2009

Architecture of a Coat for the Nuclear Pore Membrane

Kuo-Chiang Hsia

ARCHITECTURE OF A COAT FOR THE NUCLEAR PORE MEMBRANE

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Kuo-Chiang Hsia

June 2009

ARCHITECTURE OF COAT FOR the NUCLEAR PORE MEMBRANE

Kuo-Chiang Hsia, Ph.D.

The Rockefeller University 2009

Nuclear pore complexes (NPCs) reside in the nuclear membrane and mediate macromolecules exchange between the nucleus and cytoplasm. Although the protein components of NPCs, termed collectively nucleoporin, have been identified, how nucleoporins arrange in NPC, however, is still an enigma. NPC is a highly symmetric protein complex, which contains an eight-fold rotational symmetry across the center of the pore and a two-fold symmetry in the plane of the nuclear envelope. In addition, according to electron microscopic reconstruction model, NPC can also be considered schematically as a series of concentric cylinders. A peripheral cylinder coating the nuclear pore membrane contains a heptameric Nup84 complex, which harbors Nup145C/Sec13 complex in the center. Hence, X-ray crystallographic approach was employed to determine the Nup145C/Sec13 structure. Notably, in the two crystal forms I obtained, Nup145C/Sec13 oligmerized to form a slightly bent hetero-octamer rod. In the meantime, the

oligomerization state could be observed in solution as well. Furthermore, as indicated by structural analyses, the interacting interfaces not only are evolutionarily conserved but also cover huge area. Taken together, the Nup145C/Sec13 hetero-octamer rod could be physiologically relevant without much doubt. In concordance with the stoichiometric considerations of NPC, then we confidently propose a model with respect to a coat for the nuclear pore membrane. In the model, the coat is constructed of four rings and each ring is composed of eight heptamers. Moreover, due to the three axes on the Nup145C/Sec13 rod, the hetero-octamer is also places vertically in order to connect these four anti-parallel rings. Remarkably, the Nup85/Seh1p crystal structures that were determined recently indicated that the assembly of this pair bears reseblance in its shape and dimension to that of another pair, Nup145C/Sec13. Thus, Nup85/Seh1 does form another hetero-octamer pole as we predicted in the model. Surprisingly, in contrast to the rigidity of the Nup145C/Sec13 hetero-octamer, the Nup85/Seh1 structures reveal a hinge motion that may facilitate the conformational change in the NPC for importing the integral membrane proteins.

I DEDICATE THIS THESIS TO MY DAUGHTER AND SON, JUDY AND JOSHUA, TO MY WIFE, SUYI, AND TO MY PARENTS FOR THE UNDERSTANDING AND ENCOURAGEMENT THEY PROVIDED DURING ALL THESE YEARS OF STUDY

ACKNOWLEDGEMENTS

Four and half year graduate student life is like a trek and, on the trek, I am fortunate to have come in touch with many people who directly or indirectly affected the course of the trek. Hence, I would like to express my deepest thanks and gratitude.

I certainly want to thank my advisor, Günter Blobel, for the valuable guidance, and constructive advice throughout the entire project. I also want to thank André Hoelz for the precious assistance and inspirations in experimental designing. Special thanks goes to the members of Blobel lab for their continued support and friendship, and in particular Elias Coutavas, Erik Debler, and Peter Stavropoulos for their critical reading and helpful discussions regarding my thesis writing; Vivien Nagy, Johanna Napetschnig, Peter Stavropoulos, Daniel Wacker, and Kimihisa Yoshida for their comments on my thesis speech; Richard Wong for the assistance of cell biology experiments. And also special thanks Lourdes Quirolgico and Tita Isberto for their administrative support and any help requested.

I am also appreciative to my committee members, Tarun Kapoor and Seth Darst, for their comments and suggestions

during my graduate study, as well as Jonathan Goldberg for taking the time to act as the external member of my thesis committee. At last, I would like to acknowledge all of my teachers and friends in my country, Taiwan, for their encouragement and opinions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Eukaryotic cells, unlike prokaryotic cells, contain about a dozen different organelles and one of which is the nucleus that possesses genomic DNA, the inheritable material of life. The nucleus is surrounded by a nuclear envelope, which serves as a barrier between the nuclear and the cytoplasmic compartments. However, macromolecules, such as pre-ribosomal sub-units and histones, need to be located in the right compartment in order to execute their functions. The only gateway that allows molecules to go through the nuclear envelope is the nuclear pore complex (NPC). The transport process of macromolecules through the NPC is highly regulated.

1.1 Selective transport of macromolecules

Molecules smaller than ~40 kDa in size are able to travel through the NPC in a diffusion-controlled manner. However, molecules larger than 40 kDa require specific transport receptors for transport (Hoelz and Blobel, 2004; Keminer and Peters, 1999; Pante and Kann, 2002). Furthermore, in order to control the transport of various sizes of cargo in and out of the nucleus, two groups of proteins are involved

in tightly regulating nuclear transport: mobile transport factors and the stable nuclear pore complex.

1.1.1 Characteristics of karyopherins and Ran

Cargo which contain either nuclear localization signals (NLS) or nuclear export signals (NES) are specifically recognized by a family of soluble transport factors called importins, exportins and transportins, collectively referred to as karyopherins (kaps) in yeast. The karyopherin family comprises at least 14 members in yeast and 20 members in humans (Cook et al., 2007), and all kaps are mainly composed of α-solenoid folds, consisting of multiple HEAT repeats, a helix–loop–helix motif.

The directionality of nucleocytoplasmic transport is governed by the small G protein Ran, which has two nucleotide-bound states, guanosine triphosphate (GTP) or guanosine diphosphate (GDP). The different nucleotide-bound states of Ran have significant conformational differences, altering the affinity to their binding partners and, eventually, triggering the cargo transport cycle. For instance, the affinity of import kap-β2 for RanGTP is about 10^4-fold higher than for RanGDP (Richards et al., 1995). Hence, RanGTP, which is abundant in the nucleus, interacts

with a cargo/kap-β2 complex in the nucleus and further triggers the release of cargo.

Compared with RanGDP, the RanGppNHp structure, the structure of Ran in complex with a non-hydrolyzable analog of GTP, shows large structural rearrangements in the nucleotide switch regions, switch I, II, and the C-terminal extension. The conformational changes result in different binding affinities (Chook and Blobel, 1999; Vetter et al., 1999) (Fig 1). Upon binding to both the C- and N-terminal arches of the import kap-β2, RanGTP locks the import kap-β2 into a conformation that is incapable of cargo binding (Lee et al., 2005).

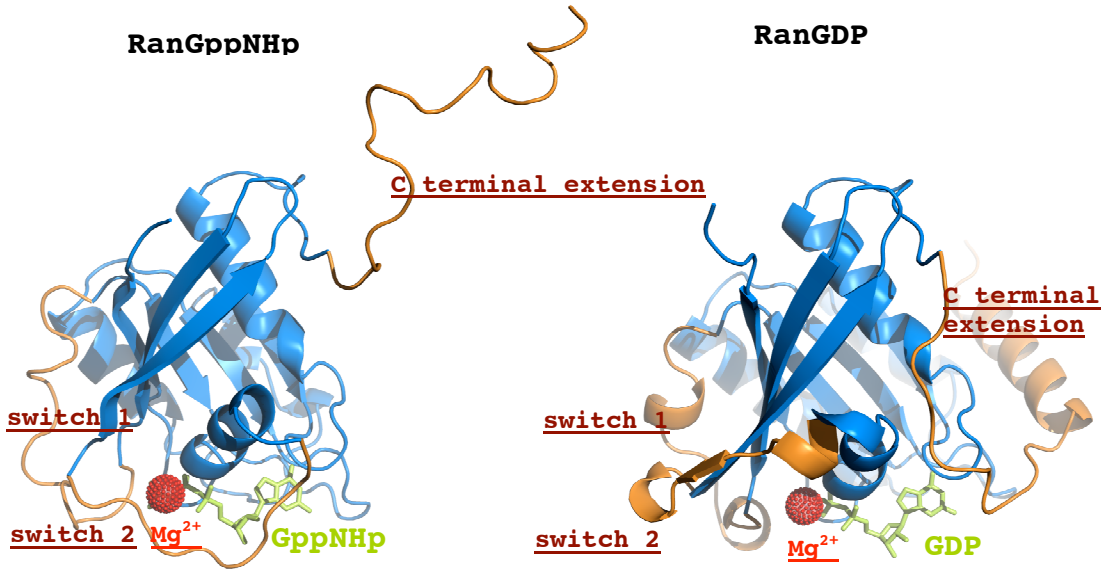


Figure 1. Structural differences between RanGppNHp and RanGDP

The regions that differ between RanGppNHp (Protein data bank (PDB) code 1qbk, Chook and Blobel, 1999) and RanGDP (PDB code 1byu, Stewart et al., 1998) are shown in brown. These regions include switch 1, switch 2, and the C-terminal extension. A magnesium ion is shown in red, and GppNHp and GDP are shown light green. Image adapted from Chook et al., 1999.

1.1.2 Mechanism of cargo transport

RanGTP is present in high concentration in the nucleus. Upon import of cargo from the cytoplasm, RanGTP interacts with the kap/cargo complex, which triggers the release of cargo and recycling of the RanGTP/import kap complexes back into the cytoplasm. Similarly, export kap /cargo complexes in the nucleus interact with RanGTP and are then ferried to the cytoplasm. Once RanGTP was hydrolyzed to RanGDP by RanGAP (GTPase-activating protein) in the cytoplasm, cargo is released from the exporting complex, facilitated by complex formation with Ran-binding protein 1 (RanBP1). The RanGDP produced from this process returns to the nucleus by binding to the nuclear transport factor, NTF2. Once in the nucleus, RanGDP interacts with a RanGEF (guanosine nucleotide exchange factor), which replaces the GDP with GTP, resulting again in RanGTP, and beginning the cycle anew (Fig 2).

Altogether, through an asymmetric distribution of RanGTP/GDP in the cell established by Ran GAP and GEF in the cytoplasm and nucleus, respectively, different kaps carry different cargoes in and out of the nucleus (Fig 3)(Hoelz et al., 2004; Cook et al., 2007). Although our understanding of the mobile transport machinery is

extensive, the structural and functional understanding of the NPC remains elusive.

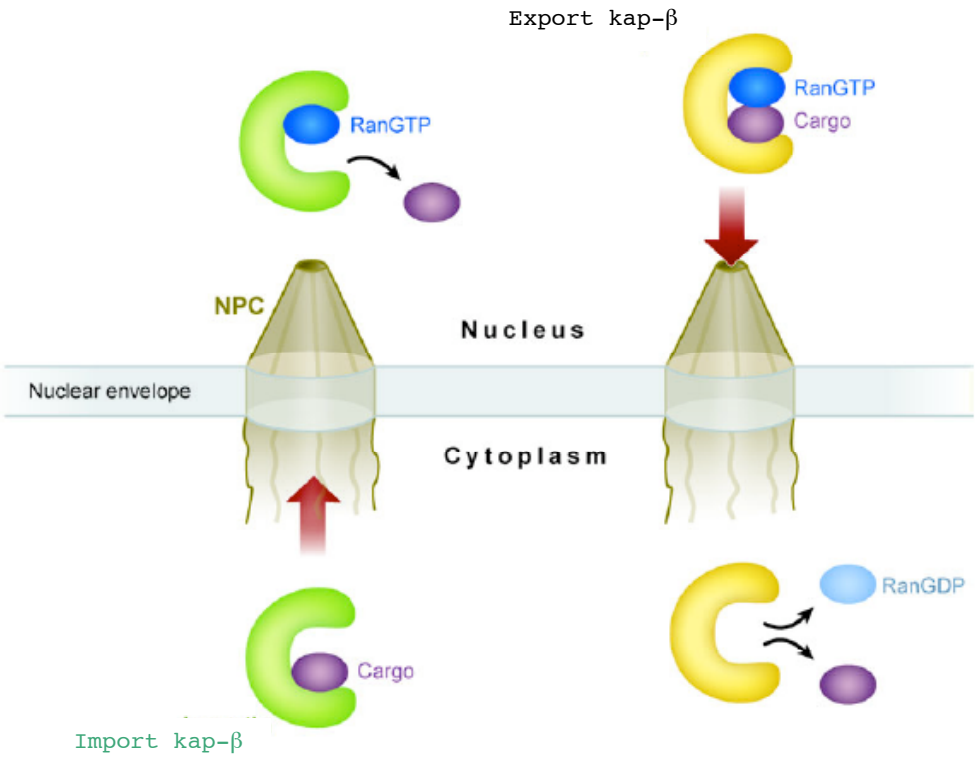


Figure 2. The nuclear transport cycle

The two nuclear pore complexes (NPCs) show the characteristic nuclear basket and cytoplasmic filaments. An import kap-β binds cargo in the cytoplasm and releases it upon binding of RanGTP in the nucleus. An export kap-β binds both cargo and RanGTP in the nucleus and releases both of them once Ran is hydrolyzed its bound GTP to GDP in the presence of RanGAP. Image taken from Cook et al., 2007.

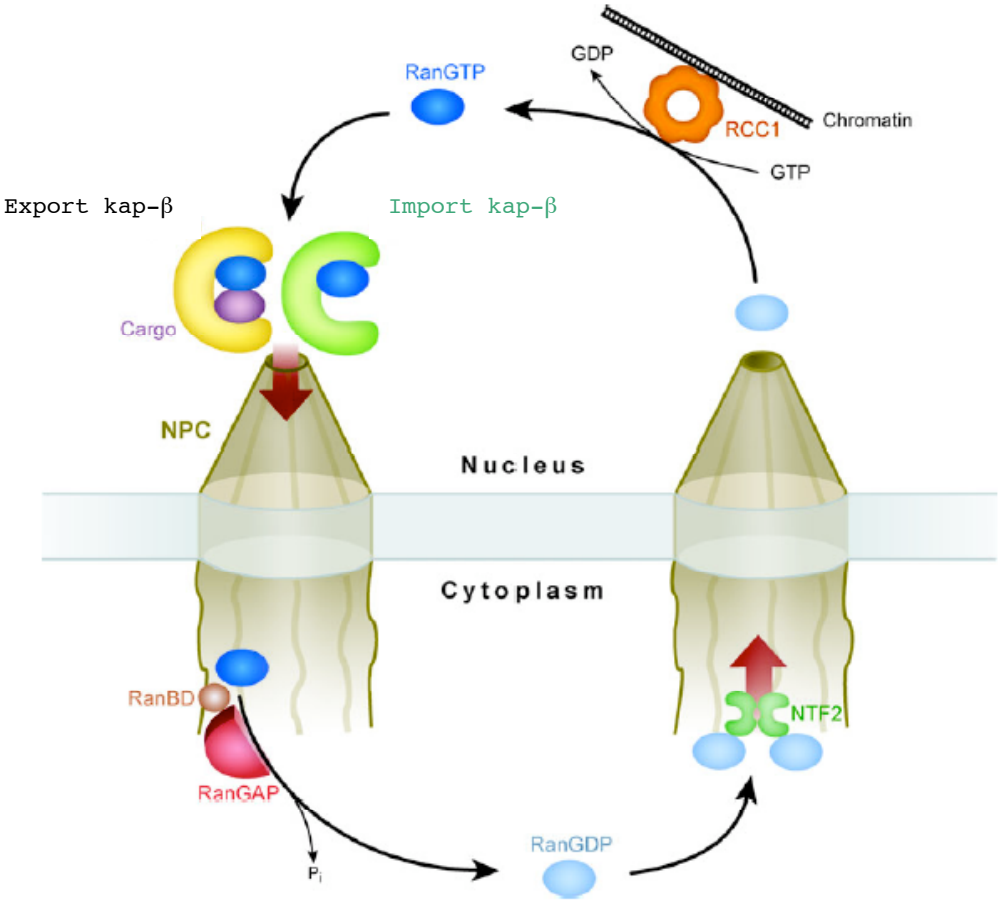


Figure 3. The RanGTP/GDP cycle

The high concentration of RanGDP in the cytosol is maintained by RanGAP, which is bound to the cytoplasmic filaments of the nuclear pore complex. With the help of RanBPs, RanGAP dissociates the cargo/RanGTP/export kap-β or RanGTP/import kap-β complex that enters the cytoplasm by hydrolyzing Ran bound GTP to GDP. The RanGDP returns to the nucleus by binding to NTF2. The high concentration of RanGTP in the nucleus is maintained by the chromatin-bound RanGEF, which converts RanGDP to RanGTP. Image taken from Cook et al., 2007.

1.2 The nuclear pore complex, a selective gate

The main role of the NPC in the interphase of the cell cycle is the regulated selective exchange of macromolecules between the nucleus and cytoplasm (Fig 4). The NPC is constructed from multiple copies of ~30 nucleoporins (nups) (Rout et al., 2000; Cronshaw et al., 2002), which are embedded in the nuclear envelope at sites in which the inner and outer nuclear membranes connect through sharply curved sections of membrane. These nups not only form the scaffold of the NPC, but many of them also contain phenylalanine-glycine (FG) repeats, which have been suggested to be responsible for mediating the selectivity of the NPC against large macromolecules by forming a permeability barrier in the center of the NPC (Shulga and Goldfarb, 2003).

The FG-repeats account for approximately 12-20% of the NPC mass (Rout et al., 2000; Rout and Wente, 1994) and fulfill redundant but consistency functions (Strawn et al., 2004). They also appear to be unstructured (Denning et al., 2003) and, in general, comprise more than 50 FG-repeat units, in which a short cluster of hydrophobic residues is surrounded by more hydrophilic spacer sequences (Denning and Rexach, 2007). Although the characteristics of individual FG-nups have been well studied, the biophysical

nature of the permeability barrier, which is constituted by FG-nups, and the mechanism of translocation through the NPC are still heavily debated.

Key models of how FG-repeats form the selectivity barrier were proposed by Frey et al., Rout et al., Macara, and Patal et al. and these models are discussed in detail in the sections below.

