure protein.

For on-column protease elution, proteins were bound to the Ni⁺² matrix as before. The column was plugged from the bottom and 1-2ml of Wash buffer (50Tris pH7.5, 100-200mM NaCl, 2mM DTT, 5-10mM Imidazole ± 250mM sucrose) was added. Untagged protease in appropriate amount (100-1000 fold less protease than estimated substrate amount) was added to the matrix and mixed. The digestion was allowed to proceed at 4°C for 30min-1h after which the flow through was collected. The matrix was washed with 2-3X bed-volume of Wash buffer and 500µl sized fractions were collected. The eluted fractions were analyzed by amido-black staining and peak fractions were pooled together.

Gel filtration chromatography

Gel filtration (*a.k.a.* size-exclusion chromatography) purifies proteins by separating them according to differences in their molecular size and shape. Thus, protein aggregates and degradation products can be efficiently removed from the target protein. The matrix used in gel-filtration chromatography comprises of porous beads, which based on their cut-off size, allow small molecules to enter, whereas large molecules are excluded and therefore, elute first from the column. Superdex™ 75 (separates proteins in the range of 3-70kDa) and Superdex™ 200

5. Material and Methods

(separates proteins in the range of 10-600kDa) columns (Amersham) were used for gel filtration chromatography on Äkta Purifier/Explorer system (GE Healthcare).

The columns were equilibrated with 1.5-2 CV of appropriate buffer (typically 20mM Tris-HCl pH 7.5, 100-200mM NaCl, 1-2mM DTT) prior to sample injection. The protein was concentrated (see protein concentration) to a final volume of 1-5ml (for HiLoad 16/60 columns) and aggregates were removed by ultracentrifugation at 30,000 rpm for 30min at 4°C. The sample was applied via the sample loop with the help of a syringe into the system. The gel filtration was carried out according to manufacturer's instructions and samples were collected after the void volume and peak fractions were analyzed by SDS-PAGE electrophoresis.

Anion-exchange chromatography

Ion-exchange chromatography resolves proteins according to differences in their surface charge. The method relies on the fact that surface charge of protein molecules is dependent upon the ionization state of amino acid residues, which in turn depend upon the pH of the buffer. Proteins tend to have a net negative surface charge at pH above their isoelectric point (pI) and a net positive charge at the pH below it. Therefore, a protein molecule at a pH above its pI binds positively charged matrix such as Q-sepharose. Different anion-exchange columns such as HiTrap Q FF, Mono Q HR 5/5 GL, Mini-Q 4.6/50 PE (GE healthcare) and Souce-15 Q (self-made by J.Schuenemann) were tested for binding and maximum resolution.

In general, the column was equilibrated with 5-10 CV of low-salt binding buffer or Buffer A (20mM Tris pH 7.5, 50mM NaCl, 1mM DTT) followed by high-salt elution buffer or Buffer B (20mM Tris pH 7.5, 2M NaCl, 1mM DTT) and again brought back to Buffer A. Before loading onto the column, the salt concentration in the protein-solution was adjusted to 50mM NaCl either by dilution or buffer-exchange using PD10 column. Protein was slowly applied to the column at the concentration of 0.5-1 mg/ml and a flow rate of 0.5-1ml/min. After binding, the column was washed thoroughly with 10-15 CV of Buffer A at the default flow-rate. Elution was done by gradual increase in the ionic strength by applying a gradient of 0-30 % Buffer B over 10 CV. 500µl fractions were collected during the gradient application. Peak fractions were analyzed by SDS-PAGE electrophoresis and pooled together for further analysis.

Laemmli SDS-PAGE electrophoresis

Denaturing sodiumdodecylsulfate-polyacrylamide (SDS-PAGE) electrophoresis separates protein according to their molecular weights and is a standard method for visual analysis of protein samples (Laemmli, 1970). SDS binds proteins and confers them with equal charge:mass ratio causing them to migrate under the influence of an electric field towards the positive electrode

5. Material and Methods

based upon their molecular weight. The proteins can then be visualized by staining the gel with Coomassie stain.

The composition of 7.5-16% gradient gels prepared in the lab by G.Kopp is as follows:

	Stacking Gel (4.5%) (100ml)*	Separating (7.5%) (200ml)*	gel	Separating gel (16%) (200ml)*
2M Tris pH8.8	-	40ml		40ml
0.5M Tris PH6.8	15ml	-		-
ddH ₂ O	68ml	107ml		32ml
2M Sucrose	-	-		10ml
Glycerol (87%)	-	-		8ml
SDS (10%)	1ml	2ml		2ml
Rotiphorese Gel 30	15ml	51ml		108ml
TEMED	100µl	120µl		120µl
APS (10%)	1ml	2x 580µl		2x 580µl

TEMED= Tetramethylethylenediamine
APS= Ammonium persulfate
* (Recipe for 10 gels)

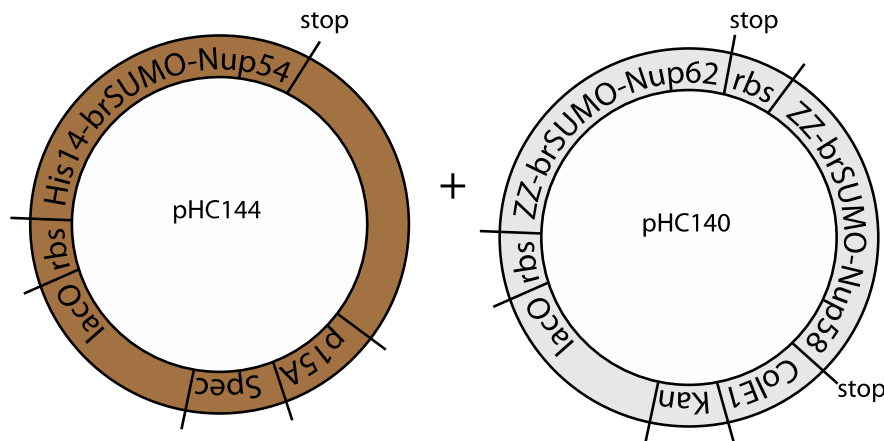
Prior to application on the gel, protein samples were mixed with equal volumes of 2X SDS sample buffer and boiled at 95°C for 5-8 min while shaking. The samples were centrifuged shortly and 5-15µl was applied to the gel. Electrophoresis was carried out at a constant current of 50mA till the loading front reached the bottom of the gel. Proteins were fixed by heating in 3% acetic acid and stained with approx. 1:1000 dilution of Coomassie stock solution. Gels were destained in ddH₂O and scanned using Perfection V700 photo scanner (Epson, Tokyo, Japan).

5.2.3. Specific purification protocols

5.2.3.1. *xlNup62·58·54* complex

Nup62·58·54 complex was purified by co-expression of His₁₄ brSUMO_Nup54 (pHC140) with ZZbrSUMO_Nup62_rbs_ZZbrSUMO_Nup58 (pHC140) in *E.coli* NEB Express (New England Biolabs) expression strain. Transformed cells were grown overnight in TB medium at 28°C in the presence of 25µg/ml Kanamycin and 50µg/ml Spectinomycin. Expression culture was grown by inoculating 30ml of the primary culture into fresh TB medium supplemented with antibiotics. Protein expression was induced at an OD_{600nm} of 2.5-3.0 by addition of 200µM IPTG. Cells were grown overnight at 18°C and processed as described before.