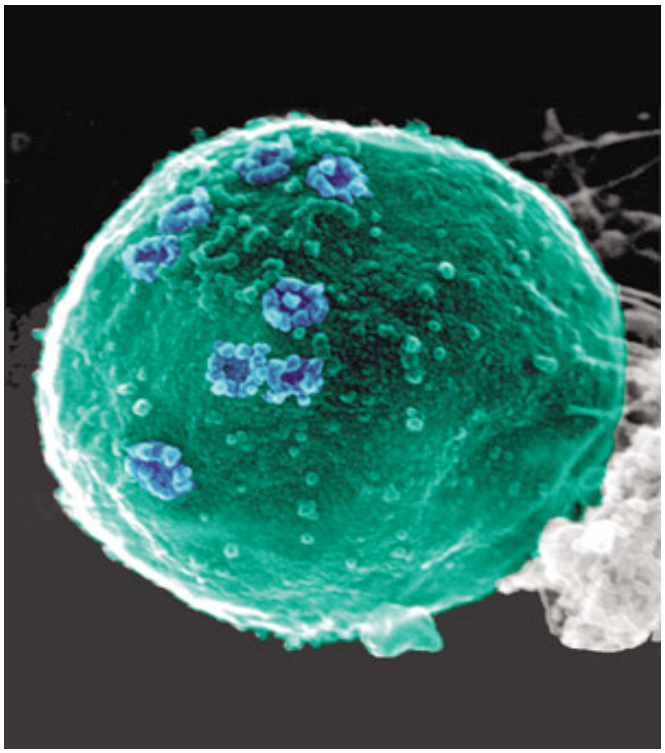


Figure 4. Nuclear pore complexes on the nucleus
Several nuclear pore complexes (NPCs) on the surface of *Saccharomyces cerevisiae* nucleus were visualized by scanning electron microscopy (SEM). Image taken from Kiseleva et al., 2007.

1.2.1 Model 1: The selective phase model

The permeability barrier of the NPC could be physical, energetic, or both, and several models have been proposed with respect to these characteristics.

If the FG-repeats act as a physical barrier, the FG-repeats can be envisioned to attract each other via hydrophobic and/or π-π inter-FG-repeat interactions to form a hydrophobic gel, called a hydrogel (Frey et al., 2006). Moreover, in order to form the hydrogel, the amount of FG-repeats has to reach a "saturated" concentration, which is in a range of millimolar concentration. Like a molecular sieve, the saturated hydrophobic gel would allow the passage of small molecules but restrict the entry of inert molecules, which exceed the mesh size of approximately 30 kDa. Hence, large hydrophilic proteins would be prevented from entering the hydrophobic meshwork. However, since kaps process multiple binding sites for FG-repeats, the kap/cargo complexes would be able to bind and dissolve the sieve-like gel by disengaging the inter-FG-repeat interactions. The process would allow their translocation through the NPC. As a result, this physical barrier provides a selective phase in the NPC channel (Fig 5A).

1.2.2 Model 2: The virtual gating or oily spaghetti model

In the case of an energetic barrier, the peripheral FG-nups would not be able to interact with each other. Instead, they would function as “entropic bristles” or “brushes”. By the movement of the bristles, the available space to access the pore channel is substantially reduced, resulting in an entropic barrier. Hence, the barrier for large molecules to cross the NPC is high in the absence of transport factors due to their inability to enter a narrow central conduit of the NPC that is physically obstructed by FG-repeats. As a result, the NPC would function as a repulsive gate, repelling non–FG-binding molecules. However, in the kap/cargo complexes, kaps can bind to the FG-nups and, by increasing the residence time, the complexes are able to overcome the entropic barrier gradually (Rout et al., 2000; Lim et al., 2006). Eventually, the possibility of entry into the NPC would be dramatically enhanced and cargo transport would occur (Fig 5B).

A similar model, termed the oily spaghetti model, also has been proposed. According to this model, the channel of the NPC is occupied by hydrophobic FG-repeats that could be pushed aside by kap/cargo complexes (Macara, 2001).

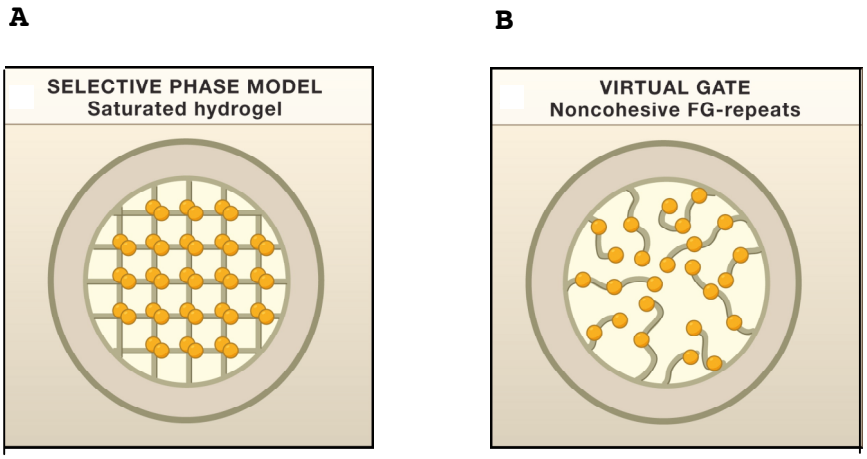


Figure 5. Models for the selective gate

(A) The physical barrier. Once the amount of FG-repeats reaches the "saturated" concentration, the FG-repeats interact with each other to form a highly ordered hydrogel-like structure in the center of the pore. (B) Virtual gating or oily sphaghetti model. The unstructured non-cohesive FG-nups exclude cargo, which are not bound by kaps. FG-repeats (orange) do not mechanically restrict the access of cargo, but binding of kaps to FG-repeats would significantly speed up the translocation process. Image taken from Weis, 2007.

1.2.3 Model 3: The two-gate model

Recently, a model has been proposed which combines features of both the physical and energetic barrier models (Patel et al., 2007). According to this model, FG-nups in the periphery and the center of the NPC have completely different functions. The peripheral FG-nups, which contain FXFG repeats, may behave as non-cohesive entropic bristles, but the centrally located FG-nups, which harbor GLFG repeats, would form a cohesive meshwork.

Therefore, the kap/cargo complexes could bind to the peripheral FG-nups and overcome the entropic barrier in order to reach the central channel. Upon arriving at the central channel, the kap/cargo complexes could further dissolve and penetrate the meshwork, which is held together by hydrophobic attractions between central FG-nups. Hence, in the two-gate model, the NPC contains two types of gates that simultaneously control the diffusion of proteins into and out of the nucleus.

1.3 Overall structural features of the nuclear pore complex

Watson at Rockefeller University first discovered the NPC by electron microscopy (EM) (Watson 1954) and the overall shape of the NPC observed by transmission electron microscopy (TEM) in 1967 appeared to be octagonal (Gall

1967) (Fig 6). Advances and innovations in electron microscopy technology dramatically improved the resolution of NPC imaging, which has increased our understanding of the overall architecture of this macromolecular assembly. According to recent cryo-electron tomography (CET) studies, the NPC is a highly symmetric protein complex, which contains an eight-fold rotational symmetry along the central nucleoplasmic axis and a two-fold symmetry in the plane of the nuclear envelope. Moreover, the NPC consists of an approximately cylindrical symmetric core, which is asymmetrically decorated with filaments on the cytoplasmic side and a “nuclear basket” structure on the nucleoplasmic side (Fahrenkrog et al., 2004) (Fig 7).

In addition, based on scanning transmission electron microscopy (STEM) studies, the mass distribution of the yeast NPC is estimated to be between 55 and 66 MDa, which is much smaller than its vertebrate counterpart with an estimated size of 125 MDa (Yang et al., 1998). Furthermore, as determined by cryo-electron microscopy (CEM), the diameter and height of the NPC in yeast are approximately 1000 and 300 Å, respectively, whereas the NPC in the vertebrate nucleus is significantly larger with an estimated diameter of ~1500 Å and height of ~800 Å (Fig 8). In spite of the size and mass differences between the yeast

and vertebrate NPCs, the general morphology and main features of the NPC still appear to be evolutionarily conserved, such as the overall doughnut-shaped structure and the highly symmetric central core.

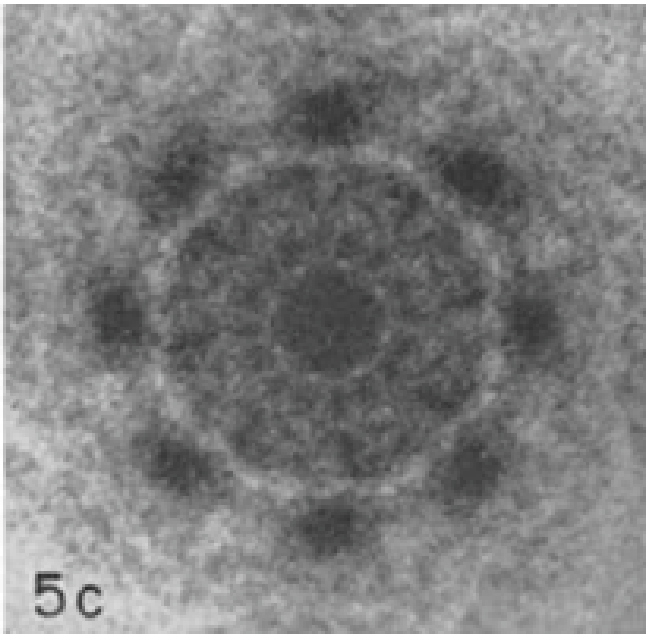


Figure 6. The octagonal shape of the nuclear pore complex

The overall shape of the nuclear pore complex was first observed by transmission electron microscopy from *Triturus* envelopes. The octagonal shape could be clearly identified. Image taken from Gall, 1967.

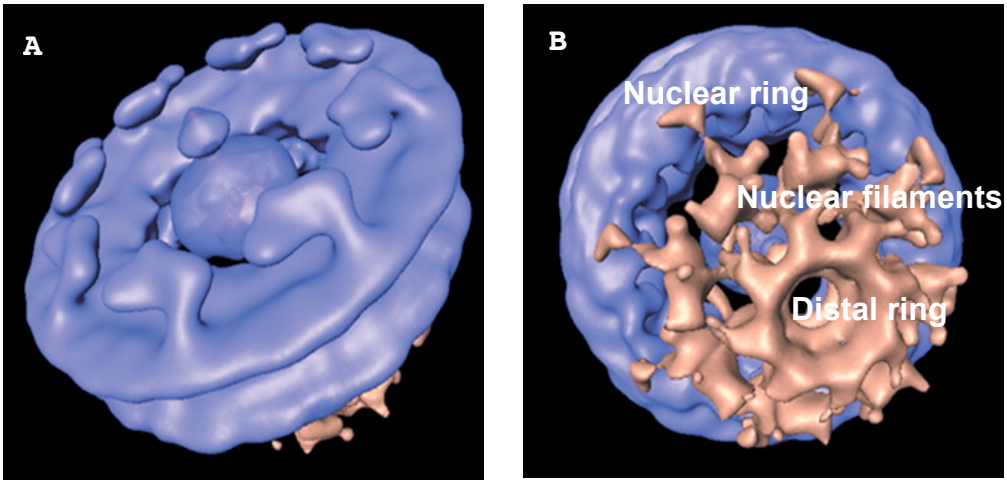


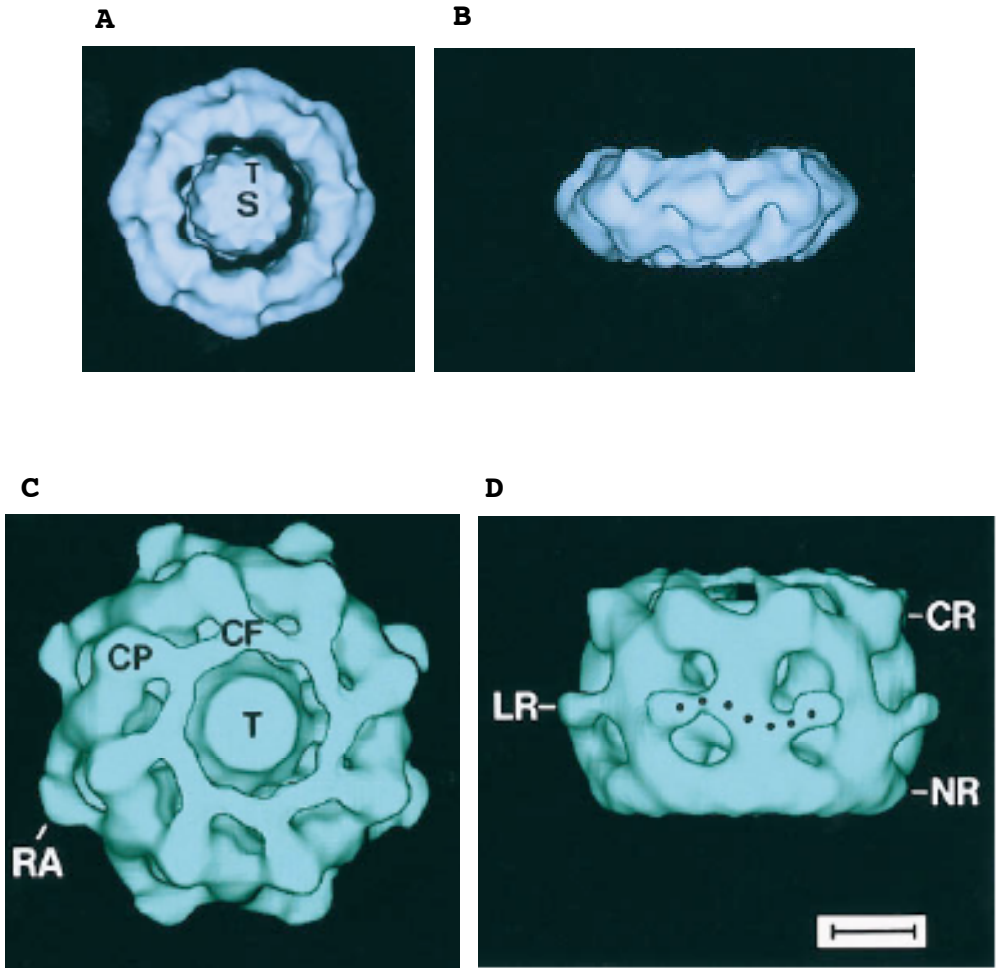
Figure 7. The topology of the nuclear pore complex of *Dictyostelium discoideum*

Overall nuclear pore complex structure was determined by cryo-electron tomography at approximately 90 Å resolution, illustrating that the NPC is a highly symmetric protein assembly. The central core of the NPC is shown in blue and the nuclear basket is shown in brown. (A) The cytoplasmic face of the NPC. The NPC contains an eight-fold rotational symmetry across the center of the pore and a two-fold symmetry in the plane of the nuclear envelope. (B) The nucleoplasmic face of the NPC. The nuclear filaments link the distal ring to the nuclear ring, forming a basket structure, namely nuclear basket. Image taken from Beck et al.,

Figure 8. 3-Dimensional structures of the yeast and vertebrate nuclear pore complexes

The vertebrate NPC is significantly larger than the yeast NPC. (A) Top view from the cytoplasmic surface, and (B) side view of the yeast NPC. The transporter, a weak smeared density in the center of the pore and are suggested to be FG-repeats and cargo/kap complexes at present, is marked by (T), and putative substrates are marked by (S). (C) Top view from the cytoplasmic surface, and (D) side view of the vertebrate NPC. Besides the inner spoke ring that exists in both yeast and vertebrate NPCs, lumenal ring (LR), cytoplasmic ring (CR), and nuclear ring (NR) are unique features discovered in vNPC. The LR is supposed to anchor the NPC to the nuclear envelope and radial arm (RA) structure is part of the LR. In the central pore of the NPC, A ring of collapsed cytoplasmic filaments (CF) originates from the cytoplasmic particles (CP). Scale BAR = 300 Å. Image taken from Yang et al., 1998.

Figure 8



1.4 Density and distribution of nuclear pore complexes

The total number of NPCs in the nucleus varies tremendously with different cell sizes. In total, there are ~200 NPCs per nucleus in yeast S. cerevisiae (Rout and Blobel, 1993), ~2000 NPCs per nucleus in human cells (Maul et al., 1972), and in a mature *Xenopus* oocyte, the number of NPCs can reach ~50 million per nucleus (Cordes et al., 1995). Surprisingly, NPCs on the yeast nucleus are not equidistantly spread but rather gather into regions of higher density (Winey et al., 1997). Meanwhile, the sum of NPCs increases steadily during the cell cycle and reaches its maximum number in S phase during the cell cycle (Winey et al., 1997).

1.5 Composition of the nuclear pore complex: nucleoporins

1.5.1 Nucleoporins in yeast

Based on biochemical analyses, the NPC has been demonstrated to be comprised of only about 30 distinct proteins (Rout et al., 2000). Furthermore, in vivo studies suggested that different nups have distinct roles in the functions of the NPC. For instance, some nups containing FG-repeat motifs provide docking sites for kaps (Strawn et al., 2004). Three nups have been identified as integral

membrane proteins, which may be involved in anchoring the NPC to the nuclear membrane (Wozniak et al., 1994; Chial et al., 1998; Rout et al., 2000). Many of the nups function as structural elements that interact with other nups to constitute the scaffold of the NPC.

Unlike eukaryotic cells that undergo open mitosis, yeast cells undergo a closed mitosis, and, hence, the disassembly of the NPC into distinct sub-complexes does not occur in yeast. Nevertheless, sub-complexes can be isolated and defined also from yeast by the dissociation of isolated intact NPCs in different ranges of salt, pH, and detergent conditions. Moreover, these sub-complexes have been confirmed by biochemical and genetic studies (Suntharalingam and Wente, 2003). For example: the Nup84 sub-complex, which contains Nup85, Nup120, Nup145C, Nup84, Nup133, Sec13, and Seh1, build the core scaffold of the NPC. The Nsp1 sub-complex, which contains Nic96, Nsp1, Nup57, and Nup49, probably lines the central channel of the pore. Among all the sub-complexes in yeast, the Nup84 sub-complex is the best characterized.

1.5.2 Nucleoporins in vertebrates

Despite the fact that yeast and vertebrate NPCs are constructed by approximately the same number of nups and

that these nups are evolutionarily conserved, the nups in yeast and vertebrates are not completely analogous. For instance, in addition to the seven proteins in the yeast Nup84 sub-complex, two additional components, Nup37 and Nup43, have been identified in the vertebrate counterpart, the Nup107-Nup160 sub-complex (Loiodice et al., 2004). Only two membrane proteins, gp210 and Pom121, were found in vertebrates instead of three, Pom152, Pom34 and Ndc1, that were found in yeast.

Due to nuclear envelope break down (NEBD) and open mitosis, the human NPC has been found to disassemble during mitosis into several sub-complexes, which reassemble at the end of telophase (Vasu and Forbes, 2001) (Fig 9). For instance, in human cells, the nuclear basket contains Nup153/Nup50 and Nup98/Gle2 sub-complexes. The central scaffold of the pore has two major sub-complexes: the Nup107-Nup160 sub-complex, containing Nup85, Nup160, Nup96, Nup107, Nup133, Seh1 and Sec13 and the Nup188 sub-complex, containing Nup205, Nup188, Nup155, Nup93, and Nup35. The Nup62 sub-complex, including Nup93, Nup62, Nup58, Nup54 and Nup45, is suggested to be located closest to in the central transporter channel. The cytoplasmic filaments have been determined to contain Nup214, Nup88, and Nup358.

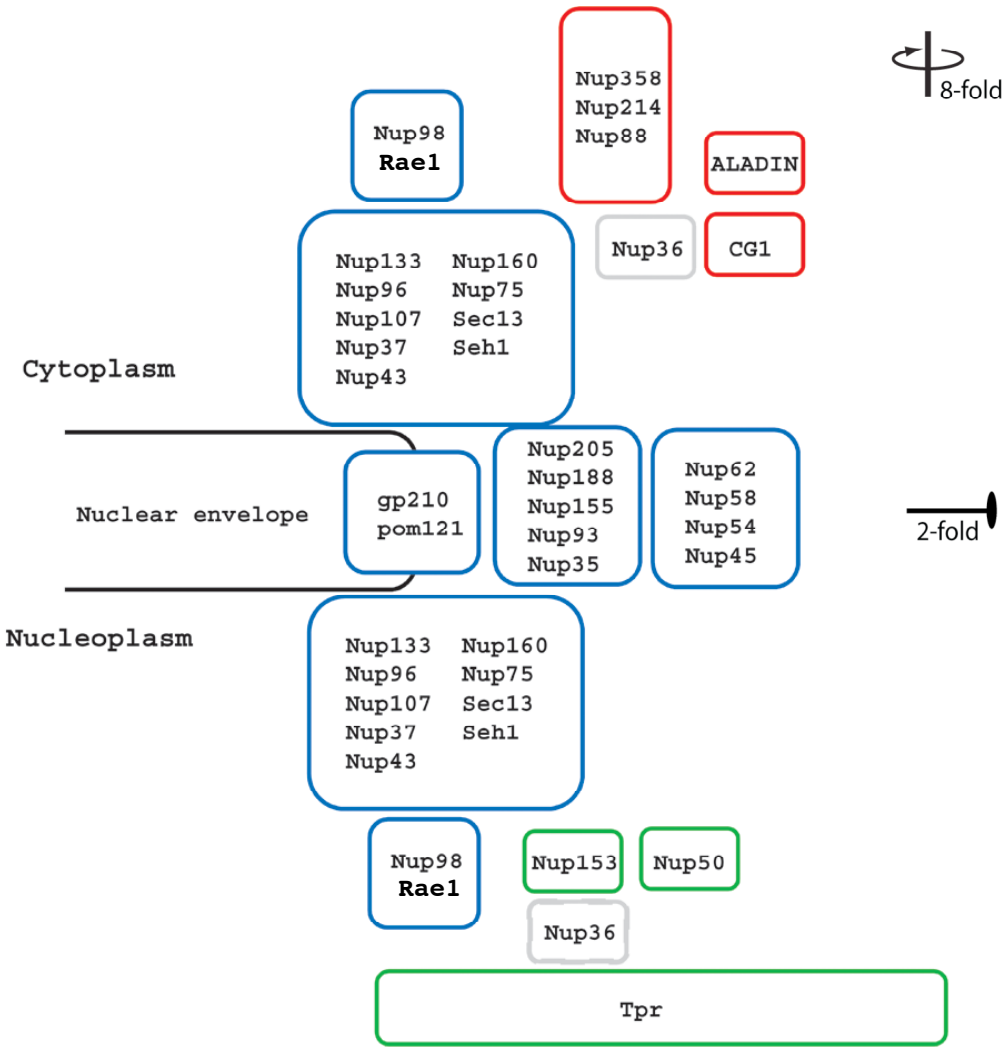


FIG 9: Schematic representation of the major nucleoporin sub-complexes in the vertebrate NPC

The NPC can disassemble into several sub-complexes during mitosis and reassemble at the end of mitosis. Sub-complexes are shown in boxes. Blue (symmetric core), green (nuclear basket), and red (cytoplasmic filament) boxes indicate the locations of the nucleoporins. Nup36 has not been identified the position and is indicated by gray box. Eight-fold and two-fold symmetries are indicated. Image adapted from Bapteste et al., 2005.

1.6 Structure of nucleoporins

According to a protein folding and modeling prediction analysis of nups, only ten fold types are assigned to all nups: the α-solenoid fold, the FG-repeat, the β-propeller fold, the coiled-coil motif, the transmembrane helices (TMH) motif, the cadherin fold, the autoproteolytic Nup98 domain, the RNA recognition motif (RRM), the Zinc finger motif and the Ran-binding domain (Fig 10) (Devos et al., 2006). Among these ten domains and motifs, the β-propeller fold, covering approximately 16% of the residues, the α-solenoid fold, covering the most residues (~38%), and the coiled-coil motif, covering approximately 5%, are the most prominent folds.

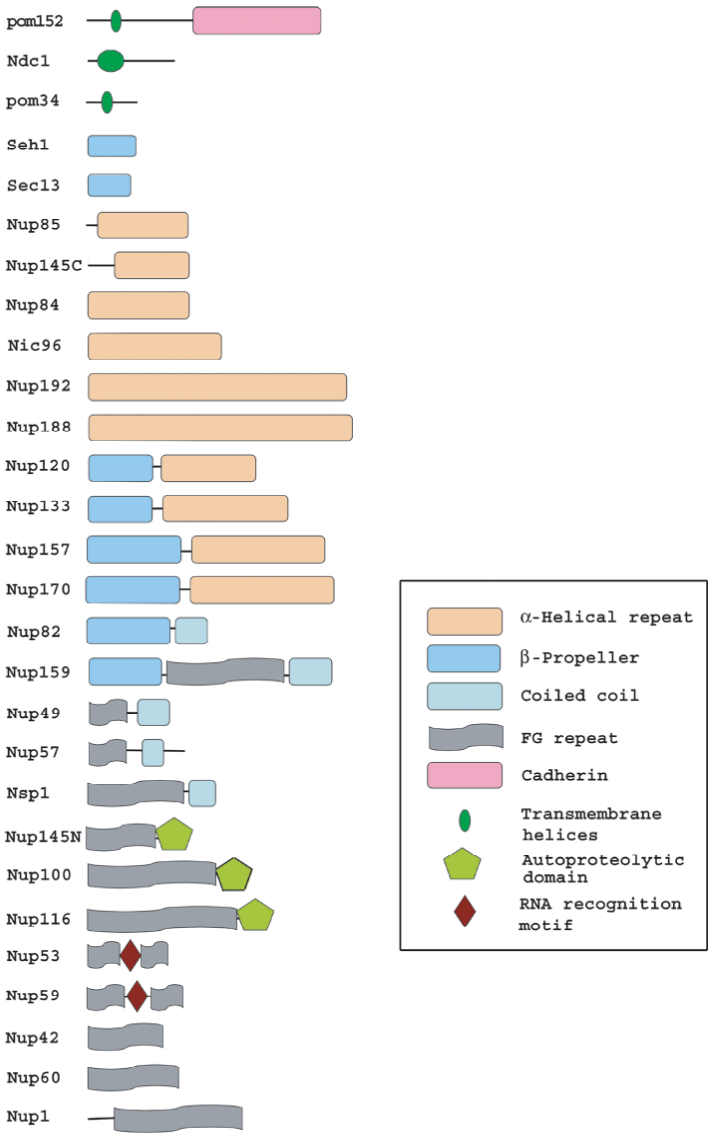


Figure 10. Domain architecture of nucleoporins in the the yeast NPC

Different motifs and domains are indicated by different colors. In addition to the FG repeats, the β-propeller fold, the α-solenoid fold, and the coiled-coil motif are the most prominent structural elements in the NPC. Image adapted form Devos et al., 2004.

1.6.1 The β-propeller fold

The β-propeller fold is a type of all-β strand protein architecture, which contains 4 to 8 blade-shaped β sheets arranged around the central axis (Fig 11). Each sheet typically has four anti-parallel β strands twisted so that the predominantly hydrophobic interaction between the sheets provides most of the required structural stability. Most of the known β-propeller proteins are closed circular structures, achieved by forming a "Velcro closure" structure, in which the N-terminal blade is completed by one or two strands at the C-terminus of the protein.

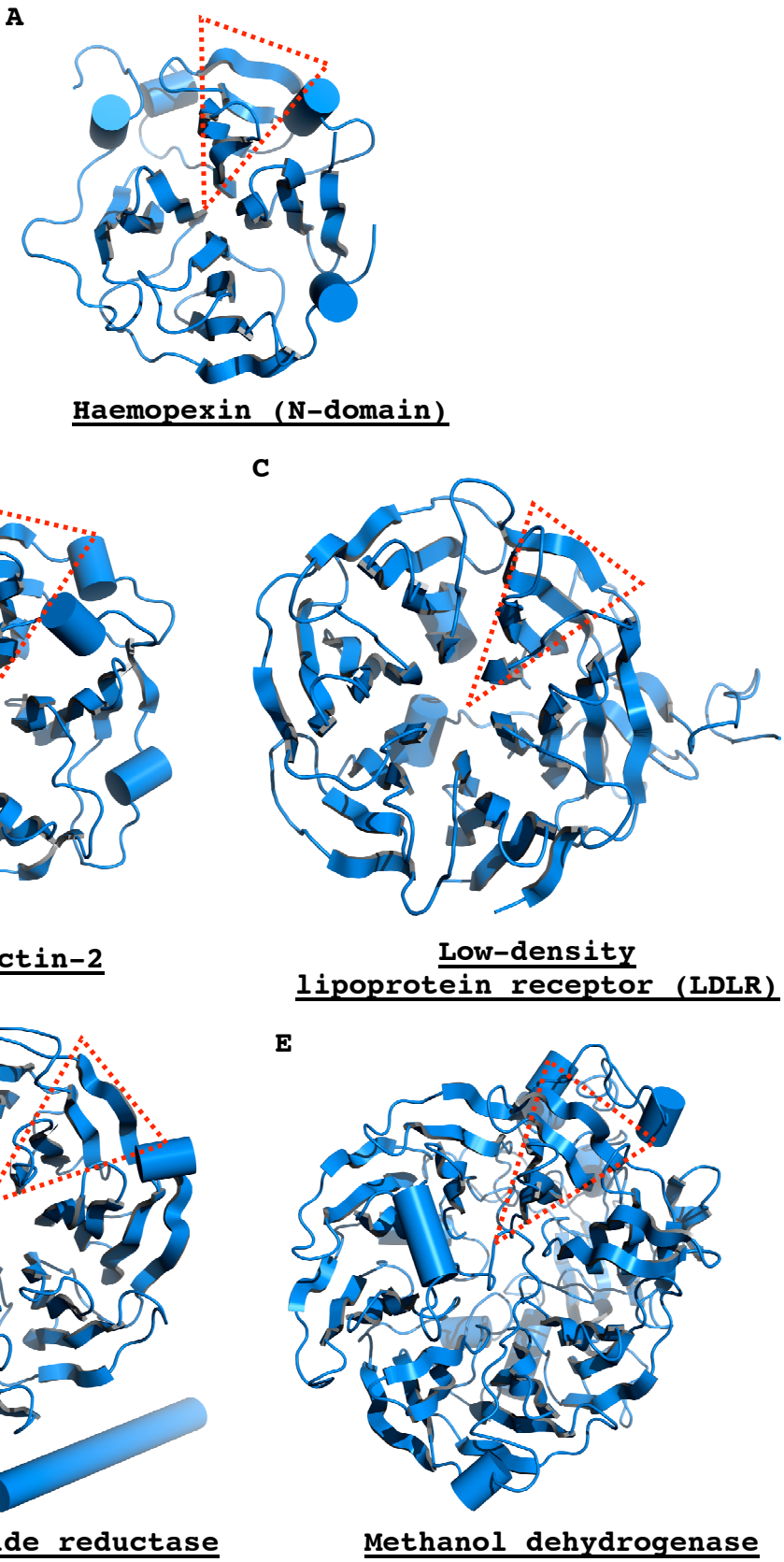
The β propeller proteins can be divided into several families. Currently the largest single family, which also has the most diverse functions, is the WD repeat β-propeller, which contains 4 or more copies of the conserved WD (tryptophan-aspartate) sequence motif (Ron et al., 1994), which is approximately 40 amino acids long and often ends with a tryptophan and an aspartate. In contrast to the WD repeat propellers, none of which have enzymatic activity in the propeller domain, non-WD repeat β-propellers are frequently enzymes (Faber et al., 1995). The non-WD repeat β-propellers might be less rigid than WD repeat β-propellers to allow for the flexibility that is necessary for catalytic activity.

The β-propeller structure creates a stable platform for protein-protein interactions, due to the three potential interacting surfaces, the top, the bottom, and the rims. Hence, the β-propeller fold in the NPC appears to be mainly involved in protein-protein recognition. For instance, the N-terminal β-propeller domain of yeast Nup159 functions in mRNA export as a platform for recruiting Dbp5, a DEAD box RNA helicase, to the nuclear side of the NPC. Eventually, the Dbp5 unwinds mRNA in mRNA protein particles (mRNP), which triggers the release of proteins from mRNPs that travel through the NPC (Zhao et al., 2002).

Figure 11. Typical β-propeller folds

(A) Four-bladed structure of the N-terminal domain of haemopexin (PDB code 1qhu, Paoli et al., 1999); (B) The five-bladed structure of tachylectin-2 (PDB code 1tl2, Beisel et al., 1999); (C) The six-bladed the YWTD protein low-density lipoprotein receptor (PDB code 1ijq, Jeon et al., 2001); (D) The seven-bladed structure of the β subunit of the G protein, Gβ(PDB code 1got, Lambright et al., 1996); (E) The eight-bladed propeller of methanol dehydrogenase (PDB code 4aah, Xia et al., 1996). The triangle indicates one of the blades in each β-propeller fold.

Figure 11


A
Haemopexin (N-domain)
B
Tachylectin-2
C
Low-density lipoprotein receptor (LDLR)
D
Nitrous oxide reductase
E
Methanol dehydrogenase

1.6.2 The α-solenoid fold

The basic structural unit of α-solenoid proteins is a hairpin-like structure with two or three repeats of an α-helix, which stack on top of each other to form a super-helix domain (Fig 12). The canonical Armadillo repeat, such as kap-α, has three α-helices, while the HEAT repeat, such as kap-β1, harbors two anti-parallel α-helices. These repeating units interact with each other mainly via hydrogen-bonding and hydrophobic amino acid residues. Due to the lack of stabilizing contacts between repeating units distant from each other, α-solenoid proteins show relatively high flexibility in their structure.

Moreover, most of the α-solenoid proteins are rod-like helical structures with right-handedness, which means that the polypeptide chain winds along the super-helical axis like B form DNA. Two additional structural features, twist and curvature, also have tremendous influence on the overall shape of the solenoid fold. Twist is determined by how neighboring repeats stack on top to each other, while curvature is generated by the rotations of repeating units along the concave side of the α-solenoid fold.

Since the non-globular shape of the solenoid fold provides large and diverse surfaces, the major functions of this protein fold is supposed to mediate multivalent

protein-protein interactions. For instance, kap β1, an α-solenoid protein containing HEAT repeats, mediates nuclear import and has been identified as a platform for binding various proteins. Firstly, kap β1 is able to bind the N-terminal domain of kap α, an adaptor protein for NLS recognition. Kap α is wrapped within the kap β1 super-helix (Cingolani et al., 1999). Secondly, kap β1 also can interact with RanGppNHp (Vetter et al., 1999) Moreover, comparison of kap α bound to kap β1 with RanGppNHp bound kap β1 reveals that kap α and Ran share common contact residues on HEAT repeat 7 and an acidic loop of kap β 1, suggesting competition of the two proteins for a common area in kap β1. Thirdly, two FG-repeats of Nsp1 and a FG-repeat peptide are bound to the convex side of the kap β1 arch (Bayliss et al., 2000). As a result, the α-solenoid fold of kap β1 provides multiple sites for binding different substrates in order to execute various biological functions (Fig. 13).

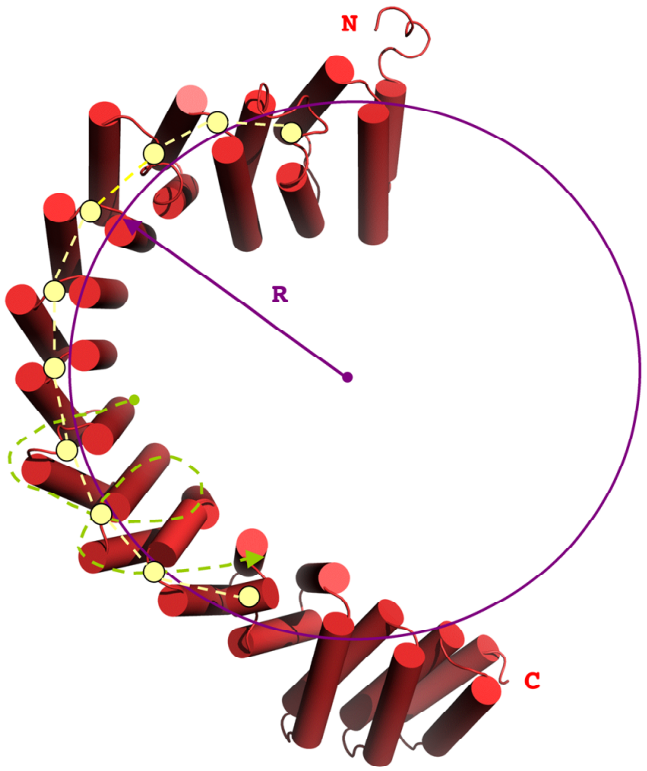


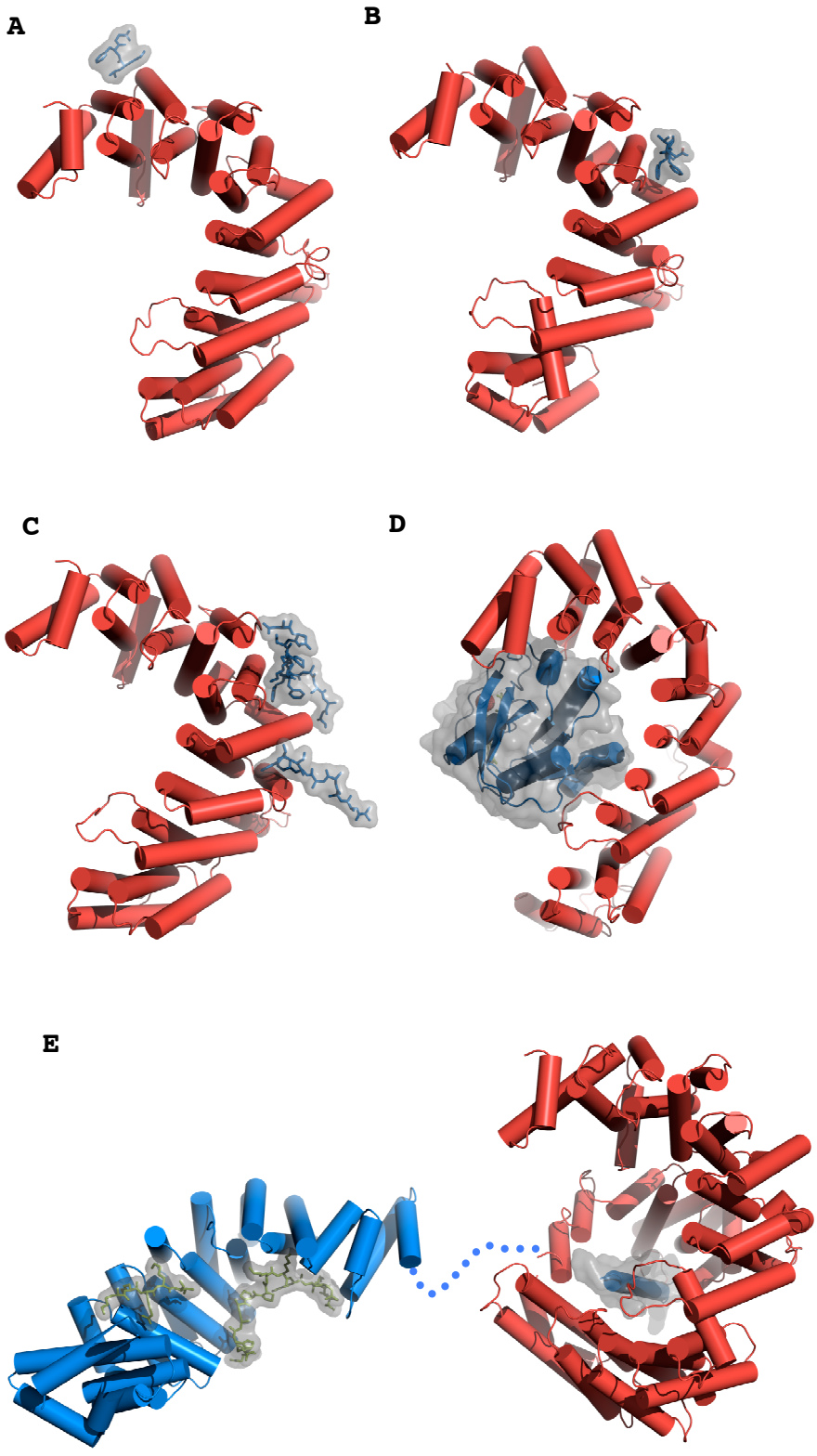
Figure 12. A schematic representation of a typical α-solenoid protein structure