5. Material and Methods



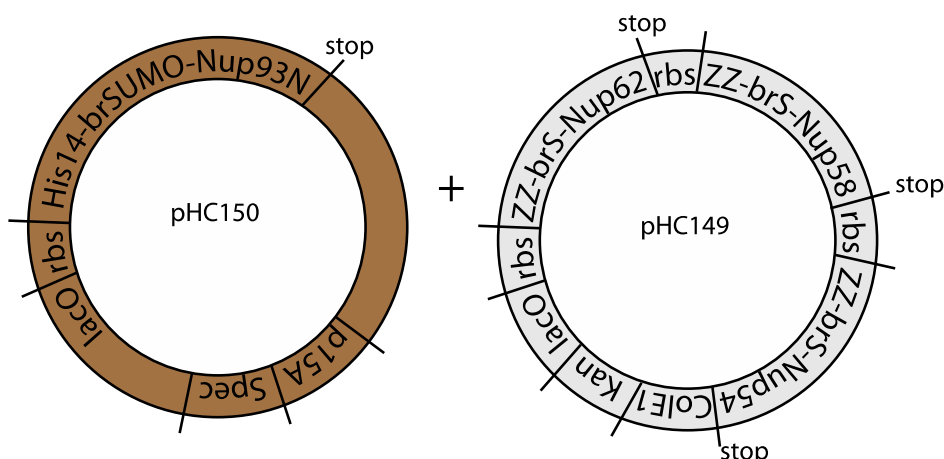
Steps followed for the purification of Nup62·58·54 complex:

1. **IMAC** (50mM Tris-HCl pH7.5, 200mM NaCl, 5mM DTT)
2. **On-column protease elution** (50mM Tris-HCl pH7.5, 200mM NaCl, 5mM DTT, 10mM Imidazole)
3. **Gel filtration chromatography** (20mM Tris-HCl pH7.5, 100mM NaCl, 2mM DTT) in Superdex™ 200 gel filtration column (GE Healthcare).
4. **Anion exchange chromatography** (Buffer A- 20mM Tris-HCl pH7.5, 50mM NaCl, 2mM DTT; Buffer B- 20mM Tris-HCl pH7.5, 2M NaCl, 2mM DTT). The complex eluted at 250-300 mM NaCl during gradient elution from MonoQ HR 10/10 anion-exchange column (GE healthcare).

5.2.3.2. *x*Nup62·58·54·93N complex

Nup62·58·54·93N complex was purified by coexpression of His₁₄ brSUMO_Nup93N (pHC150) with ZZbrSUMO_Nup62_rbs_ZZbrSUMO_Nup58_ZZbrSUMO_Nup54 (pHC149) in *E.coli* NEB Express (New England Biolabs) expression strain. Transformed cells were grown overnight in TB medium at 28°C in the presence of 25µg/ml Kanamycin and 50µg/ml Spectinomycin. Expression culture was grown by inoculating 30ml of the primary culture into fresh TB medium supplemented with antibiotics. Protein expression was induced at an OD_{600nm} of 2.5-3.0 by addition of 200µM IPTG. Cells were grown overnight at 18°C and processed as described before.

5. Material and Methods



Steps followed for the purification of Nup62·58·54·93N complex:

1. **IMAC** (50mM Tris-HCl pH7.5, 300mM NaCl, 5mM DTT)
2. **On-column protease elution** (50mM Tris-HCl pH7.5, 300mM NaCl, 5mM DTT, 10mM Imidazole)
3. **Gel filtration chromatography** (50mM Tris-HCl pH7.5, 200mM NaCl, 2mM DTT) in Superdex™ 200 gel filtration column (GE Healthcare).

5.2.3.3. Nanobody expression and purification

For periplasmic expression, nanobody sequences were cloned in pHEN6c vectors containing N-terminal secretion signal sequence fused in-frame to Nb sequence. Cytoplasmic expression of nanobodies was carried out in *E. coli* NEB SHuffle strains (New England Biolabs). His14-NEDD8-Nb constructs were transformed and grown overnight at 30°C in 2YT medium supplemented with 40% glycerol and 30mM K₂HPO₄. For nanobody expression, the culture was diluted with fresh medium to OD_{600nm} 0.8-1.0 and induced with 300µM IPTG and grown further for 5h. The cells were harvested and processed as before. The nanobodies were purified by IMAC in a buffer lacking DTT (50mM Tris-HCl pH 7.5, 300mM NaCl) and eluted with imidazole. The eluent was applied to Superdex™ 75 column equilibrated in 50mM Tris-HCl, 200mM NaCl and the peak fractions were collected.

5.3. Nanobody conjugation to xINup62·58·54 complex

For complex formation between Nup62·58·54 complex and Nb15, purified Nup62·58·54 complex (10-20 mg) and His14-NEDD8-Nb15 were mixed together in 1:1 molar ratio and incubated at 25°C for 30min. The complex was then bound to 2-3 ml Ni²⁺-EDTA affinity matrix (pre equilibrated with 50 Tris-HCl pH7.5, 200mM NaCl, 1mM DTT) and eluted via on-column cleavage by NEDD8 protease (1:100 protease: substrate ratio). The complex was applied to HiLoad Superdex™ 200 16/60 gel filtration column equilibrated in 20mM Tris-HCl pH 7.5, 100mM NaCl, 1mM DTT). The peak fractions from gel filtration column were analyzed by SDS PAGE, pooled

5. Material and Methods

together, and concentrated by ultrafiltration using Vivaspin 6 centrifugal concentrators (Sartorius, Göttingen) according to manufacturer's instructions. The concentrated samples (8-10 mg/ml) were frozen in 100µl aliquots and used for crystallization.

5.4. Limited proteolysis of α /Nup62·58·54 complex

For probing the structure of core Nup62·58·54 complex, limited proteolysis with chymotrypsin was performed. For analytical scale digestion, 1mg/ml Nup62·58·54 complex was incubated with chymotrypsin (Roche) in 1:1000(w/w) protease: substrate ratio at 4°C. Samples were taken out before protease addition and at indicated time intervals after protease addition, mixed with equal volumes of SDS sample buffer and boiled immediately at 95°C for 10min. The samples were analyzed on SDS-PAGE to identify stable fragments.

For large scale limited proteolysis, 1ml of 10mg/ml Nup62·58·54 complex was incubated with chymotrypsin (1:1000 w/w) for 2h at 4°C. The digested sample was immediately loaded on HiLoad Superdex™ 200 16/60 (Pharmacia, Uppasala, Sweden) column pre equilibrated with gel filtration buffer (20mM Tris-HCl pH 7.5, 100mM NaCl, 2mM DTT). Gel filtration was performed as above and eluted fractions were analyzed by SDS-PAGE. The gel was stained with colloidal coomassie and indicated bands were analyzed by Mass spectrometry.