A typical α-solenoid structure, phosphatase 2A PR65/A subunit (PDB code 1b3u, Groves et al., 1999), shows the basic parameters of solenoid structures: curvature (magenta); handedness (green); and twist (yellow). Curvature could be described by the radius (R) of a circle, which is drawn based on the super-helical axis. The smaller the radius is, the bigger curvature occurs. Handedness indicates how the α-solenoid winds along the super-helical axis. Most of the α-solenoid proteins show right-handedness instead of left-handedness. Twist is shown by the yellow points in consecutive repeating units. Image adapted from Kobe AND Kajava, 2000.

Figure 13. Karyopherin β1 functions as a multivalent binding platform

(A) (B) Human kap β1/GLFG peptide complexes (PDB code 1o6o and 106p, Bayliss et al., 2002). Human kap β1 is α helical molecule constructed from 19 HEAT repeats. The N-terminus of kap β1, containing ~440 residues, was co-crystallized with GLFG peptide. Two sites were occupied by this peptide on kap β1. All kap β1s are shown in red and binding substrates, GLFG, Nsp1, kap α, and RanGppNHp, are shown in blue. (C) The human kap-β1 complexed with the FG-repeat region of yeast Nsp1 (PDB code 1f59, Bayliss et al., 2000). Twenty-one residues were observed and bind to the convex side of N-terminal arch of the kap β1. (D) The human kap β1/Ran·GppNHp complex (PDB code 1ibr, Vetter et al., 1999). (E) The complex of human full-length kap β1 and N-terminal α-helix region of kap α(PDB code 1qgr, Cingolani et al., 1999). The C-terminal half of kap β1 wraps around the helix and an acidic loop of kap β1 further binds the extended chain. In addition, the structure of the truncated mouse kap α, shown in blue, and SV40 T antigen NLS, shown in gray, complex (PDB code 1q1s, Fontes et al., 2003) is connected by a dashed line to the N-terminus of kap α..

Figure 13



1.6.3 The coiled-coil motif

The coiled coil is a common structural motif, formed by approximately 3 to 5% of all residues in proteins (Wolf et al., 1997) and consists of two to five α-helices wrapped around each other into a left-handed helix to form a supercoil. The interacting surfaces between the helices often contain hydrophobic residues, such as leucines in the leucine zipper.

Unlike the regular α-helices containing 3.6 residues per turn, the left-handed coiled-coil only contains around 3.5 per turn. Hence, a heptad repeat, $(abcdefg)_n$, often appears every two turns in the coiled coil motif and typically has hydrophobic residues at position **a** and **d**, and polar residues at position **e** and **g** (Lupas 1997). The coiled-coil is the most widespread motif involved in oligomerization as well as binding to target proteins. It has been identified in various proteins, from leucine zipper transcription factors such as Fos and Jun transforming proteins (Glover and Harrison, 1995), involved in cell growth and proliferation, to matrilins, involved in the development of cartilage and bone (Frank et al., 2002).

An intriguing function of a coiled-coil motif has been proposed in the NPC based on a recent structural study (Melcak et al., 2007). The α-helical region of mammalian

Nup58/45, two FG-nups that belong to the Nup62 sub-complex and form the central tranport channel of the NPC, was determined by x-ray crystallography. According to the structures, Nup58/45 forms distinct tetramers, each consisting of two anti-parallel hairpin dimers. The intradimeric interface is hydrophobic, whereas dimer-dimer association occurs through large hydrophilic residues. By superposition of the different crystal structures, the rigid Nup58/45 dimers are sequentially shifted along the continuous electrostatic tetramerization interface by a distance of ~11 Å, suggesting an intermolecular sliding mechanism (Fig 14).

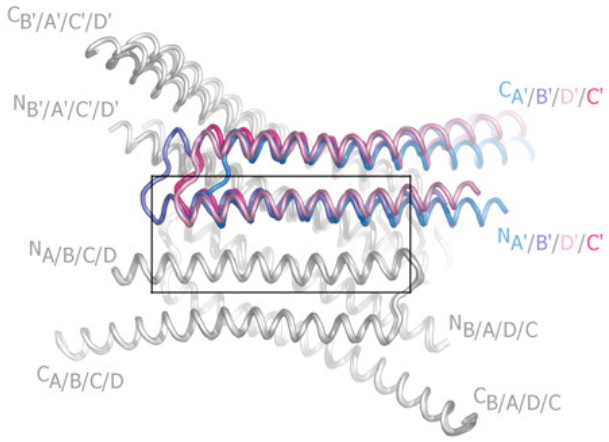
Remarkably, the central channel of the NPC shows the ability of altering its diameter. Hence, the sliding between the two Nup58/45 dimers could result in an overall extension of circumference of the pore, and, ultimately, the NPC then could facilitate the transport of cargo of relatively large size such as pre-ribosomal units (Fig 15).

Figure 14. Intermolecular sliding mechanism of two Nup58/45 dimers

(A) Nup58/45 forms distinct tetramers, each consisting of four anti-parallel hairpin monomer. All monomers of two tetrameric Nup58/45 assemblies are superimposed. The shift is observed between the different monomers, shown in colors. The intradimeric interface is hydrophobic, whereas the area involved in dimer-dimer association contains hydrophilic residues, which are highlighted in the box, and is suggested to be capable of sliding. (B) The direction and the approximate sliding distance are indicated by black arrows. From states I to IV, the two sliding helices shifts approximately 11 Å. The interface residues are labeled and colored in all four states according to state I. The electrostatic interactions between are represented by red dashed lines, and water molecules are shown as blue spheres. Image taken from Melcak et al., 2007.

Figure 14

A


CB'/A'/C'/D'
NB'/A'/C'/D'
CA'/B'/D'/C'
NA'/B'/D'/C'
NA/B/C/D
NB/A/D/C
CA/B/C/D
CB/A/D/C

B

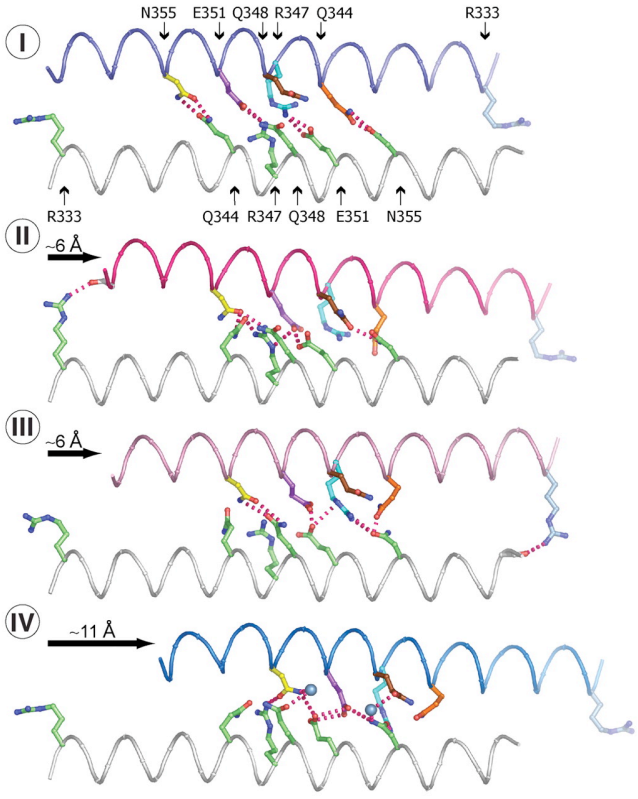
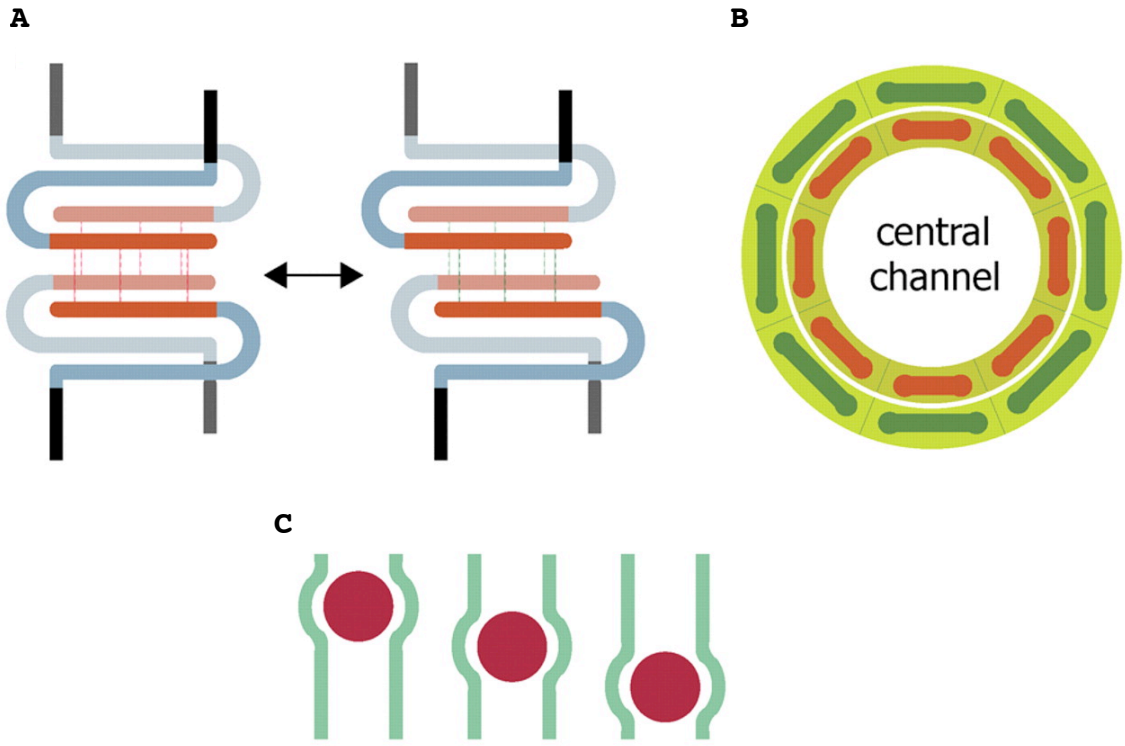

N355
E351
Q348
R347
Q344
R333
I
R333
Q344
R347
Q348
E351
N355
II
~6 Å
III
~6 Å
IV
~11 Å

Figure 15. The Nup58/45 sliding mechanism results in pore expansion

(A) Schematic representation of the Nup58/45 sliding module. The four N-terminal helices that generate the tetramerization interface (orange), the C-terminal helices (light blue), and the C-terminal FG-repeats (black) are indicated. Due to the hydrophilic interface, the sliding of the Nup58/45 dimer is mediated by a number of hydrogen bonds, which are indicated by red and green thin lines. (B) Schematic representation of the central channel of the NPC. Nup58/45 are FG-nups and belong to the Nup62 sub-complex, which is suggested to be located in the central channel and facilitates cargo transport. According to the eight-fold symmetry across the center of the pore and relative localization, eight Nup58/45 tetrameric assemblies are circularly arranged to form a ring (red bars). Once the Nup58/45 dimers slides, the entire Nup58/45 tetramer (green bars) also extends, which causes the expansion of the circumference of the pore. (C) The adjustment of the circumference of the pore, accomplished by sliding of the Nup58/45 tetramers, could facilitate the transport of cargo (red spheres) across the central channel (green). Image taken from Melcak et al., 2007.

Figure 15



1.7 Structure and function of the Nup84 sub-complex

In yeast, the Nup84 sub-complex is assembled from seven proteins, Nup84, Nup85, Nup120, Nup145C, Seh1, Sec13 and Nup133 and, thus, is also named the heptameric complex (Fig 16).

In vivo dissection from yeast, employing the tagging of proteins and subsequent pullout experiments, initially identified a pentameric Nup84 sub-complex consisting of Nup84, Nup120, Nup85, Sec13, and Seh1 (Siniossoglou et al., 1996). Subsequent pullout experiments resulted in the identification of another two members, Nup145C and Nup133, and established a heptameric Nup84 sub-complex (Allen et al., 2001; Siniossoglou et al., 2000). Eventually, a stable heptameric Nup84 sub-complex could be purified from yeast, which was observed to be a Y-shaped structure by negative stain electron microscopy (Siniossoglou et al., 2000). In order to reconstitute the heptameric complex from recombinant proteins, Lutzmann et al. co-expressed up to three nups in *E. coli* and purified the resulting complexes. Finally, the heptameric complex was obtained by mixing the recombinant protein complexes (Lutzmann et al., 2002). Furthermore, by subjecting individual sub-complexes and the assembled heptameric complex to shape analysis using negative stain microscopy, the relative position of the

seven components in the Nup84 sub-complex was established. Nup133 is at the base, followed by Nup84, the Nup145C/Sec13 complex in the center and Nup120/Nup85/Seh1 on the top. The entire Nup84 sub-complex measures about 400 Å in length and forms a Y-shaped structure.

Although the spatial arrangement of the heptameric complexes in the symmetric core of the NPC is currently unknown, the components of the heptameric complex have been demonstrated to be involved in various cellular functions.

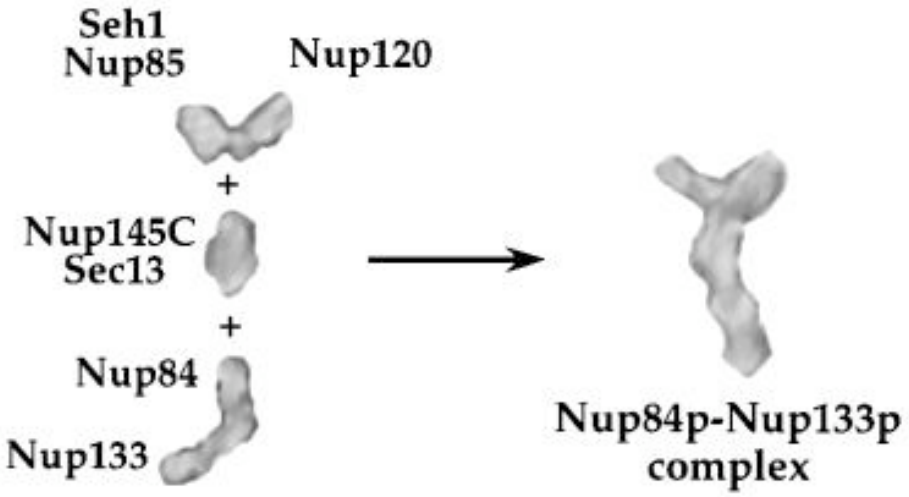


Figure 16. Negative stain EM of the Nup84 sub-complex
Relative positions of the components of the Nup84 sub-complex were established. Nup133 is at the base, followed by Nup84, the Nup145C/Sec13 complex in the center and Nup120/Nup85/Seh1 on the top. Note that the depicted EM pictures are processed extensively and thus should be considered only as models. Image taken from Lutzmann *et al.*, 2002.

1.7.1 The Nup84 complex is a major building block in the symmetric core of the NPC

The Nup84 sub-complex consists of seven proteins and has a molecular weight of approximately 600 kDa (Siniossoglou et al. 2000). In addition to the seven components in the yeast Nup84 complex, the vertebrate counterpart contains two extra proteins, Nup43 and Nup37 (Loiodice et al., 2004). Hence, among the 30 nups of the NPC, nearly one-third can be assigned to this Nup84 complex that thus represents a major building block of the NPC. Deletion or immuno-depletion of any nups from this sub-complex in yeast or other eukaryotic cells results in dramatic impairment of the organization of the NPC (Bai et al., 2004; Boehmer et al., 2003; Dockendorff et al., 1997; Siniossoglou et al., 1996; Siniossoglou et al., 2000).

1.7.2 The Nup84 sub-complex has been proposed to serve as a membrane-curving module

According to immuno-gold electron microscopy studies, the Nup84 complex appears to be localized adjacent to the pore membrane domain (POM) of the nuclear envelope, which is the highly curved region linking the inner and outer nuclear membrane (Krull et al., 2004; Rout et al., 2000).

Interestingly, a component identified in the COPII coat, Sec13, was also discovered in the Nup84 sub-complex. The sharing of Sec13 in two systems further implies that the Nup84 sub-complex may have similar functions as the COPII coat, which has the ability to enforce membrane curvature during vesicle budding. Moreover, due to similar protein folds in the Nup84 sub-complex and the vesicle coats, clathrin, COPI and COPII, the Nup84 sub-complex was also proposed to be capable of shaping the pore membrane domain (Devos et al., 2004). Altogether, the Nup84 sub-complex is suggested to function as a membrane curving module to coat or maintain the membrane curvature in the NPC (Fig 17).

Interestingly, Nup133, a component of the Nup84 complex, has been shown to contain a membrane-curvature-sensing motif, which is an amphiphatic α-helix containing hydrophobic and polar residues on two opposite faces, which matches well the chemical characteristics of the membrane interface. The N-terminus of Nup133 is a seven bladed β-propeller fold decorated by several loops between the blades, as determined by x-ray crystallography (Berke et al., 2004).

Bioinformatics predicted that the membrane-curvature-sensing motif is located in one of the loops in the β-propeller domain of the Nup133. Moreover, according to

circular dichroism (CD) experimental results, a synthetic peptide, which contains the same protein sequence as in this particular loop, can fold into an α-helix upon interacting with small liposomes (Drin et al., 2007). In order to detect the membrane, Nup133 has to be close to the nuclear envelope. Thus, the localization of the Nup84 sub-complex, which is proposed to be near the membrane pore domain within the NPC, is further supported by this discovery.

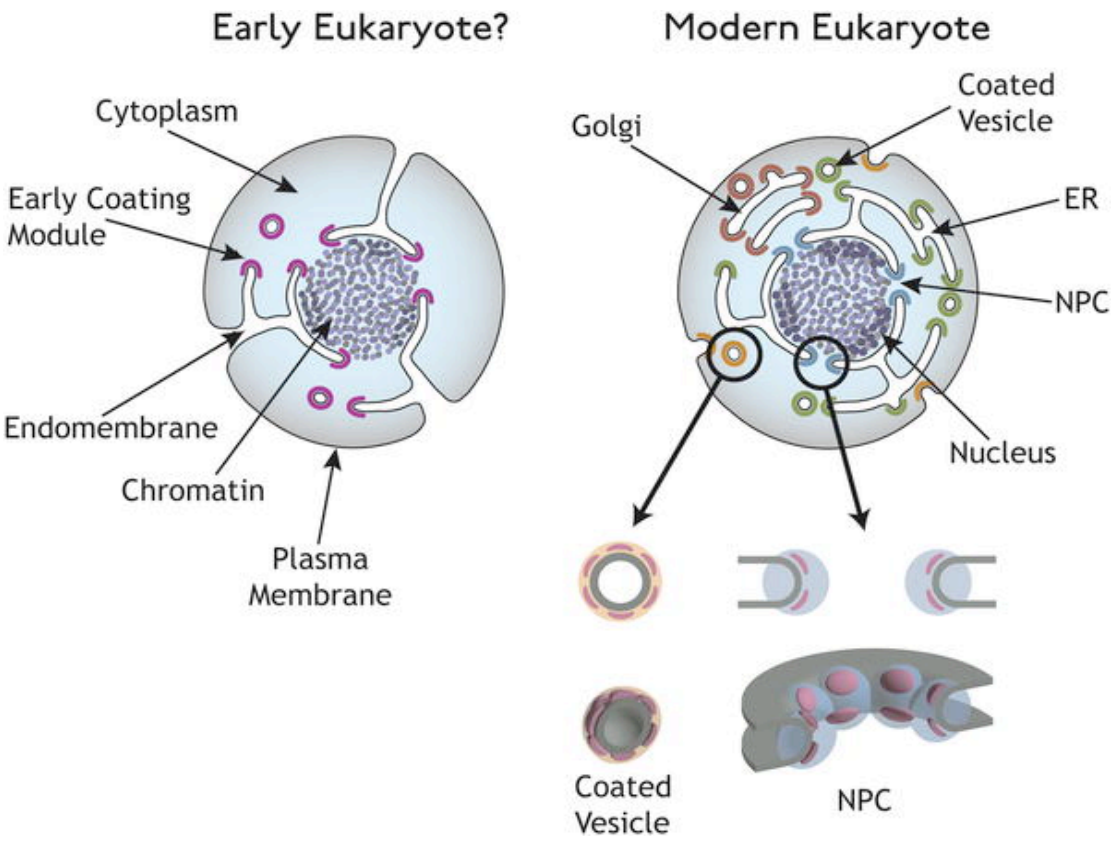


Figure 17. A model for the evolution of the coated vesicle and the nuclear pore complex

In order to form the early endomembrane systems, early eukaryotic cells employed a membrane-curving protein module (purple), also named protocoatmer, to induce curvature of the plasma membrane. During divergent evolution, various coated vesicles (orange, green and brown) and the Nup84 sub-complex of the NPC (blue) in modern eukaryotic cells would have evolved from this protocoatomer in order to fulfill similar functions in diverse contexts. Hence, this module (pink) is evolutionarily conserved in both NPCs and coated vesicles for a purpose of stabilizing membrane curvature. Image taken from Devos et al., 2004.

1.7.3 The Nup84 sub-complex has been shown to regulate gene expression

20 years ago, the gene gating hypothesis was proposed by Günter Blobel (Blobel, 1985) based on the two observations that, firstly, NPCs do not distribute randomly around the nuclear envelope, and, secondly, abundant chromatin is attached to the NPC-rich regions of the nuclear envelope. The NPCs were envisioned to serve as gene-gating organelles capable of interacting specifically with transcriptional active chromatin in the nucleus and to regulate gene expression.

Interestingly, Galy et al. showed that Nup145C, a component of Nup84 sub-complex, is a potential docking site in the NPC for recruiting several proteins involving in gene silencing, such as repressor activator protein 1 (Rap1), silent information regulator 3 (Sir3), and silent information regulator 4 (Sir4) and eventually organizes perinuclear chromatin domains, an area of transcriptional repression, close to the nuclear envelope (Gotta et al., 1996; Galy et al., 2000).

1.8 Characteristics and functions of individual nucleoporins of the Nup84 sub-complex

1.8.1 Nup145C

Nup145 is expressed as a 145kDa precursor protein and auto-catalytically cleaved in vivo into two functionally distinct proteins (Teixeria et al., 1997 and 1999). The N-terminal half of Nup145, Nup145N, contains multiple FG-repeats and is similar to Nup100 and Nup116, which are FG-nups responsible for kap binding (Iovine and Wente, 1997). The C-terminal half, Nup145C, assembles into the Nup84 sub-complex (Teixeria et al., 1997) and forms a complex with Sec13 (Lutzmann et al., 2002). Moreover, Nup145C was predicted to fold into an α-helical domain (Devos et al., 2004).

1.8.2 Sec13

Sec13 was first identified by temperature-sensitive mutations, which block vesicle formation at the non-permissive temperature (Kaiser and Schekman, 1990) and further proved to be a component of the COPII vesicle coat (Salama et al., 1993). In the COPII coat, Sec13 forms a complex with Sec31 and acts as a structural scaffold to gather adjacent Sec23/24 complexes into a coat lattice

during the biogenesis of COPII-coated vesicles (Matsuoka et al., 2001). Subsequently, Sec13 was also discovered as a component at the Nup84 sub-complex of the NPC that has a binding partner, Nup145C (Siniossoglou et al., 1996; Rout et al., 2000; Lutzmann et al., 2000). In addition, Sec13 was predicted to be a six-blade β-propeller based on the presence of six WD40 repeats (Garcia-Higuera et al., 1996).

1.8.3 Nup85

Nup85 was initially identified by screening for mutants defective in mRNA export from the nucleus (Goldstein et al., 1996). Nup85, an approximately 85 kDa α-solenoid protein, has been proposed to serve as a scaffold protein in the NPC (Allen et al., 2001; Siniossoglou et al., 2000; Lutzmann et al., 2000) and tightly associates with Seh1 and Nup120 in the Nup84 sub-complex (Lutzmann et al., 2000).

1.8.4 Seh1

Seh1, Sec13 homolog 1, was identified as a component of the Nup84 sub-complex within the NPC and tightly associated with Nup85 in this sub-complex (Siniossoglou S, et al. 1996; Rout et al., 2000; Cronshaw et al., 2002; Lutzmann et al., 2000). Despite being larger than Sec13, Seh1 is highly homologous to Sec13 and also folds into a six-bladed β-

propeller based on the presence of six WD40 repeats (Deovs et al., 2004).

Although the Seh1 null mutant grows slowly, it is still viable, especially at low temperatures. Since Sec13 and Seh1 belong to the same sub-complex and contain a similar protein fold in the NPC, Sec13 was also proposed to be able to rescue the functions of Seh1 in yeast, once the Seh1 gene was knocked out (Siniossoglou, et al. 1996).

1.9 The evolution of the Nup84 sub-complex

The nucleus is one of the major characteristics in eukaryotic cells and is surrounded by the nuclear envelope (NE), contiguous with the endoplasmic reticulum (ER) to form a portion of the internal membrane systems. The complicated internal membrane system, which include the Golgi apparatus, ER, and NE, were proposed to develop from endomembrane systems in early eukaryotic cells by curving membranes (Blobel, 1980).

Based on the composition and functions of the vesicle-coated complexes, three major kinds of transport vesicles are responsible for macromolecule delivery between internal membranes and plasma membrane: COPI, delivering proteins from the *cis* end of the Golgi apparatus to the rough ER;

COPII, transporting proteins from the rough ER to the Golgi apparatus; and clathrin coat, responsible for endocytosis of material at the cell surface. Similar to the vesicle-coated complexes of COPI, COPII and clathrin, which bend the membrane in order to facilitate budding during vesicle trafficking, the Nup84 sub-complex was also proposed to bend the pore membrane domain (Devos et al., 2004).

Since the components in these four complexes, COPI, COPII, clathrin and the Nup84 sub-complex, share a similar molecular architecture, consisting of β-propeller and α-solenoid domains, and since they perform alike functions involved in membrane bending, they could have originated from a common ancestor, a protocoatomer (Devos et al., 2004) (Fig 17). The protocoatomer might have existed in early eukaryotic cells as a module that induced curvature in order to form the early endomembrane systems. During divergent evolution, the vesicle complexes and the Nup84 sub-complex would evolve from this protocoatomer in order to fulfill similar functions in diverse contexts.

In spite of the similarity of protein folds in these four complexes, the arrangement of α-solenoid and β-propeller domains in the COPII and clathrin complexes is different. The clathrin complex contains a polyhedral (soccer-ball-shaped) lattice made up of 36 triskelion

assembly units. The central trimerization domain of the triskelion forms the vertex of the surface and the three α-solenoid legs of the triskelion extend toward the three adjacent vertices. The β-propeller domain in the triskelion is not involved in the coat assembly but faces the inner shell for interaction with cargo adaptor proteins such as β-arrestin (Kirschhausen et al., 2000; Fotin et al., 2004) (Fig 18).

The architecture of the COPII cage has been observed as a cuboctahedron, which contains eight triangular faces and six square faces, and a rhombicosidodecahedron, which has 20 triangular faces, 30 square faces and 12 pentagonal faces (Stagg et al., 2008). Moreover, the β-propeller and α-solenoid domains of Sec31 are involved in forming the vertex and edge portions of the COPII cage, respectively (Fath et al., 2007) (Fig 19). Notably, another β-propeller protein, Sec13, was also suggested to play a crucial role in the assembly of the cage structure (Stagg et al., 2008). In addition, unlike the clathrin coat, which employs the β-propeller domain for cargo adapter protein binding, an un-structured fragment of Sec31, which is in the proline rich region and not supposed to be the architectural core, has been demonstrated to interact with a multivalent cargo recruitment complex, Sec23/Sec24 (Bi et al., 2007).

Therefore, due to architectural differences between the clathrin and COPII complexes, further structural studies of the Nup84 sub-complex and COPI are necessary in order to evaluate the evolutionary relationships among these complexes.

Another interesting fact is that the coated vesicle complexes cover a spherical surface, which is convex in all directions. However, there are two kinds of curvatures in the pore membrane domain, a convex curvature of the NE from the inside to the outside of the nucleus and a concave curvature surrounding the center of the membrane pore (Fig 20). Which curvature is selected and covered by the membrane-curving module, the Nup84 sub-complex, is still unclear. If the concave curvature is indeed the surface topology that the Nup84 sub-complex covers, another consideration refers to the question of how similar structural modules could be used in different surface topologies. Hence, more experimental testing has to be carried out to determine the evolutionary origin of the coated vesicle complexes and the Nup84 sub-complex.

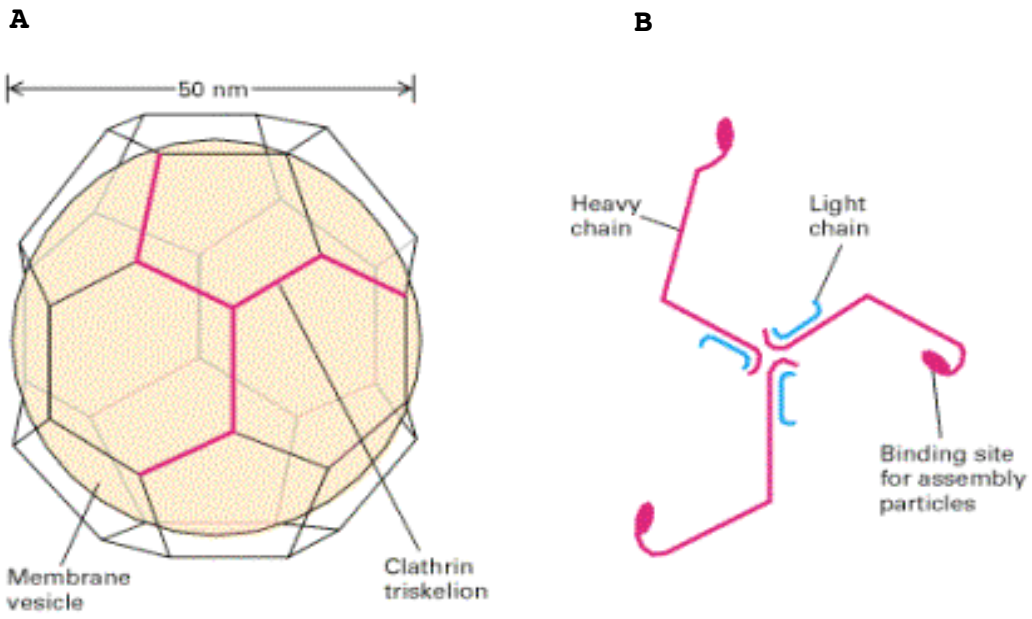


Figure 18. Structure of a clathrin-coated vesicle
(A) A typical clathrin coat is approximately 40 nm in diameter and, like a soccer ball, contains a polyhedral lattice with 12 pentagons and 8 hexagons. In total, the coat is constructed of 36 clathrin triskelions, one of which is shown in red. One clathrin triskelion is centered on each of the 36 vertices of the coat. (B) Detail of a clathrin triskelion structure. Each of the three clathrin heavy chains has a sharp bent structure. A clathrin light chain is attached to each heavy chain near the center; The N-terminal β-propeller domain is at each distal tip.

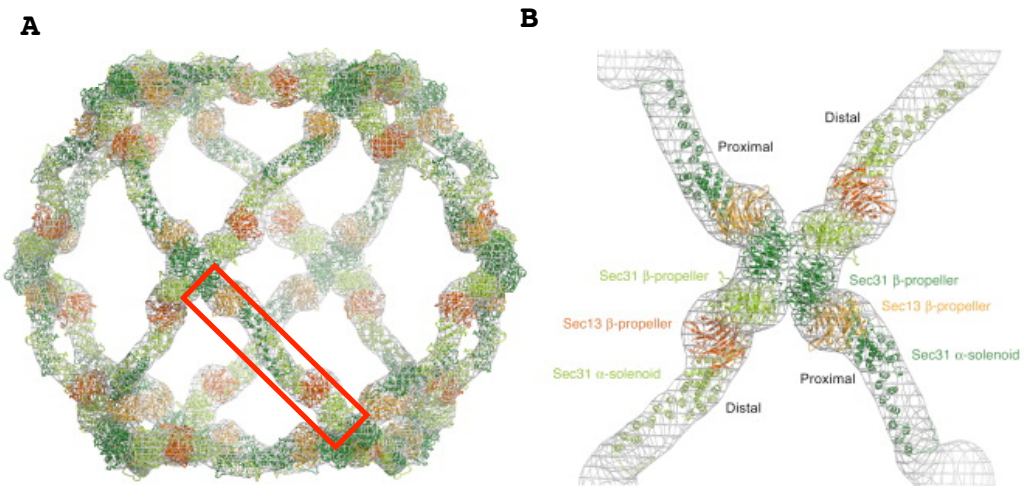


Figure 19. The organization of the COPII cuboctahedron cage

(A) A model of the COPII cage comprises 24 copies of the assembly unit with octahedral symmetry. The assembly unit, a rod-like structure including two pairs of Sec31/Sec13 complexes, is indicated by the rectangle. (B) The view of a complete vertex. One symmetry-related pair, which contains two pairs of Sec13/Sec31 complexes (colored dark green and orange), converges at the vertex and is labeled proximal; the other symmetry-related pair (light green and red) is labeled distal. Image taken from Fath et al., 2007.

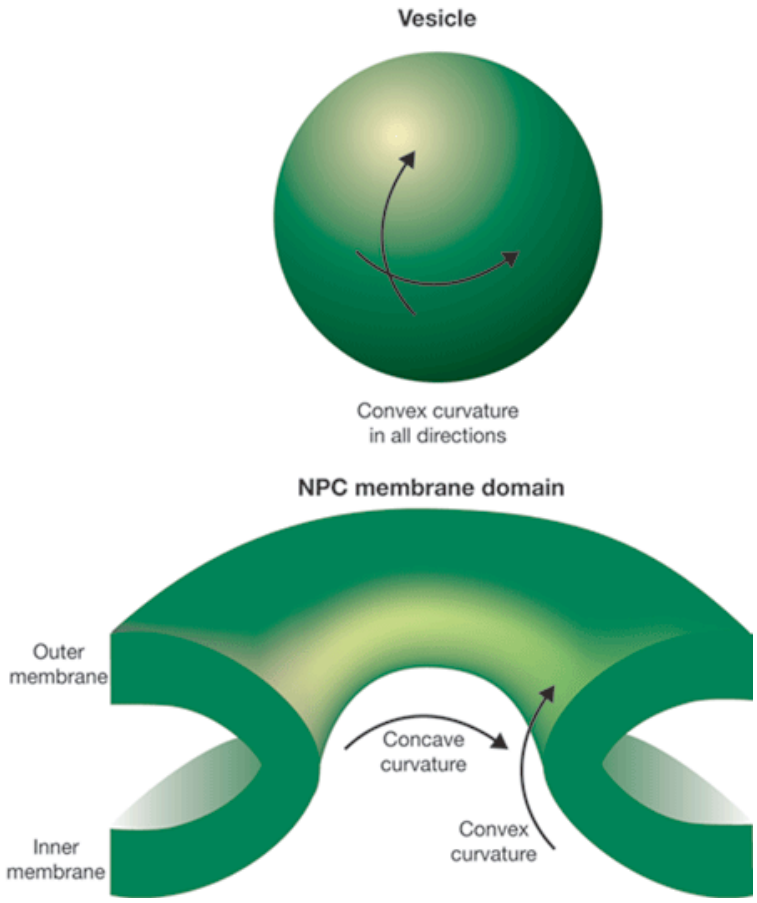


Figure 20. The different surface topologies in vesicles and nuclear pores

Like a sphere, the surface of a vesicle only possesses a convex curvature in all directions, whereas the nuclear pore membrane has two curvatures, a convex curvature of the NE from the inside to the outside of the nucleus and a concave curvature surrounding the center of the pore. Image taken from Antonin and Mattaj, 2005.