5.5. Molecular weight determination by Static light scattering (SLS)

Multi-angle Light scattering (MALS) is an analytical technique to determine the absolute molecular mass, size and conformation of protein molecules based upon scattering of light by them *in solution*. In MALS system, a beam of polarized light is focused onto the sample molecule and the scattered light is detected with photo detectors at various angles to the incident light. To determine the molecular weight (MW) and root mean square radius (rms) of the protein, static light scattering coupled with SEC (size exclusion chromatography) was used. 200µl of purified protein sample (1mg/ml) was loaded to Superdex™ 200 10/300 GL (GE healthcare) column and gel filtration chromatography was performed as described above. As the protein eluted from the column, several online detectors coupled to the column analyzed it. The first detector in series is the refractive index (RI) detector (Shodex RI-101, Showa Denko KK, Japan) With this detector, the refractive index and the UV-absorbance at 280nm of the protein sample was measured and the concentration of the protein was calculated. The scattered light was then recorded at three different angles by in-built detectors in the miniDAWN™ TREOS detector (Wyatt technology Europe). From the intensity of scattered light at different angles and the protein concentration, the absolute molecular weight and rms radius was determined using Astra® 6 software (Wyatt Technology).

5. Material and Methods

5.6. Polydispersity analysis using Dynamic Light scattering (DLS)

DLS measures the intensity of scattered light by molecules in solution as a function of time. Random Brownian motion of macromolecules leads to time-dependent fluctuations in the intensity of the scattered light. In DLS, a fast photon counter measures these fluctuations. The fluctuations can be quantified and directly relate to the hydrodynamic radius of the macromolecule. In general, large molecules diffuse slower than the small molecules. Thus, DLS can be used to determine the quality of a protein solution. A solution containing different particle populations such as protein aggregates, higher order oligomers and monomeric protein is considered polydisperse and not suitable for crystallization whereas a sample with a single homogeneous population of target protein is monodisperse and a requirement for crystallization.

20µl of protein sample was pipetted in disposable plastic microCuvettes (Wyatt technology) and inserted into the sample holder of DynaPro NanoStar™ (Wyatt Technology) for batch-mode measurement. Measurements were performed according to the manufacturer's instructions and data analysis was done using Astra® 6 software (Wyatt Technology).

5.7. Mass spectrometry

Mass-spectrometric analysis was used for the identification of the protease-resistant regions of Nup62·58·54 complex subjected to limited proteolysis and for the identification of neighboring cysteine residues in the Nup62·58·54 complex after crosslinking. In each case, SDS PAGE analysis of the sample followed by colloidal Coomassie staining was used as a starting material for MS analysis (done in the lab of Prof H.Urlaub).

5.8. Protein crystallization

Samples of highest possible purity at high concentrations were used for obtaining protein crystals for X-ray crystallography. Vapor-diffusion method was used as a method of choice as it is the most commonly used method and can be used in combination with automated set ups for screening large amounts of crystallization conditions. In this method, a small volume of the concentrated protein sample is mixed with the same volume of crystallization solution (e.g. buffer, salt or precipitant; generally a combination of all three) and placed in a sealed reservoir containing the crystallization solution. In vapor diffusion with hanging-drop method, the sample is placed on a cover slip that is used to seal the reservoir, whereas in the sitting-drop method, the sample is placed inside a well surrounded by the reservoir solution. The difference in the concentration of components between the sample and reservoir solution causes vapor diffusion of the solvent until equilibrium is reached. During equilibrium, super saturation of the protein is achieved in the drop, which may lead to crystallization. All crystallography experiments were

5. Material and Methods

performed in the central crystallization facility of MPI for biophysical chemistry headed by Dr. V.Pena.

5.8.1. Sparse matrix screening with 96-well MRC sitting drop plates

For obtaining initial crystallization hits, large numbers of commercially available crystallization screens were tested (See Materials section). Crystallization drops were set up in 96 well sitting drop MRC plates (Hampton Research) using Cartesian Microsys (Genomic Solutions /Digilab) Nano dispenser robot. 60nl of reservoir solution (containing the precipitant) was pipetted followed by 60nl of protein (8-16 mg/ml) into the round wells of MRC plates. Two different protein concentrations were used for the initial screens. After pipetting, the plates were sealed tightly with a clear sealing tape and stored at 20°C or 4°C (both temperatures were screened for each condition) in automated robotic imager (Formulatrix). Crystal growth was monitored over time using RockImager software (Formulatrix), which was used to take images of the crystallization drops at regular intervals.

5.8.2. Grid screening with 96-well MRC sitting drop plates

Initial crystallization hits obtained from sparse matrix screens were further optimized for increased crystal quality and size by screening around the initial crystallization conditions (grid screening). Grid screens were designed using RockMaker software (Formulatrix) by varying the concentration and pH of components in the initial hit condition around their original concentration and pH. The plates were pipetted as above and monitored on regular intervals.

5.8.3. 24-well Linbro hanging drop plates

For further refinement of crystal quality and to obtain larger crystals, manual drops were set up with hanging drop method using 24-well Linbro plates. Successful crystallization conditions from the grid screen were pipetted in the wells of 24-well plates. 1-2µl of reservoir solution was mixed with 1-2µl of protein solution on a cover slip. Generally, the volume ratio between reservoir and protein drop was also varied between 1:1, 1:1.5 or 1:2 and 3-4 drops were set per coverslip. After the drops were placed, the coverslip was inverted and placed on the well containing 1ml reservoir solution. The upper boundary of the wells was pre greased with silicone grease (Jena Bioscience) for sealing the crystallization chamber. The plates were left undisturbed at 20°C or 4°C and monitored at regular time intervals by visual examination under Leica MZ6 microscope (Leica Microsystems).

5.8.4. Additive screen set up

Additives such as traces of metal ions, organic polymers, and alcohols etc. are known to improve crystal quality by manipulating protein-protein or protein-solvent interactions. Commercially

5. Material and Methods

available 96 tube format additives screens (Hampton research) were used to set up crystallization plates with optimized precipitant conditions. 16 μ l of additive was mixed with 64 μ l of precipitant in the reservoir. 120nl of reservoir containing the additive and 120nl of protein were added to the wells of 96-well MRC plate using the Nano dispenser robot as described above. The plates were sealed and imaged at regular intervals with the automated robotic imager (Formulatrix).

5.8.5. Flash cooling in liquid nitrogen and data collection

Protein crystals were carefully fished out of the mother liquor using Nylon loops mounted on a magnetic head (mounted CryoLoops, Hampton Research) of 0.02-0.5 μ M in size. Prior to freezing, crystals were first pre equilibrated in a suitable cryo-protectant (glycerol, ethylene glycol, xylitol, sucrose, PEG 400) for preventing ice-crystal formation inside or surrounding the protein crystal during freezing. Typically, a series of drops containing increasing concentration of cryo-protectant in the reservoir solution were pipetted. The crystal was transferred from one drop to another with a 30sec-3min incubation-period per drop. The crystal was finally fished out of the drop and plunged in liquid nitrogen for freezing. Diffraction data from cryo-cooled crystals were collected at 100 K on a Pilatus detector at the Swiss Light Source (Switzerland) beamline PXII.

5.8.6. In-situ proteolysis

In situ-proteolysis was performed as described in (Dong et al., 2007). α -chymotrypsin (SigmaC3142) dissolved at a concentration of 1mg/ml in a solution containing 1mM HCl and 2mM CaCl₂, was added to purified protein on ice immediately prior to crystallization trials at a ratio of 0.01-1 μ g of chymotrypsin per mg of protein, dissolved at 10-15 mg/ml in 20mM Tris pH7.5 and 100mM NaCl. Crystallization trays were set up immediately as described in section 6.10.1.