CHAPTER 2: MATERIALS AND METHODS

2.1 Bacterial strains, microbiological techniques, plasmids and DNA manipulations

DH5α (Invitrogen), BL21-CodonPlus (DE3)-RIL (Stratagene) are employed for DNA manipulation and heterologous protein expression respectively. DNA manipulations, such as restriction analysis, ligation, PCR amplification and plasmid isolation and so on, and microbiological techniques, such as competent cell preparation and transformation, are performed according to standard protocols.

2.2 Protein expression and purification

2.2.1 Sec13 variants and Nup145C/Sec13 complex

DNA fragments of yeast Nup145C (residue 1-711 and 125-555) were amplified by PCR and cloned into the pETDuet-1 (Novagen) expression vector (Table 1). The resulting Nup145C fusion proteins contained an N-terminal 6-His tag. The amplified PCR fragments of human or yeast Sec13 (residues 1-322 and 1-316 of hSec13 and residues 1-297 of ySec13) were cloned into pET24b (Novagene) for the expression of untagged proteins. The full-length hSec13

(residues 1-322) fragment was also cloned into a modified pET28a vector (Nogagene) that contained a PreScission protease site directly after the N-terminal 6-His tag (Hoelz et al., 2003).

For the expression of various Nup145C/Sec13 complexes, *E. coli* BL21-CodonPlus (DE3)-RIL cells were co-transformed with the appropriate expression vectors, using the untagged Sec13 variants. Protein expression was carried out in Louria-Bertani (LB) medium containing 100 μg/ml amplicillin, 50 μg/ml kanamycin and 34 μg/ml chloramphenicol, and induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Gold Biothechnology) at 18°C for 16 hours. Cells were harvested by centrifugation and resuspended in a buffer containing 50 mM K-phosphate, pH 7.4, 150 mM NaCl, 5 mM β-mercaptoethanol and protease inhibitor (Roche). The cells were lysed with a cell disruptor (Avestin) and the lysate was centrifuged for 90 min at 40,000g. The crude lysate was then applied to a His-Select nickel column (Sigma) and eluted via an imidazole gradient. Fractions containing the Nup145C/Sec13 complex were pooled and dialyzed against a buffer containing 20 mM HEPES, pH7.4, 50 mM NaCl and 3 mM DTT. The protein was purified over a HiTrap Q HP column (GE Healthcare) via a NaCl gradient and further purified over a 16/60 Superdex

200 column (GE Healthcare). Fractions containing the pure Nup145C/Sec13 complex were pooled and concentrated to 12 mg/ml for crystallization (Fig 21).

Expression and purification of Sec13 protein were carried out using an identical expression and purification protocol. The N-terminal 6-His tag of Sec13 was removed by PreScssion protease cleavage for 36 hours.

A

Protein	Species	Residues	Expression vector	Restriction sites 5′, 3′	Tag
Nup145C	yeast	1-711	pETDuet-1	BamHI, NotI	HIS
Nup145C	yeast	125-555	pETDuet-1	BamHI, NotI	HIS
Sec13	yeast	1-297	pET24b	NdeI, XhoI	No tag
Sec13	human	1-316	pET24b	NdeI, XhoI	No tag
Sec13	human	1-322	pET24b	NdeI, XhoI	No tag
Sec13	Human	1-316	PET28-pp	NdeI, XhoI	HIS
Nup145C/Sec13	yeast yeast	1-316(Sec13) 125-555(Nup145C)	pETDuet-1	BamHI, NotI NdeI, XhoI	HIS No tag
Nup145C/Sec13	yeast yeast	1-316(Sec13) 1-777(Nup145C)	pETDuet1& pCDFDuet-1	BamHI, NotI NdeI, XhoI	HIS No tag
Nup85/Seh1	yeast yeast	1-349(Seh1) 1-570(Nup85)	pETDuet-1& pRSFDuet-1	BamHI, NotI NdeI, XhoI	HIS No tag
Nup85/Seh1	yeast yeast	1-349(Seh1) 1-744(Nup85)	pETDuet-1& pRSFDuet-1	BamHI, NotI NdeI, XhoI	HIS No tag
Nup85	yeast	566-744	pRSFDuet-1	BamHI, NotI	HIS
Nup120	yeast	1-1037	pET8c	NheI, BamHI	HIS
Sec31	yeast	370-763	pETDuet-1	BamHI, NotI	HIS

B

Protein	Species	Residues	Expression vector	Restriction sites 5′, 3′	Tag
Sec13	yeast	1-297	pRS414	BamHI, EcoRI	No tag
Sec13	human	1-316	pRS414	BamHI, EcoRI	No tag
Sec13	human	1-322	pRS414	BamHI, EcoRI	No tag
Seh1	Yeast	1-347	pRS414	BamHI, EcoRI	No tag

Table 1: Expression constructs (A) Bacteria over-expression constructs (B) Yeast expression constructs.

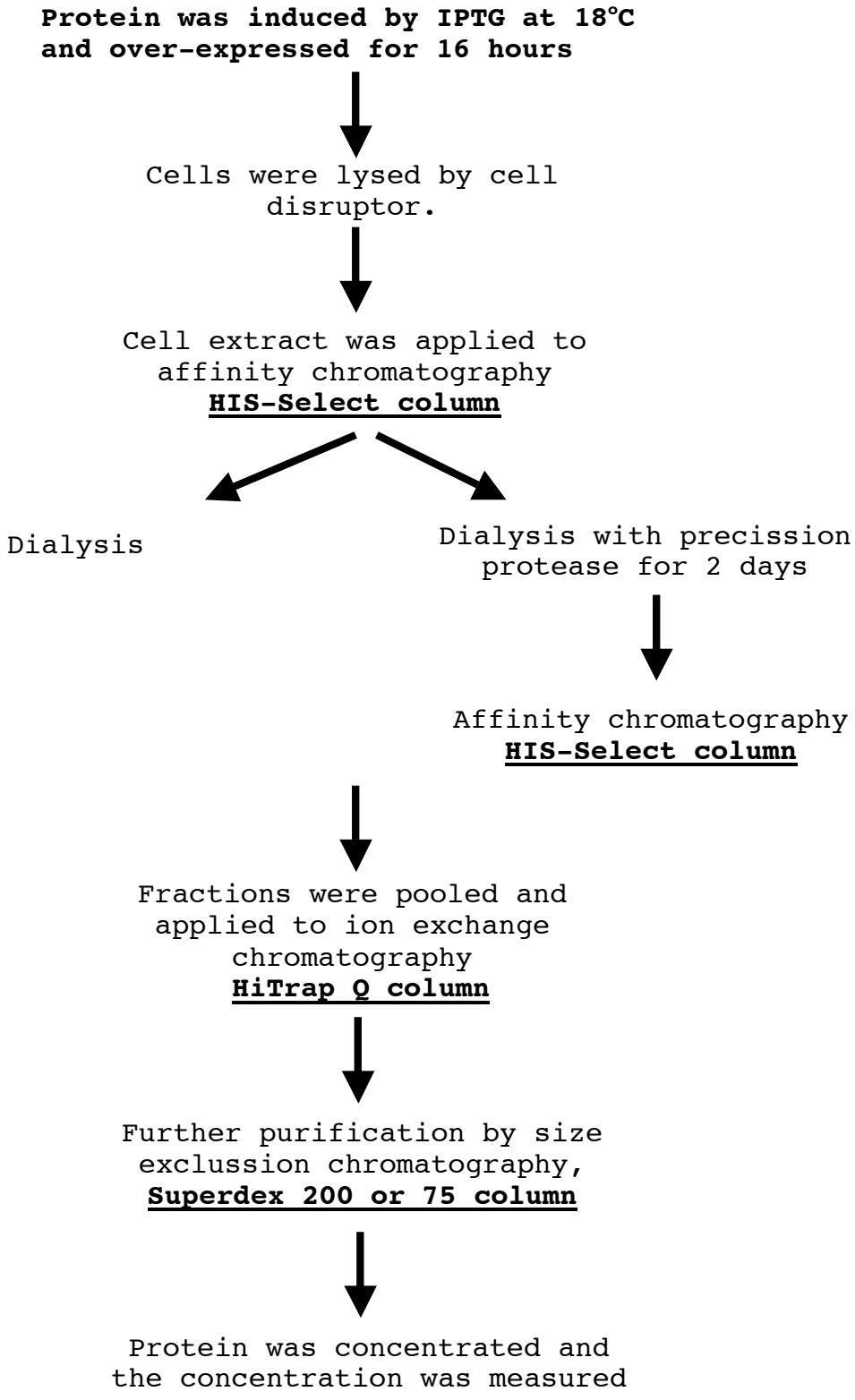


Figure 21: Protein purification flowchart

2.2.2 Other nucleoporins and Sec31

The coding regions of full-length yeast Seh1 (residues 1-349) and Nup85 (residues 1-570) were cloned into the two multiple cloning sites of the bacterial bicistronic expression vectors, pETDuet-1 and pRSFDuet-1 (Novagene) as well as Seh1 and full-length Nup85 (residues 1-744). The C-terminus yeast Nup85 (residues 566-744) fragment was also cloned into the pRSFDuet-1 expression vector. The solenoid portion of yeast Sec31 (residues 370-763) was cloned into the expression vector pETDuet-1. Full-length yeast Nup120 in pET8c vector is a gift from Dr. Ed Hurt.

Except for the full-length Nup120/Nup85$^{566-744}$ complex, other proteins or complexes were expressed and purified using a similar protocol. In the case of over-expressing full-length Nup120/ Nup85$^{566-744}$ complex, terrific broth (TB) medium was used instead of LB medium to increase the protein expression level.

2.3 Western blot analysis

Concentration of protein samples was determined by the Lowry method (Bio-Rad). Proteins were fractionated on 12.5% SDS-PAGE and then transferred onto PVDF membrane (Millipore) by semi-dry electrophoretic transfer cell (Bio-Rad). Membranes were further probed with primary

antibodies, penta-histidine antibody (Qiagen) or anti-glutathione-s-transferase (GST) antibody (GE Healthcare). After incubation with appropriate secondary antibodies (GE Healthcare), proteins were visualized by chemiluminescence according to the manufacturer's instructions (Pierce).

2.4 Limited proteolysis

A total 50 μg full-length Nup145C/Sec13 protein complex was mixed separately with 1 μg protease, such as trypsin (Roche), elastase (Sigma) and papain (Roche), at room temperature at various times. Edman sequencing and mass spectrometer then were employed to determine the protein fragments generated by protease digestion.

2.5 Crystallization and X-ray data collection

Nup145C/hSec13$^{1-316}$ (12 mg/ml) were grown at 21°C in hanging drops containing 1 μl of the protein and 1 μl of a reservoir solution containing of 16% (w/v) PEG 3350, 400mM NaCl, 100mM Na-K tartrate, and 100mM BIS-TRIS, pH7.7. Crystals belonged to the orthorhombic space group $C222_1$ and the monoclinic space group C2. Both crystal forms contain a Nup145C/Sec13 hetero-octamer in the asymmetric unit according to the Matthew coefficient calculation. They grew to their maximum size, 1000 x 300 x 200 μm within 16 days.

For cryo-protection, crystals were stabilized in 16% (w/v) PEG 3350, 400mM NaCl, 100mM Na-K tartrate, 100mM BIS-TRIS, pH 7.7, and 22% (v/v) glecrol and flash frozen in liquid nitrogen-cooled liquid propane. X-ray diffraction data were collected at the Advanced Light Source (ALS) and the National Synchrotron Light Source (NSLS). X-ray intensities were processed using HKL2000 (Otwinowski and Minor, 1997), and the CCP4 (CCP4, 1994) program package was used for subsequent calculations.

2.6 Structure determination and refinement

Initial phase was determined using a $[Ta_6Br_{12}]^{2+}$ cluster derivative and MAD measurements of the orthorhombic crystals. These phases were used to locate the heavy atom sites of sodium ethylmercurithiosalicylate (EMTS), potassium osmate (K_2OsO_4), and di-potassium tetrachloroplatinate (K_2PtCl_4)-derivatized and SeMet-labelled protein crystals (Fig 22). Combined phasing using isomorphous EMTS, K_2OsO_4, and SeMet MAD and native datasets was carried out in SHARP, followed by density modification in DM, with solvent flattening and histogram matching. This procedure yielded an interpretable electron density map of high quality (Fig 23). Phasing of the X-ray diffraction data derived from the monoclinic crystals was carried out

independently with a similar phasing protocol, yielding an electron density map with comparable quality.

The initial model was built into the electron density map of the orthorhombic crystal by using O (Jones et al., 1991) and refined using CNS (Brunger et al., 1998). The final model was refined to a 3.0 Å resolution with an R_{work} of 24.7% and an R_{free} of 28.9%. This final model was docked into the experimental electron density map of the monoclinic crystal form and refined using CNS. The final model was refined to a 3.15 Å resolution with an R_{work} of 24.7% and an R_{free} of 29.2%. The stereochemical quality of both final models was assessed with PROCHECK (Laskowski et al., 1993) There are no residues in the disallowed region of the Ramachandran plot (Fig 24 and 25). Data collection and refinement statistics are shown Table 2 and 3.

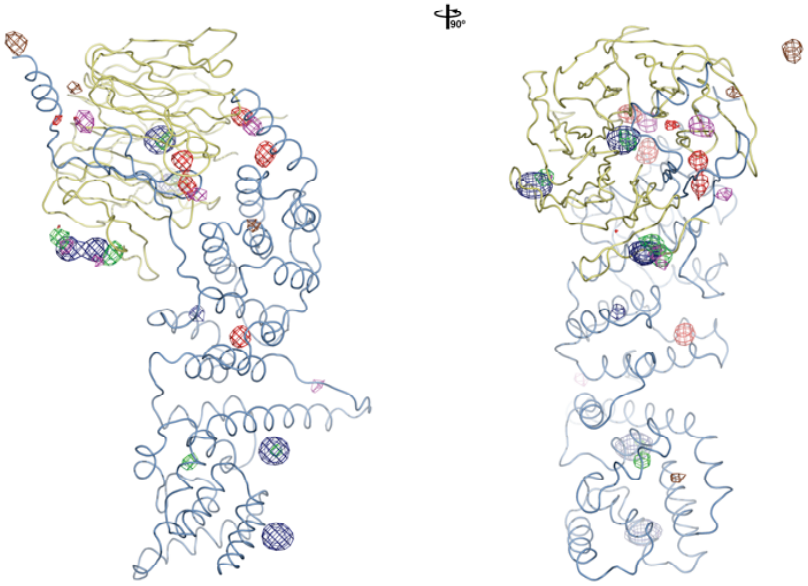


Figure 22. Experimental phasing

Anomalous difference Fourier maps, illustrating selenium (red), osmium (blue), mercury (green), platinum (magenta) and $[Ta_6Br_{12}]^{2+}$ cluster (orange) sites, calculated from x-ray diffraction data obtained from the orthorhombic crystal form. For charity, only one of the four Nup145C/Sec13 complexes in the asymmetric unit is shown (Image taken from Hsia et al., 2007).

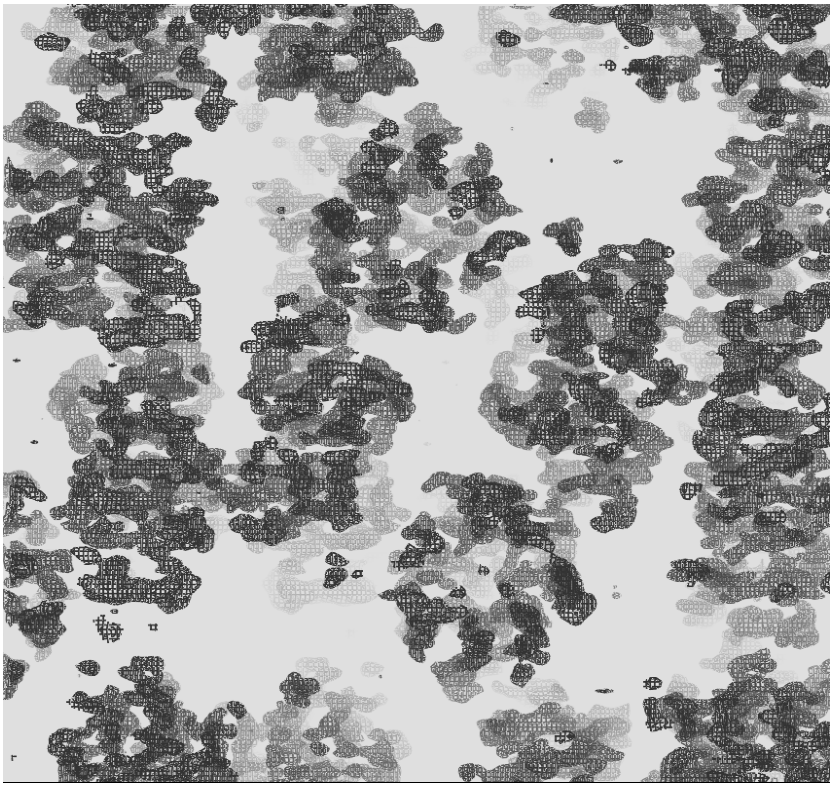


Figure 23. Experimental electron density map
A 3.0 Å resolution experimental electron density map (contoured at 1.5 σ) of the orthorhombic crystal form was calculated by using multiple anomalous dispersion (MAD) following by density modification.

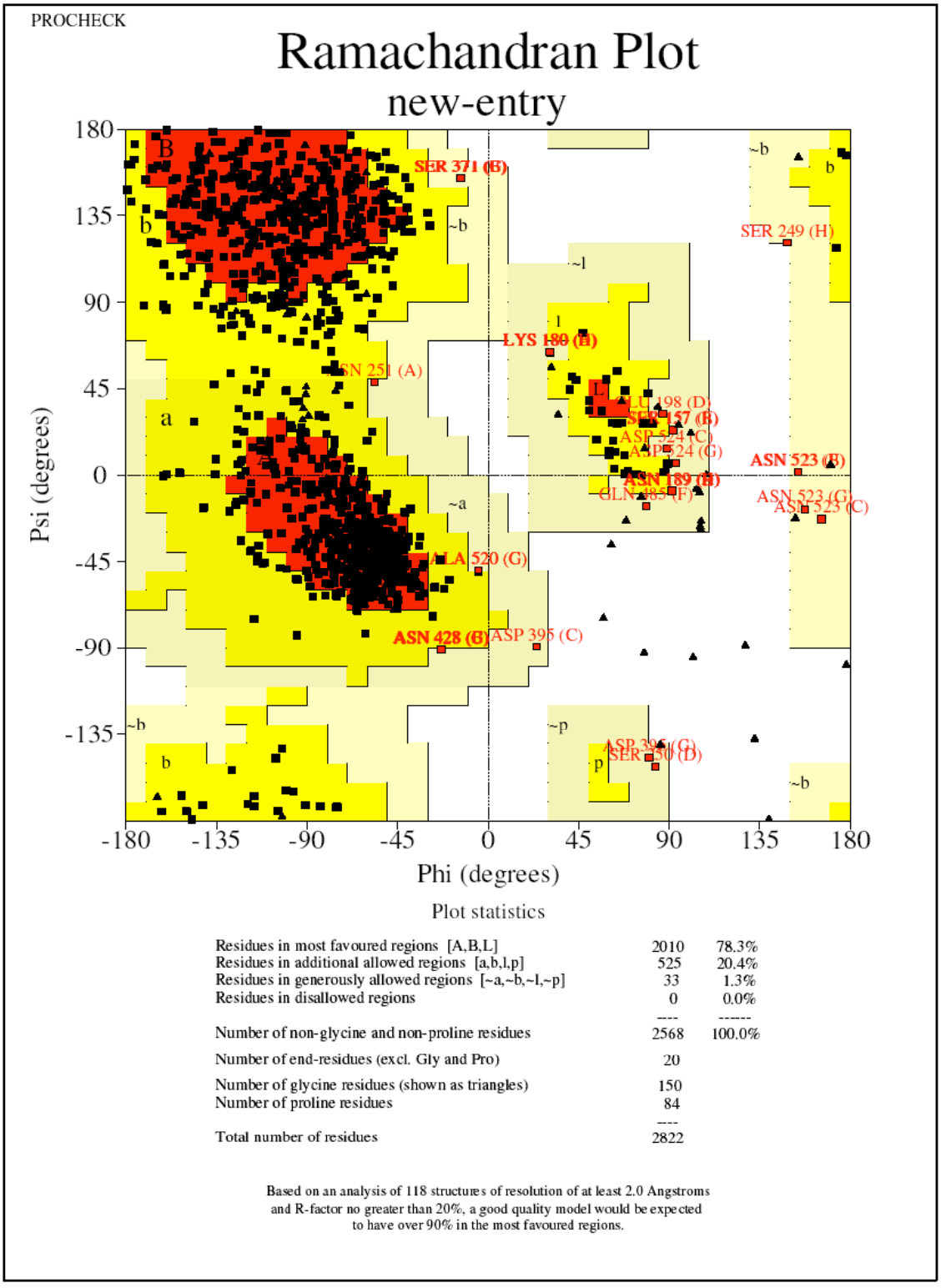
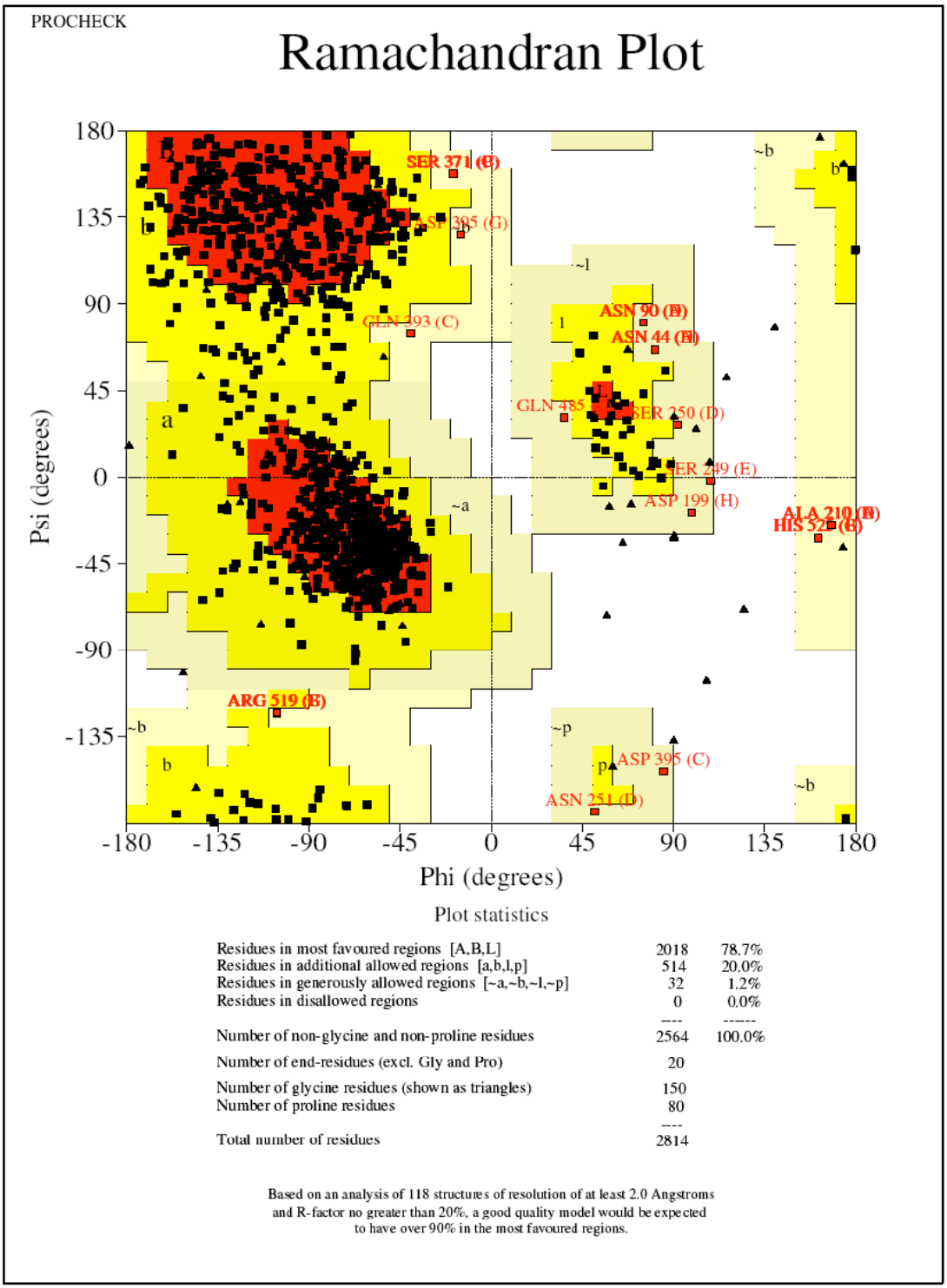


Figure 24. Ramachandran plot for the orthorhombic crystal form of the Nup145/Sec13 complex



Plot statistics

Residues in most favoured regions [A,B,L]	2018	78.7%
Residues in additional allowed regions [a,b,l,p]	514	20.0%
Residues in generously allowed regions [~a,~b,~l,~p]	32	1.2%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	2564	100.0%
Number of end-residues (excl. Gly and Pro)	20	
Number of glycine residues (shown as triangles)	150	
Number of proline residues	80	
Total number of residues	2814	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 25. Ramachandran plot for the monoclinic crystal form of the Nup145/Sec13 complex

	Crystal 1 Native	Crystal 2 $[Ta_6Br_{12}]^{2+}$	Crystal 3 SeMet			Crystal 4 EMTS[a]	Crystal 5 $KOsO_4$			Crystal 6 K_2PtCl_4
Data collection										
Synchrotron	ALS	ALS	NSLS	NSLS	NSLS	ALS	ALS	ALS	ALS	ALS
Beamline	BL8.2.2	BL8.2.1	X29	X29	X29	BL8.2.2	BL8.2.1	BL8.2.1	BL8.2.1	BL8.2.1
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
Cell dimensions										
a, b, c (Å)	a=181.4,	a=180.8,	a=180.7,	a=181.0,	a=181.9,	a=180.6,	a=181.0,	a=181.0,	a=180.5,	a=179.9,
	b=216.8,	b=216.5,	b=216.5,	b=217.3,	b=217.4,	b=217.0,	b=217.1,	b=217.6,	b=217.5,	b=216.1,
	c=192.6	c=193.4	c=192.9	c=193.6	c=194.2	c=191.6	c=193.5	c=193.9	c=194.0	c=187.3
α, β, γ (^{o})	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=90^0$	$\alpha=\beta=\gamma=9$
		Ta Peak	Se Peak	Se Inflection	Se Remote	Hg Peak	Os Peak	Os Inflection	Os Remote	Pt Peak
Wavelength	1.00000	1.25490	0.97900	0.97930	0.95000	1.00910	1.14010	1.14050	1.10700	1.07210
Resolution (Å)	20.0-3.0	50.0-4.0	30.0-3.4	30.0-3.65	30.0-3.8	50.0-3.9	50.0-3.8	50.0-3.95	50.0-4.45	50.0-4.7
R_{sym} (%)[b]	11.0(64.6)	11.2(55.2)	13.6(66.4)	14.6(72.3)	14.5(69.2)	8.8(60.1)	10.4(47.7)	12.1(65.5)	10.8(82.8)	9.8(79.4)
$<1/\sigma>$[b]	19.3(1.7)	11.6(2.2)	9.0(1.7)	8.6(1.7)	7.5(1.8)	12.5(2.3)	11.8(2.9)	10.2(2.0)	10.3(1.7)	13.3(1.3)
Completeness (%)[b]	97.3(92.4)	97.6(97.7)	99.1(99.1)	98.8(99.0)	95.2(93.1)	91.8(90.2)	99.4(96.6)	99.4(97.8)	99.4(97.9)	94.8(86.8
Redundancy	11.2	3.1	3.3	3.2	2.8	3.2	3.7	3.7	3.7	5.0

Refinement	
Resolution (Å)	20-3.0
No. reflections	53,141
Test set	4,327(8%)
R_{work}/R_{free} (%)	24.7/28.9
No. atoms	22,618
R.m.s. deviations	
Bond length (Å)	0.009
Bond angles (^{o})	1.6

[a]EMTS, ethylmercurithiosalicylate
[b]Highest-resolution shell is shown in parenttheses

Table 2: Data collection and refinement statistics table of orthorhombic crystals

	Crystal 1 Native	Crystal 2 $[Ta_6Br_{12}]^{2+}$			Crystal 3 $KOsO_4$		Crystal 4 EMTS[a]		Crystal 5 SeMet
Data collection									
Synchrotron	ALS	ALS	ALS	ALS	NSLS	NSLS	ALS	ALS	ALS
Beamline	BL8.2.2	BL8.2.1	BL8.2.1	BL8.2.1	X29	X29	BL8.2.1	BL8.2.1	BL8.2.2
Space group	C2	C2	C2	C2	C2	C2	C2	C2	C2
Cell dimensions									
a, b, c (Å)	a=181.8,	a=180.7,	a=180.8,	a=181.4,	a=181.1,	a=181.2,	a=181.1,	a=181.4,	a=181.7,
	b=216.6,	b=216.9,	b=217.5,	b=218.0,	b=218.0,	b=217.3,	b=217.4,	b=218.0,	b=217.2,
	c=101.0	c=101.0	c=101.5	c=101.5	c=101.2	c=100.8	c=102.1	c=102.3	c=101.7
a, b, g (°)	$\alpha=\gamma=90^0$ $\beta=108.1^o$	$\alpha=\gamma=90^0$ $\beta=108.2^o$	$\alpha=\gamma=90^0$ $\beta=108.2^o$	$\alpha=\gamma=90^0$ $\beta=108.2^o$	$\alpha=\gamma=90^0$ $\beta=108.0^o$	$\alpha=\gamma=90^0$ $\beta=107.9^o$	$\alpha=\gamma=90^0$ $\beta=107.9^o$	$\alpha=\gamma=90^0$ $\beta=107.9^o$	$\alpha=\gamma=90^0$ $\beta=108.0^o$
		Ta Peak	Ta Inflection	Br Peak	Os Peak	Os inflection	Hg Peak	Hg Inflection	Se Peak
Wavelength	1.00000	1.25480	1.25490	0.91980	1.14000	1.14040	1.00910	1.00930	0.97950
Resolution (Å)	20.0-3.15	20.0-5.0	50.0-5.0	20.0-6.5	20.0-3.75	20.0-3.75	20.0-3.6	20.0-4.0	50.0-4.0
R_{sym} (%)[b]	14.5(74.6)	8.2(40.9)	7.5(42.8)	9.1(49.8)	14.2(83.4)	11.6(59.2)	9.9(90.7)	6.6(31.1)	7.5(31.6)
<1/s>[b]	17.9(1.9)	10.4(1.8)	9.8(1.8)	10.6(1.6)	9.6(1.4)	12.2(1.4)	8.0(2.1)	12.1(2.0)	10.6(2.2)
Completeness (%)[b]	95.7(77.7)	96.5(96.2)	96.8(93.7)	98.1(96.1)	98.7(93.4)	97.5(89,4)	97.3(89.7)	97.1(89.5)	80.2(49.5)
Redundancy	10.8	1.9	1.9	1.9	3.5	3.4	1.8	1.7	1.8

Refinement	
Resolution (Å)	20-3.15
No. reflections	55,955
Test set	2,837(5%)
R_{work}/R_{free} (%)	24.7/29.2
No. atoms	22,566
R.m.s. deviations	
Bond length (Å)	0.009
Bond angles (°)	1.6

[a]EMTS, ethylmercurithiosalicylate

[b]Highest-resolution shell is shown in parenttheses

Table 3: Data collection and refinement statistics table of monoclinic crystals

2.7 Multi-angle light scattering

Protein at various concentrations was injected onto a Superdex 200 HR 10/300 column equilibrated in a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM DTT. The chromatography system was coupled to an 18-angle light scattering detector (DAWN HELEOS) and a refractive index detector (Optiab rEX) (Wyatt Technology). Data were collected every 0.5 s at a flow rate of 0.4 ml/min. Data analysis was carried out using the program ASTRA, yielding the molar mass and mass distribution (polydispersity) of the sample.

2.8 Illustrations and figures

Figures were generated using PyMOL (www.pymol.org). The molecular surfaces were calculated using MSMS (Sanner et al., 1996) and the electrostatic potential was calculated using APBS (Baker et al., 2001). Sequence alignments were generated using ClustalX (Jeanmougin et al., 1998) and colored with Alscript (Barton, 1993).

2.9 Plasmid shuffle analysis

The *Saccharomyces cerevisiae* Sec13 shuffle strain, *sec13::HIS3* (pRS316-URA3-SEC13-Myc), in the RS453 background, *MATa,ade2,his3,leu2,trp1,ura3*, was provided by

Dr. Ed Hurt. Full-length human Sec13$^{1-322}$, tail-less human S Sec13$^{1-316}$ and yeast Sec13 were cloned into yeast expression vector, pRS414 (Stratagene), which contains a constitutive promoter from alcohol dehydrogenase 1 and a TRP1 maker gene. These constructs were transformed into the yeast Sec13 shuffle strain and then cells were spread on ura$^-$ trp$^-$ medium in order to maintain two plasmids, pRS316-URA-ySec13 and pRS414-TRP1-"Sec13 variants". Once transformants were apparent, they were picked and further transferred to 5-fluoroorotic acid (5-FOA) (American bioanalytica) containing medium for selecting against pRS316-URA3-Sec13. Different transformants were spotted at 10 fold serial dilutions (from ~10^5 to 10) on plates of YPD or minimum medium containing 5-FOA.

2.10 Reconstitution of the Nup84 sub-complex

Full-length proteins of Nup120, Nup145C/Sec13 complex and Nup85/Seh1 complex were purified separately based on the protocols used for the Nup145C/Sec13 complex. In order to in vitro assemble the Nup84 sub-complex, these five proteins were mixed in approximately euqimolar amounts and dialyzed against a buffer containing 25 mM HEPES, pH7.5, 300 mM NaCl, 150 mM K-acetate, 2 mM Mg_2-acetate and 3 mM

DTT. After 16-hour incubation, the mixture was applied to the Superose 6 10/300 column (GE Healthcare) and the fractions were analyzed by SDS-PAGE.

CHAPTER 3: RESULTS

3.1 Characterization of the full-length Nup145C/Sec13 complex

In yeast, Nup145C is generated by the autoproteolytic cleavage of a nascent precursor polypeptide, yielding Nup145N and Nup145C (Nup98 and Nup96 in vertebrates) (Fontoura et al., 1999; Teixeira et al., 1997). According to secondary structure predictions, the Nup145C polypeptide chain (residues 1 to 711) can be divided into two regions: an N-terminal, 125-residue region that lacks any apparent structural elements followed by a 586-residue region that has been predicted to fold into an α-helical solenoid domain (Fig 26). Yeast Sec13 contains six WD40 repeats that form a β-propeller (Fath et al., 2007; Saxena et al., 1996)

For the purpose of expressing the full-length yeast Nup145C and Sec13 protein complex, the genes of these two proteins were cloned into bacterial expression vectors, pETDuet-1 and pET24b, and then co-transformed into BL21-codon plus (DE3) RIL cells. Furthermore, according to the characterization of Nup145C/Sec13 complex, appropriate affinity, ion exchange and gel filtration chromatography steps were chosen to purify the protein complex. Eventually, the purity was analyzed by SDS-PAGE (Fig 27)

and the yields of the complex were measured spectrophotometrically. The final yield was ~2 mg per liter of culture. The protein was then concentrated and set up for crystallization. By using various crystal screen kits, however, no initial hits or crystals were observed in the initial conditions.

The native conformation of protein provides some steric hindrance to enzymatic attack of buried regions, but leaving the exposed and flexible regions accessible to proteases. Furthermore, flexible regions of proteins and protein degradation are two unfavorable factors that could highly disturb protein packing, thus hampering crystal formation. Therefore, limited proteolysis was employed to remove unstructured or flexible regions and to further avoid protein degradation by defining the most compact and stable regions of the Nup145C/Sec13 complex.