5.9. Binding assays for CRM1 mutants

20 μ l IgG-sepharose 6 FastFlow (Amersham Biosciences) was washed twice with 10CV of binding buffer (50mM Tris-HCl pH7.5, 300mM NaCl), 2X with 10CV of 200mM Glycine-HCl pH2.0 followed by equilibration in binding buffer. ZZ-tagged baits (PKI-NES or RanGTP Q69L¹⁻¹⁸⁰) were mixed to saturate the beads and excess bait was washed away by 2X wash in the binding buffer. Binding mix was prepared by adding CRM1 to PKI-NES or RanGTP (1 μ M final concentration of each component in 500 μ l binding buffer) in the presence of 1X E-mix. Bait-saturated IgG beads were incubated in the binding mix for 2-3hrs after which beads were washed 3X with the binding buffer. Bound material was eluted from the column by addition of 50 μ l SDS sample buffer.

5. Material and Methods

5.10. Live export assays in HeLa nuclei

HeLa nuclei were prepared as described in (Stuven et al., 2003). High-speed *X. laevis* egg-extract was depleted for nuclear transport receptors by passing it through phenyl-sepharose column (low-substitution, Pharmacia) using 500 μ l beads per 1ml of the extract. For export experiments, 1 μ l HeLa nuclei were incubated with 7 μ l of depleted egg-extract in the presence of 1XRan-mix (Ran-regenerating system) and 4 μ M fluorescently-labeled cargo for 30 minutes at 23°C. 1 μ M wild-type CRM1 or CRM1 mutant was added to the mix and incubated for further 10 minutes. Just before imaging, 1 μ M mCherry and 2 μ g/ml DAPI were added to the nuclei. 1 μ l of the mix was placed on the wells of a 10-well multi-test slide and covered with a cover slip. Imaging was performed with a confocal laser-scanning microscope system (SP5, Leica) using 63X- numerical aperture 1.3 Plan-Apochromat glycerol objective lens (Leica).

6. Bibliography

- Aitchison, J. D., Rout, M. P., Marelli, M., Blobel, G., and Wozniak, R. W. (1995). Two novel related yeast nucleoporins Nup170p and Nup157p: complementation with the vertebrate homologue Nup155p and functional interactions with the yeast nuclear pore-membrane protein Pom152p. *J Cell Biol* *131*, 1133-1148.
- Akey, C. W., and Radermacher, M. (1993). Architecture of the *Xenopus* nuclear pore complex revealed by three-dimensional cryo-electron microscopy. *J Cell Biol* *122*, 1-19.
- Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2007). The molecular architecture of the nuclear pore complex. *Nature* *450*, 695-701.
- Allen, N. P., Patel, S. S., Huang, L., Chalkley, R. J., Burlingame, A., Lutzmann, M., Hurt, E. C., and Rexach, M. (2002). Deciphering networks of protein interactions at the nuclear pore complex. *Mol Cell Proteomics* *1*, 930-946.
- Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D. P., Arumugam, M., Bork, P., and Hurt, E. (2011). Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell* *146*, 277-289.
- Andersen, K. R., Onischenko, E., Tang, J. H., Kumar, P., Chen, J. Z., Ulrich, A., Liphardt, J. T., Weis, K., and Schwartz, T. U. (2013). Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* *2*, e00745.
- Antonin, W., Ellenberg, J., and Dultz, E. (2008). Nuclear pore complex assembly through the cell cycle: regulation and membrane organization. *FEBS Lett* *582*, 2004-2016.
- Arts, G. J., Fornerod, M., and Mattaj, I. W. (1998). Identification of a nuclear export receptor for tRNA. *Curr Biol* *8*, 305-314.
- BAHR, G. F., and 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.
- Bailer, S. M., Balduf, C., and Hurt, E. C. (2001). The Nsp1p carboxy-terminal domain is organized into functionally distinct coiled-coil regions required for assembly of nucleoporin subcomplexes and nucleocytoplasmic transport. *Mol Cell Biol* *21*, 7944-7955.
- Bailer, S. M., Balduf, C., Katahira, J., Podtelejnikov, A., Rollenhagen, C., Mann, M., Pante, N., and Hurt, E. C. (2000). Nup116p associates with the Nup82p-Nsp1p-Nup159p nucleoporin complex. *J Biol Chem* *275*, 23540-23548.
- Bailer, S. M., Berlin, W. K., Starr, C. M., and Hanover, J. A. (1995). Characterization of nuclear pore protein p62 produced using baculovirus. *Protein Expr Purif* *6*, 546-554.
- Beck, M., Förster, F., Ecke, M., Plitzko, J. M., Melchior, F., Gerisch, G., Baumeister, W., and Medalia, O. (2004). Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* *306*, 1387-1390.
- Bilokapic, S., and Schwartz, T. U. (2012). 3D ultrastructure of the nuclear pore complex. *Curr Opin Cell Biol* *24*, 86-91.
- Bischoff, F., Klebe, C., Kretschmer, J., Wittinghofer, A., and Ponstingl, H. (1994). RanGAP1 induces GTPase activity of nuclear ras-related Ran. *Proc. Natl. Acad. Sci. USA* *91*, 2587-2591.
- Bischoff, F. R., Krebber, H., Smirnova, E., Dong, W. H., and Ponstingl, H. (1995). Coactivation of RanGTPase and inhibition of GTP dissociation by Ran GTP binding protein RanBP1. *EMBO J.* *14*, 705-715.
- Bischoff, F. R., and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* *354*, 80-82.
- Boehmer, T., Jeudy, S., Berke, I. C., and Schwartz, T. U. (2008). Structural and functional studies of Nup107/Nup133 interaction and its implications for the architecture of the nuclear pore complex. *Mol Cell* *30*, 721-731.
- Brohawn, S. G., Leksa, N. C., Spear, E. D., Rajashankar, K. R., and Schwartz, T. U. (2008). Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. *Science* *322*, 1369-1373.
- Brohawn, S. G., Partridge, J. R., Whittle, J. R., and Schwartz, T. U. (2009). The nuclear pore complex has entered the atomic age. *Structure* *17*, 1156-1168.
- Brohawn, S. G., and Schwartz, T. U. (2009). A lattice model of the nuclear pore complex. *Commun Integr Biol* *2*, 205-207.

6. Bibliography

- Bross, P., Andresen, B. S., Winter, V., Krautle, F., Jensen, T. G., Nandy, A., Kolvråa, S., Ghisla, S., Bolund, L., and Gregersen, N. (1993). Co-overexpression of bacterial GroESL chaperonins partly overcomes non-productive folding and tetramer assembly of *E. coli*-expressed human medium-chain acyl-CoA dehydrogenase (MCAD) carrying the prevalent disease-causing K304E mutation. *Biochim Biophys Acta* *1182*, 264-274.
- Bukata, L., Parker, S. L., and D'Angelo, M. A. (2013). Nuclear pore complexes in the maintenance of genome integrity. *Curr Opin Cell Biol* *25*, 378-386.
- Buss, F., and Stewart, M. (1995). Macromolecular interactions in the nucleoporin p62 complex of rat nuclear pores: binding of nucleoporin p54 to the rod domain of p62. *J Cell Biol* *128*, 251-261.
- CALLAN, H. G., RANDALL, J. T., and TOMLIN, S. G. (1949). An electron microscope study of the nuclear membrane. *Nature* *163*, 280.
- Cavalier-Smith, T. (1988). Origin of the cell nucleus. *Bioessays* *9*, 72-78.
- Chi, N. C., Adam, E. J., and Adam, S. A. (1995). Sequence and characterization of cytoplasmic nuclear protein import factor p97. *J Cell Biol* *130*, 265-274.
- Chung, K. Y., Rasmussen, S. G., Liu, T., Li, S., DeVree, B. T., Chae, P. S., Calinski, D., Kobilka, B. K., Woods, V. L. J., and Sunahara, R. K. (2011). Conformational changes in the G protein Gs induced by the beta2 adrenergic receptor. *Nature* *477*, 611-615.
- Cohen, S. L., Ferre-D'Amare, A. R., Burley, S. K., and Chait, B. T. (1995). Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. *Protein Sci* *4*, 1088-1099.
- Cordes, V. C., Reidenbach, S., and Franke, W. W. (1995). High content of a nuclear pore complex protein in cytoplasmic annulate lamellae of *Xenopus* oocytes. *Eur J Cell Biol* *68*, 240-255.
- Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T., and Matunis, M. J. (2002). Proteomic analysis of the mammalian nuclear pore complex. *J Cell Biol* *158*, 915-927.
- D'Angelo, M. A., and Hetzer, M. W. (2008). Structure, dynamics and function of nuclear pore complexes. *Trends in Cell Biology* *18*, 456-466.
- D'Angelo, M. A., Anderson, D. J., Richard, E., and Hetzer, M. W. (2006). Nuclear pores form de novo from both sides of the nuclear envelope. *Science* *312*, 440-443.
- Davis, L. I., and Blobel, G. (1986). Identification and characterization of a nuclear pore complex protein. *Cell* *45*, 699-709.
- Davis, L. I., and Blobel, G. (1987). Nuclear pore complex contains a family of glycoproteins that includes p62: glycosylation through a previously unidentified cellular pathway. *Proc Natl Acad Sci USA* *84*, 7552-7556.
- Debler, E. W., Ma, Y., Seo, H. S., Hsia, K. C., Noriega, T. R., Blobel, G., and Hoelz, A. (2008). A fence-like coat for the nuclear pore membrane. *Mol Cell* *32*, 815-826.
- Denning, D. P., Patel, S. S., Uversky, V., Fink, A. L., and Rexach, M. (2003). Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc Natl Acad Sci USA* *100*, 2450-2455.
- Devos, D., Dokudovskaya, S., Williams, R., Alber, F., Eswar, N., Chait, B. T., Rout, M. P., and Sali, A. (2006). Simple fold composition and modular architecture of the nuclear pore complex. *Proc Natl Acad Sci USA* *103*, 2172-2177.
- Dolker, N., Blanchet, C. E., Voss, B., Haselbach, D., Kappel, C., Monecke, T., Svergun, D. I., Stark, H., Ficner, R., Zachariae, U., Grubmüller, H., and Dickmanns, A. (2013). Structural Determinants and Mechanism of Mammalian CRM1 Allosterity. *Structure* *21*, 1350-1360.
- Dong, A., Xu, X., Edwards, A. M., Chang, C., Chruszcz, M., Cuff, M., Cymborowski, M., Di Leo, R., Egorova, O., Evdokimova, E., Filippova, E., Gu, J., Guthrie, J., Ignatchenko, A., Joachimiak, A., Klostermann, N., Kim, Y., Korniyenko, Y., Minor, W., Que, Q., Savchenko, A., Skarina, T., Tan, K., Yakunin, A., Yee, A., Yim, V., Zhang, R., Zheng, H., Akutsu, M., Arrowsmith, C., Avvakumov, G. V., Bochkarev, A., Dahlgren, L. G., Dhe-Paganon, S., Dimov, S., Dombrowski, L., Finerty, P. J., Flodin, S., Flores, A., Graslund, S., Hammerstrom, M., Herman, M. D., Hong, B. S., Hui, R., Johansson, I., Liu, Y., Nilsson, M., Nedyalkova, L., Nordlund, P., Nyman, T., Min, J., Ouyang, H., Park, H. W., Qi, C., Rabeh, W., Shen, L., Shen, Y., Sukumard, D., Tempel, W., Tong, Y., Tresagues, L., Vedadi, M., Walker, J. R., Weigelt, J., Welin, M., Wu, H., Xiao, T., Zeng, H., and Zhu, H. (2007). In situ proteolysis for protein crystallization and structure determination. *Nat Methods* *4*, 1019-1021.
- Doye, V., and Hurt, E. C. (1995). Genetic approaches to nuclear pore structure and function. *Trends Genet* *11*, 235-241.

6. Bibliography

- Dultz, E., and Ellenberg, J. (2010). Live imaging of single nuclear pores reveals unique assembly kinetics and mechanism in interphase. *J Cell Biol* *191*, 15-22.
- Dultz, E., Zanin, E., Wurzenberger, C., Braun, M., Rabut, G., Sironi, L., and Ellenberg, J. (2008). Systematic kinetic analysis of mitotic dis- and reassembly of the nuclear pore in living cells. *J Cell Biol* *180*, 857-865.
- Eisele, N. B., Frey, S., Piehler, J., Görlich, D., and Richter, R. P. (2010). Ultrathin nucleoporin phenylalanine-glycine repeat films and their interaction with nuclear transport receptors. *EMBO Rep* *11*, 366-372.
- Feldherr, C. (1962). The nuclear annuli as pathways for nucleocytoplasmic exchanges. *J. Cell Biol.* *14*, 65-72.
- Finlay, D. R., Meier, E., Bradley, P., Horecka, J., and Forbes, D. J. (1991). A complex of nuclear pore proteins required for pore function. *J Cell Biol* *114*, 169-183.
- Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I. W. (1997a). Crm1 is an export receptor for leucine rich nuclear export signals. *Cell* *90*, 1051-1060.
- Fornerod, M., van Deursen, J., van Baal, S., Reynolds, A., Davis, D., Murti, K. G., Franssen, J., and Grosveld, G. (1997b). The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88. *EMBO J* *16*, 807-816.
- Frenkiel-Krispin, D., Maco, B., Aebi, U., and Medalia, O. (2010). Structural analysis of a metazoan nuclear pore complex reveals a fused concentric ring architecture. *J Mol Biol* *395*, 578-586.
- Frey, S., and Görlich, D. (2007). A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* *130*, 512-523.
- Frey, S., Richter, R. P., and Görlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* *314*, 815-817.
- Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M., and Nishida, E. (1997). CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* *390*, 308-311.
- GALL, J. G. (1954). Observations on the nuclear membrane with the electron microscope. *Exp Cell Res* *7*, 197-200.
- Gall, J. G. (1967). Octagonal nuclear pores. *J Cell Biol* *32*, 391-399.
- Goldberg, M. W., and Allen, T. D. (1993). The nuclear pore complex: three-dimensional surface structure revealed by field emission, in-lens scanning electron microscopy, with underlying structure uncovered by proteolysis. *J Cell Sci* *106*, 261-274.
- Goldberg, M. W., and Allen, T. D. (1995). Structural and functional organization of the nuclear envelope. *Curr Opin Cell Biol* *7*, 301-309.
- Görlich, D., and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* *15*, 607-660.
- Görlich, D., Pante, N., Kutay, U., Aebi, U., and Bischoff, F. R. (1996). Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J* *15*, 5584-5594.
- Görlich, D., Prehn, S., Laskey, R. A., and Hartmann, E. (1994). Isolation of a protein that is essential for the first step of nuclear protein import. *Cell* *79*, 767-778.
- Görlich, D., Vogel, F., Mills, A. D., Hartmann, E., and Laskey, R. A. (1995). Distinct functions for the two importin subunits in nuclear protein import. *Nature* *377*, 246-248.
- Grandi, P., Dang, T., Pane, N., Shevchenko, A., Mann, M., Forbes, D. J., and Hurt, E. C. (1997). Nup93, a vertebrate homologue of yeast Nic96p, forms a complex with a novel 205-kDa protein and is required for correct nuclear pore assembly. *Mol Biol Cell* *8*, 2017-2038.
- Grandi, P., Doye, V., and Hurt, E. C. (1993). Purification of NSP1 reveals complex formation with 'GLFG' nucleoporins and a novel nuclear pore protein NIC96. *EMBO J* *12*, 3061-3071.
- Grandi, P., Emig, S., Weise, C., Hucho, F., Pohl, T., and Hurt, E. C. (1995a). A novel nuclear pore protein Nup82p which specifically binds to a fraction of Nsp1p. *J Cell Biol* *130*, 1263-1273.
- Grandi, P., Schlaich, N., Tekotte, H., and Hurt, E. C. (1995b). Functional interaction of Nic96p with a core nucleoporin complex consisting of Nsp1p, Nup49p and a novel protein Nup57p. *EMBO J* *14*, 76-87.
- Grossman, E., Medalia, O., and Zwerger, M. (2012). Functional architecture of the nuclear pore complex. *Annu Rev Biophys* *41*, 557-584.
- Grote, M., Kubitschek, U., Reichelt, R., and Peters, R. (1995). Mapping of nucleoporins to the center of the nuclear pore complex by post-embedding immunogold electron microscopy. *J Cell Sci* *108*, 2963-2972.

6. Bibliography

- Guan, T., Muller, S., Klier, G., Pante, N., Blevitt, J. M., Haner, M., Paschal, B., Aebi, U., and Gerace, L. (1995). Structural analysis of the p62 complex, an assembly of O-linked glycoproteins that localizes near the central gated channel of the nuclear pore complex. *Mol Biol Cell* *6*, 1591-1603.
- Güttler, T., and Görlich, D. (2011). Ran-dependent nuclear export mediators: a structural perspective. *EMBO J* *30*, 3457-3474.
- Güttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., and Görlich, D. (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* *17*, 1367-1376.
- Heras, B., and Martin, J. L. (2005). Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr* *61*, 1173-1180.
- Hetzer, M. W., and Wenthe, S. R. (2009). Border control at the nucleus: biogenesis and organization of the nuclear membrane and pore complexes. *Dev Cell* *17*, 606-616.
- Hoelz, A., Debler, E. W., and Blobel, G. (2011). The structure of the nuclear pore complex. *Annu Rev Biochem* *80*, 613-643.
- Hulsmann, B. B., Labokha, A. A., and Görlich, D. (2012). The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* *150*, 738-751.
- Hutten, S., and Kehlenbach, R. H. (2006). Nup214 is required for CRM1-dependent nuclear protein export in vivo. *Mol Cell Biol* *26*, 6772-6785.
- Imamoto, N., and Funakoshi, T. (2012). Nuclear pore dynamics during the cell cycle. *Curr Opin Cell Biol* *24*, 453-459.
- Iovine, M. K., Watkins, J. L., and Wenthe, S. R. (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., von Kobbe, C., Mattaj, I. W., and Görlich, D. (1997). The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J* *16*, 6535-6547.
- Kampmann, M., Atkinson, C. E., Mattheyses, A. L., and Simon, S. M. (2011). Mapping the orientation of nuclear pore proteins in living cells with polarized fluorescence microscopy. *Nat Struct Mol Biol* *18*, 643-649.
- Kehlenbach, R. H., Dickmanns, A., and Gerace, L. (1998). Nucleocytoplasmic shuttling factors including Ran and CRM1 mediate nuclear export of NFAT In vitro. *J Cell Biol* *141*, 863-874.
- Kita, K., Omata, S., and Horigome, T. (1993). Purification and characterization of a nuclear pore glycoprotein complex containing p62. *J Biochem (Tokyo)* *113*, 377-382.
- Koyama, M., and Matsuura, Y. (2010). An allosteric mechanism to displace nuclear export cargo from CRM1 and RanGTP by RanBP1. *EMBO J* *29*, 2002-2013.
- Kutay, U., and Guttinger, S. (2005). Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol* *15*, 121-124.
- Kutay, U., Bischoff, F. R., Kostka, S., Kraft, R., and Görlich, D. (1997). Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* *90*, 1061-1071.
- Labokha, A. A., Gradmann, S., Frey, S., Hulsmann, B. B., Urlaub, H., Baldus, M., and Görlich, D. (2013). Systematic analysis of barrier-forming FG hydrogels from *Xenopus* nuclear pore complexes. *EMBO J* *32*, 204-218.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* *227*, 680-685.
- Lam, A. Y., Pardon, E., Korotkov, K. V., Hol, W. G., and Steyaert, J. (2009). Nanobody-aided structure determination of the EpsI:EpsJ pseudopilin heterodimer from *Vibrio vulnificus*. *J Struct Biol* *166*, 8-15.
- Laurell, E., Beck, K., Krupina, K., Theerthagiri, G., Bodenmiller, B., Horvath, P., Aebersold, R., Antonin, W., and Kutay, U. (2011). Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. *Cell* *144*, 539-550.
- Li, C., Schwabe, J. W., Banayo, E., and Evans, R. M. (1997). Coexpression of nuclear receptor partners increases their solubility and biological activities. *Proc Natl Acad Sci U S A* *94*, 2278-2283.
- Lim, R. Y., Fahrenkrog, B., Koser, J., Schwarz-Herion, K., Deng, J., and Aebi, U. (2007). Nanomechanical basis of selective gating by the nuclear pore complex. *Science* *318*, 640-643.
- Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012). SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb Cell Fact* *11*, 56.
- Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* *252*, 1162-1164.

6. Bibliography

- Macaulay, C., Meier, E., and Forbes, D. J. (1995). Differential mitotic phosphorylation of proteins of the nuclear pore complex. *J Biol Chem* *270*, 254-262.
- Maimon, T., Elad, N., Dahan, I., and Medalia, O. (2012). The human nuclear pore complex as revealed by cryo-electron tomography. *Structure* *20*, 998-1006.
- Marelli, M., Aitchison, J. D., and Wozniak, R. W. (1998). Specific binding of the karyopherin Kap121p to a subunit of the nuclear pore complex containing Nup53p, Nup59p, and Nup170p. *J Cell Biol* *143*, 1813-1830.
- Matsuoka, Y., Takagi, M., Ban, T., Miyazaki, M., Yamamoto, T., Kondo, Y., and Yoneda, Y. (1999). Identification and characterization of nuclear pore subcomplexes in mitotic extract of human somatic cells. *Biochem Biophys Res Commun* *254*, 417-423.
- Melcak, I., Hoelz, A., and Blobel, G. (2007). Structure of Nup58/45 suggests flexible nuclear pore diameter by intermolecular sliding. *Science* *315*, 1729-1732.
- Melchior, F., Paschal, B., Evans, E., and Gerace, L. (1993). Inhibition of nuclear protein import by nonhydrolyzable analogs of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. *J. Cell Biol.* *123*, 1649-1659.
- Miller, M. W., Caracciolo, M. R., Berlin, W. K., and Hanover, J. A. (1999). Phosphorylation and glycosylation of nucleoporins. *Arch Biochem Biophys* *367*, 51-60.
- Mohr, D., Frey, S., Fischer, T., Güttler, T., and Görlich, D. (2009). Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J* *28*, 2541-2553.
- Monecke, T., Güttler, T., Neumann, P., Dickmanns, A., Görlich, D., and Ficner, R. (2009). Crystal Structure of the Nuclear Export Receptor CRM1 in Complex with Snurportin1 and RanGTP. *Science* *324*, 1087-1091.
- Monecke, T., Haselbach, D., Voss, B., Russek, A., Neumann, P., Thomson, E., Hurt, E., Zachariae, U., Stark, H., Grubmuller, H., Dickmanns, A., and Ficner, R. (2013). Structural basis for cooperativity of CRM1 export complex formation. *Proc Natl Acad Sci U S A* *110*, 960-965.
- Moore, M., and Blobel, G. (1993). The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature* *365*, 661-663.
- Moore, M. S., and Blobel, G. (1992). The two steps of nuclear import, targeting to the nuclear envelope and translocation through the nuclear pore, require different cytosolic factors. *Cell* *69*, 939-950.
- Moore, M. S., and Blobel, G. (1994). Purification of a Ran-interacting protein that is required for protein import into the nucleus. *Proc. Natl. Acad. Sci. USA* *91*, 10212-10216.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1992). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986. *Biotechnology* *24*, 17-27.
- Napetschnig, J., Blobel, G., and Hoelz, A. (2007). Crystal structure of the N-terminal domain of the human protooncogene Nup214/CAN. *Proc Natl Acad Sci U S A* *104*, 1783-1788.
- Nehrbass, U., Kern, H., Mutvei, A., Horstmann, H., Marshallsay, B., and Hurt, E. C. (1990). NSP1: a yeast nuclear envelope protein localized at the nuclear pores exerts its essential function by its carboxy-terminal domain. *Cell* *61*, 979-989.
- Ohtsubo, M., Okazaki, H., and 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.
- Ori, A., Banterle, N., Iskar, M., Andres-Pons, A., Escher, C., Khanh Bui, H., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., Lemke, E. A., and Beck, M. (2013). Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol Syst Biol* *9*, 648.
- Pante, N., Bastos, R., McMorro, I., Burke, B., and Aebi, U. (1994). Interactions and three-dimensional localization of a group of nuclear pore complex proteins. *J Cell Biol* *126*, 603-617.
- Rabut, G., Doye, V., and Ellenberg, J. (2004). Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nat Cell Biol* *6*, 1114-1121.
- Radu, A., Moore, M. S., and Blobel, G. (1995). The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex. *Cell* *81*, 215-222.
- Reichelt, R., Holzenburg, A., Buhle, E. L., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J Cell Biol* *110*, 883-894.
- Rexach, M., and Blobel, G. (1995). Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell* *83*, 683-692.
- Ribbeck, K., and Görlich, D. (2001). Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* *20*, 1320-1330.

6. Bibliography

- Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M., and Görlich, D. (1998). NTF2 mediates nuclear import of Ran. *EMBO J* *17*, 6587-6598.
- Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol* *148*, 635-651.
- Rout, M. P., and Blobel, G. (1993). Isolation of the yeast nuclear pore complex. *J Cell Biol* *123*, 771-783.
- Saito, N., and Matsuura, Y. (2013). A 2.1-Å-resolution crystal structure of unliganded CRM1 reveals the mechanism of autoinhibition. *J Mol Biol* *425*, 350-364.
- Sambrook, J., Fritsch, E., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*
- Sampathkumar, P., Kim, S. J., Upla, P., Rice, W. J., Phillips, J., Timney, B. L., Pieper, U., Bonanno, J. B., Fernandez-Martinez, J., Hakhverdyan, Z., Ketaren, N. E., Matsui, T., Weiss, T. M., Stokes, D. L., Sauder, J. M., Burley, S. K., Sali, A., Rout, M. P., and Almo, S. C. (2013). Structure, dynamics, evolution, and function of a major scaffold component in the nuclear pore complex. *Structure* *21*, 560-571.
- Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M. C., Agard, D. A., Gustafsson, M. G., Leonhardt, H., and Sedat, J. W. (2008). Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* *320*, 1332-1336.
- Schlaich, N. L., Haner, M., Lustig, A., Aebi, U., and Hurt, E. C. (1997). In vitro reconstitution of a heterotrimeric nucleoporin complex consisting of recombinant Nsp1p, Nup49p, and Nup57p. *Mol Biol Cell* *8*, 33-46.
- Schrader, N., Koerner, C., Koessmeier, K., Bangert, J. A., Wittinghofer, A., Stoll, R., and Vetter, I. R. (2008a). The crystal structure of the Ran-Nup153ZnF2 complex: a general Ran docking site at the nuclear pore complex. *Structure* *16*, 1116-1125.
- Schrader, N., Stelter, P., Flemming, D., Kunze, R., Hurt, E., and Vetter, I. R. (2008b). Structural basis of the Nic96 subcomplex organization in the nuclear pore channel. *Mol Cell* *29*, 46-55.
- Siniosoglou, S., Lutzmann, M., Santos-Rosa, H., Leonard, K., Mueller, S., Aebi, U., and Hurt, E. C. (2000). Structure and assembly of the Nup84p complex. *J Cell Biol* *149*, 41-54.
- Solmaz, S. R., Blobel, G., and Melcak, I. (2013). Ring cycle for dilating and constricting the nuclear pore. *Proc Natl Acad Sci U S A* *110*, 5858-5863.
- Solmaz, S. R., Chauhan, R., Blobel, G., and Melcak, I. (2011). Molecular architecture of the transport channel of the nuclear pore complex. *Cell* *147*, 590-602.
- Stade, K., Ford, C. S., Guthrie, C., and Weis, K. (1997). Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* *90*, 1041-1050.
- Stavru, F., Hülsmann, B. B., Spang, A., Hartmann, E., Cordes, V. C., and Görlich, D. (2006a). NDC1: a crucial membrane-integral nucleoporin of metazoan nuclear pore complexes. *J Cell Biol* *173*, 509-519.
- Stavru, F., Nautrup-Pedersen, G., Cordes, V. C., and Görlich, D. (2006b). Nuclear pore complex assembly and maintenance in POM121- and gp210-deficient cells. *J Cell Biol* *173*, 477-483.
- Stuken, T., Hartmann, E., and Görlich, D. (2003). Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. *EMBO J* *22*, 5928-5940.
- Szymborska, A., de Marco, A., Daigle, N., Cordes, V. C., Briggs, J. A., and Ellenberg, J. (2013). Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging. *Science*
- Walther, T. C., Alves, A., Pickersgill, H., Loiodice, I., Hetzer, M., Galy, V., Hülsmann, B. B., Kocher, T., Wilm, M., Allen, T., Mattaj, I. W., and Doye, V. (2003). The conserved Nup107-160 complex is critical for nuclear pore complex assembly. *Cell* *113*, 195-206.
- Watson, M. L. (1954). Pores in the mammalian nuclear membrane. *Biochim Biophys Acta* *15*, 475-479.
- Wolf, E., Kim, P. S., and Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci* *6*, 1179-1189.
- Wu, J., Matunis, M. J., Kraemer, D., Blobel, G., and Coutavas, E. (1995). Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin a homologous domain, and a leucine-rich region. *J Biol Chem* *270*, 14209-14213.
- Yang, Q., Rout, M. P., and Akey, C. W. (1998). Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Mol Cell* *1*, 223-234.

7. Abbreviations

Å	Angstrom
A ₂₈₀	absorbance at λ 280nm
AL	acidic loop
ATP	adenosine 5'-triphosphate
br	<i>Brachypodium</i>
C-	carboxy terminal
cAMP	3', 5'-cyclic adenosine monophosphate
CAS	cellular apoptosis susceptibility (Exportin 2)
CC	coiled-coil
COPII	coat protein complex II
CRM1	chromosomal region maintenance 1 (Exportin 1/Xpo1p)
Cryo-ET	cryo-electron tomography
ct	<i>Chaetomium thermophilum</i>
DDT	dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
EM	electron microscopy
ER	endoplasmic reticulum
FGrepeat	phenylalanine-glycine repeat
ΔFG	deletion FG repeat
GDP	guanosine 5'-diphosphate
GFP	green fluorescent protein
GTP	guanine nucleotide-binding protein
GTPase	GTP hydrolase
HC	hydrophobic cleft
HEAT repeat	class of protein repeats (h untingtin, e longation factor3, the PR65/A subunit of protein phosphatase 2A (PP2A), lipid kinase TOR)
HIV	human immunodeficiency virus
hnRNP	heterogeneous nuclear ribonucleoprotein
HR	heptad repeat
INM	inner nuclear membrane
K _d	dissociation constant
kDa	kilodalton
LS	light scattering
MALS	multi angle light scattering
MDa	megadalton
mm	<i>Mus musculus</i>
MW	molecular weight
N-	amino terminal
Nb	nanobody
NE	nuclear envelope
NEDD8	n eural precursor cell expressed, d evelopmentally d ownregulated 8
NES	nuclear export signal
NLS	nuclear localization signal
nm	nanometer
NMR	nuclear magnetic resonance
NPC	nuclear pore complex
NTF2	n uclear t ransport f actor 2
NTR	nuclear transport receptor
Nup	nucleoporin
Nup93n	Nup93 N-terminal region 1-95
Nup54α/β	Nup54 alpha-beta region 146-372

OD	optical density
ONM	outer nuclear membrane
PKI	protein kinase inhibitor
QELS	quasi elastic light scattering
Ran	Ras -related nuclear antigen
RanBP	Ran-binding protein
RanGAP	RanGTPase activating protein
RCC1	regulator of chromosomal condensation 1
Rh	hydrodynamic radius
Rms	root mean square radius
rn	<i>Rattus norvegicus</i>
RNP	ribonucleoprotein
S	sedimentation coefficient
sc	<i>Saccharomyces cerevisiae</i>
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SPN1	Snurportin1
SPR	surface plasmon resonance
SUMO	small ubiquitin like modifier
TEV	tobacco etch virus
TM	transmembrane
μM	micromolar
<i>X.laevis</i>	<i>Xenopus laevis</i>
Y-complex	Nup107-160 nucleoporin complex (scNup84 complex)
Z (in ZZ-tag)	IgG-binding domain of <i>Staphylococcal</i> protein A

8. Acknowledgements

There are many people whom I would like to thank for their help and support.

Firstly, I feel extremely grateful and indebted to my PhD supervisor, Prof Dirk Görlich for his valuable guidance and constant encouragement during this project. His brilliant ideas have fostered my analytical thinking and his enthusiasm for research has been contagious. I also would like to thank him for being so generous with his time and efforts in this project that often inspired me to push further.

I would also like to thank the members of my PhD thesis committee, Prof Reinhard Lührmann and Prof Peter Rehling, for the stimulating discussions during the committee meetings and for their ever-encouraging and constructive feedback.

I wish to express my deepest gratitude for the IMPRS molecular biology program for selecting me in this prestigious program and providing me the opportunity to explore Germany for my PhD studies. I also wish to thank the program coordinators Dr. Steffen Burckhardt and Kerstin Grueniger for the excellent support they have provided during the last 5 years.

This work would not have been possible without the assistance of the central X-ray crystallization facility of the Max Planck institute headed by Dr. Vlad Pena. I especially wish to thank Ulrich Steuerwald for his valuable assistance and for teaching me the technical aspects of protein crystallization.

I would like to acknowledge Dr. Bastian Hülsmann for his initial studies on Nup62 complex that formed the basis of this work. A special thanks goes to Dr. Thomas Güttler and Kevser Gencalp for their initial work on CRM1 that laid the foundation of the second part of this project.

I am also deeply thankful to many people that provided great technical assistance during this project. I would especially like to acknowledge the efforts of Juergen Schunnemann for his help in protein purification, Renate Rees and Sussane Brandfass for their assistance in the cloning work. Sincere thanks goes to Gabrielle Kopp for preparing huge amounts of plasmid DNAs, SDS gels, competent cells and making life so much easier.

I am very grateful to Dr. Steffen Frey for sharing his extensive knowledge as well as scientific tools ranging from plasmid constructs to purified enzymes, all so generously.

A big 'thanks' to Heinz Juergen Dehne for his friendly presence in the lab and a happy influence. Another big thanks to Cornelia Paz for always being so kind in her help while dealing with the paper work and more so for her Indian cooking.

My heartfelt 'thanks' to all the members of department 10600 for being so nice and friendly and creating a positive and productive work environment. Aksana, Bastian, Bernard, Broder, Jenny, Kevser, Koray, Matthias, Metin, Michael, Tino, Steffen and all the others, Thank you so much!!

Finally, I wish to thank a few but very dear friends, Jatin, Kundan, Srinivas and Swathi, who made my years in Germany a lot more fun. I am in absolute awe of Rameez for his wonderful company and so many lovely memories.

Words are not enough to describe how blessed I feel for the unconditional love and support I receive from my family, my father, my siblings- Ginny and Deepu and my grandmother. You all have been the pillars of strength for me. I love you all so dearly.

Finally, I would like to dedicate this work to my beloved mother whose blessings made all this possible. I know you must be proud.

9. Curriculum vitae

Name	Hema Chug
Date of birth	March 18, 1986
Place of birth	New Delhi, India
Nationality	Indian

Education

2003-2006	B.Sc. (Hons.) Microbiology University of Delhi, New Delhi, India
2006-2008	M.Sc. Biotechnology Jawaharlal Nehru University, New Delhi, India
2008-2009	Intensive course program MSc/PhD in Molecular Biology International Max Planck Research School (IMPRS), Göttingen, Germany
2009-present	Doctoral studies in the laboratory of Prof. Dirk Görlich Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; Göttingen Graduate School for Neurosciences, Biophysics and Molecular biology (GGNB) program “IMPRS-Molecular Biology”

Research

2009-present	Doctoral thesis project “Biochemical characterization of the Nup62-58-54 nucleoporin complex and mutational analysis of exportin CRM1”
2007-2008	Master’s thesis project “Identification and characterization of immunodominant B-cell epitopes of the C-terminus of protective antigen of <i>Bacillus anthracis</i> ”

Publications

Kaur, M., **Chug, H.**, Singh, H., Chandra, S., Mishra, M., Sharma, M., and Bhatnagar, R. (2009). Identification and characterization of immunodominant B-cell epitope of the C-terminus of protective antigen of *Bacillus anthracis*. *Mol Immunol* 46, 2107-2115.