Full-length Nup145C/Sec13 was incubated with trypsin, elastase, and papain individually for various lengths of time. Interestingly, a stable complex was obtained after 30 minutes of elastase treatment (Fig 28). In order to identify the cleavage sites of the elastase-treated complex that contains two predominant protein bands, the two protein fragments were isolated from SDS-PAGE gels and PVDF membranes and analyzed by Edman sequencing and mass

spectrometry. As a result, the upper fragment was identified to be a portion of Nup145C and, surprisingly, included an unstructured region of 125-183 residues at the N-terminus and two-thirds of the α-solenoid portion; on the other hand, the lower fragment was identified as intact Sec13 even after protease digestion.

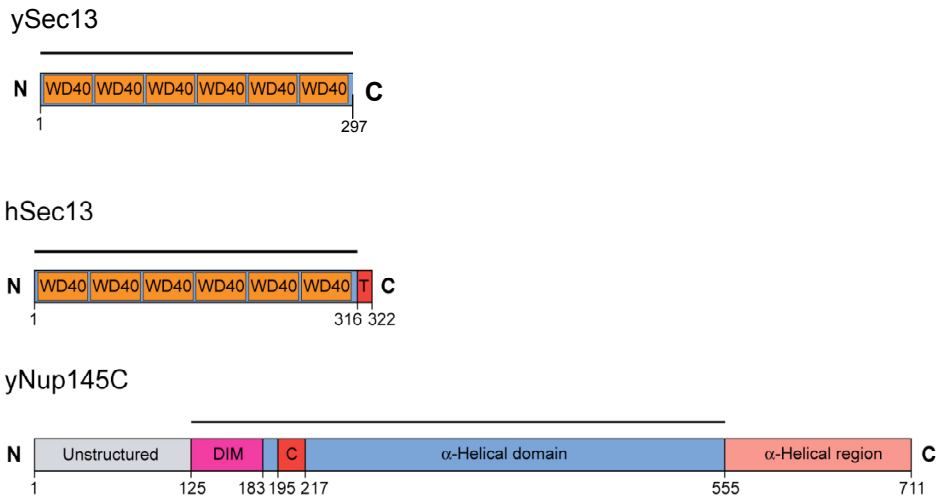


Figure 26. Domain structures of yeast and human Sec13 and yeast Nup145C

For Sec13, the WD40 repeats (orange), the C-terminal tail (T) (red), and the numbering relative to yeast and human Sec13 are indicated. For Nup145C, the unstructured N-terminal region (gray), the domain invasion motif (DIM) (magenta), the αB-αC connector (C)(red), the α-helical domain (blue), and the C-terminal α-helical region (light pink) are indicated. The numbering is relative to yeast Nup145C. The bars above the domain structures of the proteins mark the crystallized fragments.

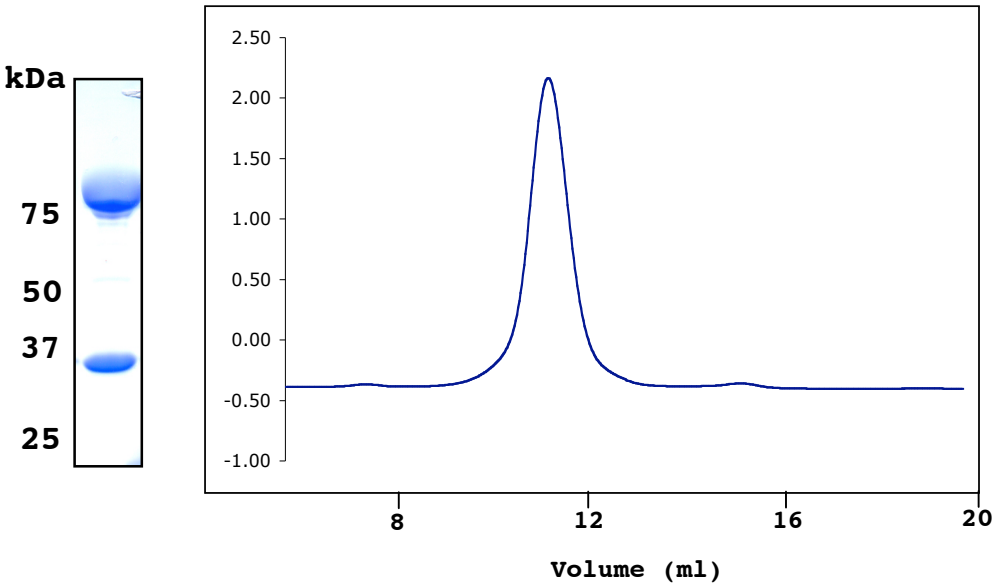


Figure 27. Purification of the yeast full-length Nup145C/Sec13 complex

Gel filtration profile of full-length yeast Nup145C/Sec13 on a Superdex 200HR 10/300 column. The peak fractions were pooled and analyzed by SDS-PAGE. The SDS-PAGE was stained with Coomassive brilliant blue.

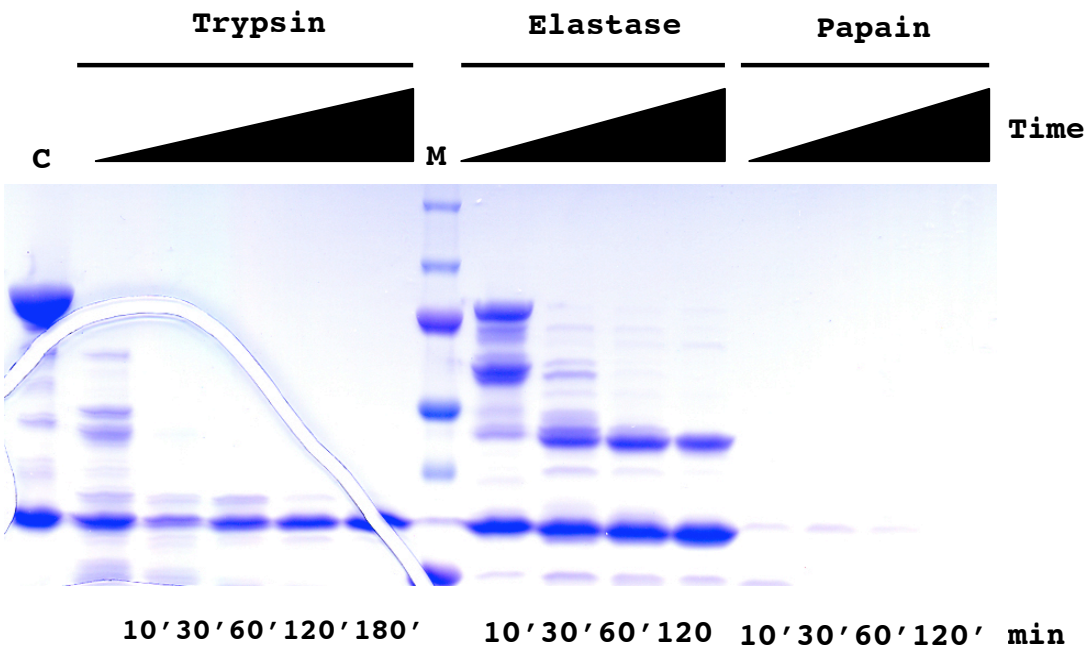


Figure 28. Limited proteolysis of full-length Nup145C/Sec13

Full-length Nup145C/Sec13 was incubated with trypsin, elastase, and papain individually at room temperature at various times. After 30 minutes, a stable complex was generated by elastase digestion. C represents the control, the full-length Nup145C/Sec13 complex without protease treatment.

3.2 Proteomic and crystallographic studies of various Nup145C$^{125-555}$/Sec13 complexes

3.2.1 Purification, crystallization, and preliminary x-ray analysis of the truncated Nup145C$^{125-555}$/yeast Sec13 complex

Based on the proteolysis data, Edman sequencing, and mass spectrometry of the full-length Nup145C/Sec13 complex, a short version of the Nup145C, Nup145$^{125-555}$, was cloned and co-expressed with full-length yeast Sec13. The purification procedure of this complex is the same to the one described in the materials and methods section. Nup145$^{125-555}$ and ySec13 formed a ~80 kDa protein complex (Fig 29). The final yield of the purified complex after all purification steps was very good with ~15 mg per liter of bacterial culture.

The Nup145$^{125-555}$/ySec13 complex was then screened for crystallization using commercially available screens. Several lead conditions were identified and they were improved by using grid screens around the original crystallization conditions. The final crystallization condition that yielded diffraction-quality crystals that subsequently used for structure determination was 0.1 M BIS-TRIS, pH 6.5, 0.2 M tri-potassium citrate monohydrate, 15% PEG3350, 0.35 M guanidine-HCl and 0.7 M potassium chloride as the reservoir solution. Crystals typically

appeared within a few days and reached their maximum size for harvesting after approximately one to two weeks. The presence of both $Nup145^{125-555}$ and ySec13 proteins was confirmed by SDS-PAGE analysis of large crystals that were dissolved after washing steps (Fig 30C).

For X-ray diffraction studies, crystals were cryo-protected by soaking in 0.1 M BIS-TRIS, pH 6.5, 0.2 M tri-potassium citrate monohydrate, 18% PEG3350, 0.35 M guanidine-HCl, 0.7 M potassium chloride, and 22% (v/v) glycerol. Cryo-protected crystals were then flash frozen in liquid nitrogen-cooled liquid propane (Fig 28A). X-ray diffraction data were collected on beamline BL8.2.2 at the Advanced Light Source (ALS) and on beamline X9 at the National Synchrotron Light Source (NSLS) (Fig 30B). Space group and unit cell dimensions of the crystal were determined using the program HKL2000. The crystals appeared to contain various large unit cells with a common long axis (750 Å). One of the potential space groups was the monoclinic space group C2 with unit cell dimensions of a = 750 Å, b = 180 Å, c = 70 Å, and $\beta = 95^{\circ}$. Furthermore, based on the Matthew coefficient calculation, the asymmetric unit in this monoclinic crystal could contain more than ten complexes, indicating more than 1 MDa in the asymmetric

unit. Due to the huge unit cell content, the difficulty of solving this structure by de-novo phasing is immense.

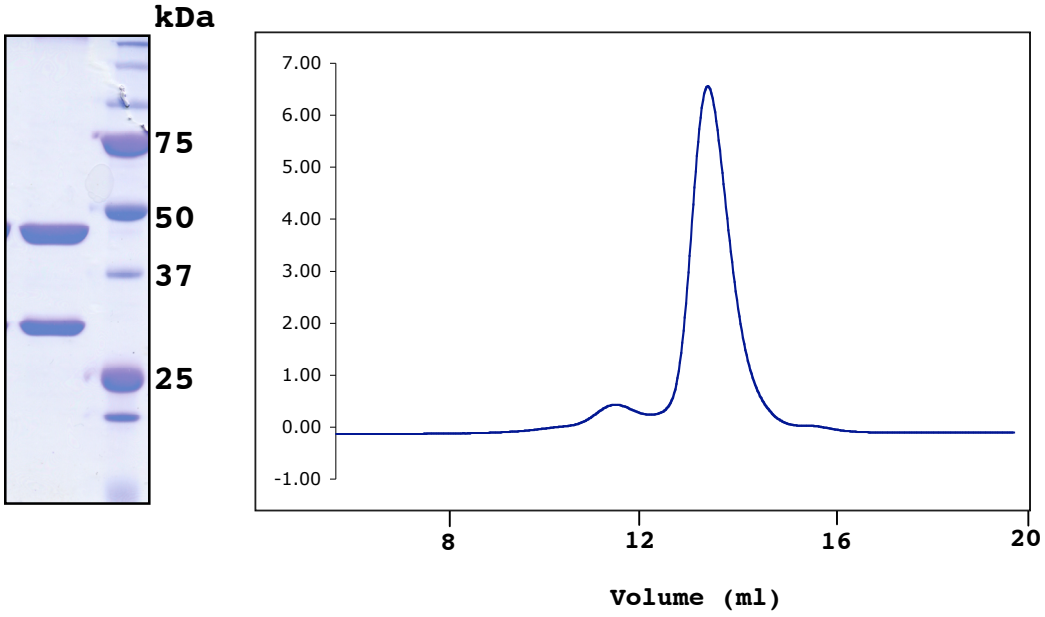


Figure 29. Purification of the truncated Nup145C$^{125-555}$/Sec13 complex

Gel filtration profile of the truncated Nup145C$^{125-555}$/Sec13 on a Superdex 200HR 10/300 column. The principal peak fractions were pooled and analyzed by SDS-PAGE. The SDS-PAGE was stained with Coomassive brilliant blue.

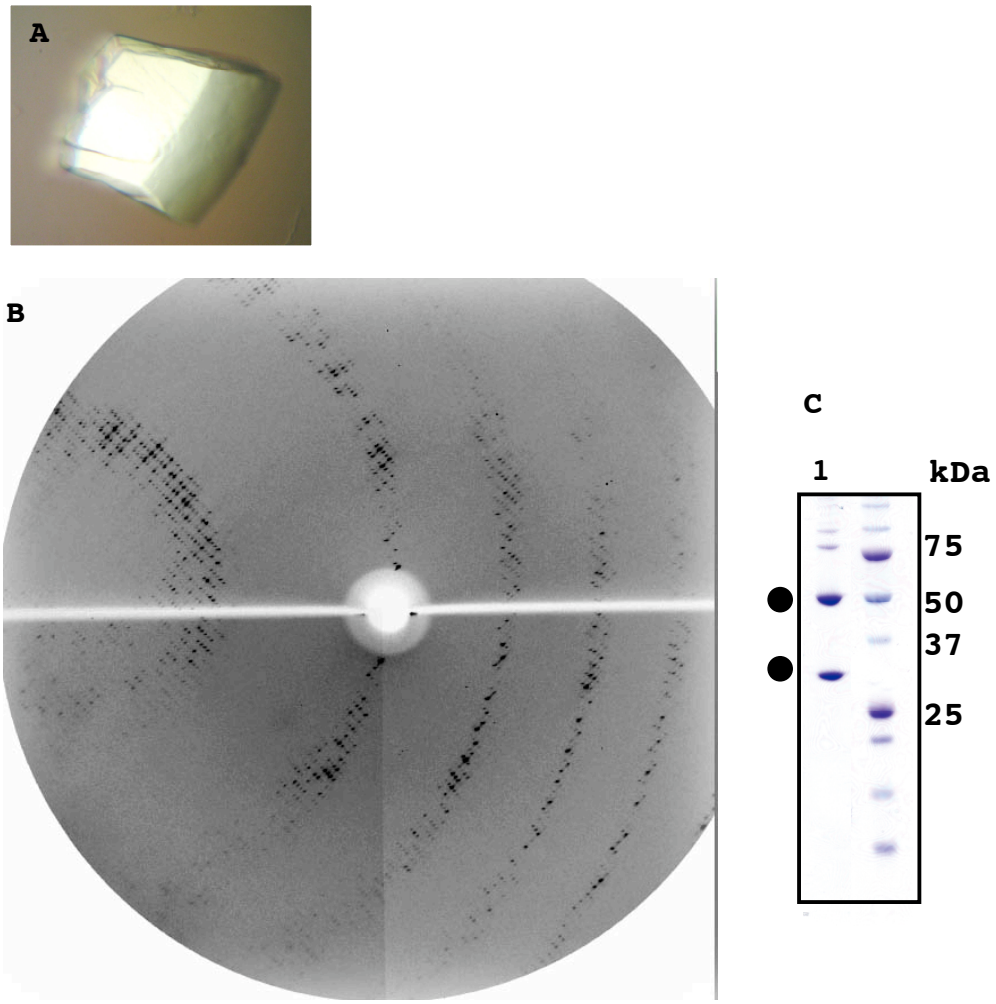


Figure 30. Crystal and typical diffraction image of the truncated Nup145C$^{125-555}$/Sec13 complex

(A) Crystal of the Nup145C$^{125-555}$/Sec13 complex. (B) X-ray diffraction pattern of this complex obtained on beamline X9 at the NLSL. The unit cell dimensions of the crystal are a=750 Å, b=180 Å, c=80 Å; $\alpha=\beta=\gamma=90^{\circ}$. (C) Crystals were dissolved and analyzed by SDS-PAGE. The black spots indicate the two proteins, Nup145C$^{125-55}$ and Sec13.

3.2.2 Purification and analyses of yeast Nup145C/human Sec13 complexes

According to protein sequence alignment analysis, yeast Sec13 shows ~74% similarity and ~48% identity to human Sec13. Thus, for the purpose of getting smaller unit cell crystals, full-length human $Sec13^{1-322}$ was utilized and co-expressed with yeast full-length Nup145C to see whether or not these two proteins would form a complex.

Indeed, full-length Nup145C and human $Sec13^{1-322}$ formed a chimeric complex. In order to probe the tertiary protein structures, the chimeric complex was treated by proteases and its proteolysis patterns were then compared to the patterns of those of the isomeric complex, the yeast Nup145C/yeast Sec13 complex, in order to rule out the possibility of nonspecific binding. Remarkably, the proteolysis patterns of the chimeric complex also showed two stable protein fragments after protease digestion, as in the isomeric complex (Fig 31). Hence, the chimera is likely to have an identical tertiary structure as the isomeric complex.

Full-length human $Sec13^{1-322}$ was then co-expressed with yeast $Nup145C^{125-555}$ and the $Nup145C^{125-555}$/$hSec13^{1-322}$ chimeric complex was purified as described in the materials and methods section. Once the pure $Nup145pC^{125-555}$/$hSec13^{1-322}$

chimeric complex was obtained (Fig 32), various crystallization-screening kits were applied for screening. However, only poor-quality crystals were obtained during the screening.

Notably, when the protein behavior of full-length hSec13$^{1-322}$ in isolation was analyzed by gel filtration chromatography, full-length hSec13$^{1-322}$ was found to exist in a dynamic equilibrium between monomers and dimers. Moreover, the dimerization of full-length hSec13$^{1-322}$ is strongly dependent on the six C-terminal residues (Fig 33A). Furthermore, the Nup145C$^{125-555}$/hSec13$^{1-322}$ heterodimer forms higher-order oligomers in a concentration-dependent manner (Fig 33B and 44). Various oligomerization states were observed in solution, and, thus, this chimeric heterodimer has a tendency to oligomerize, which may have caused the low quality of the crystals. In contrast, the Nup145C$^{125-555}$/hSec13$^{1-316}$ complex yielded only a small and barely detectable peak of hetero-octamers and the majority of the complex eluted as a second single sharp peak from a gel filtration column (Fig 33B), indicating the potential of yielding higher-quality crystals.

Further gel filtration chromatographic analyses of the Nup145C$^{125-555}$/hSec13$^{1-316}$ complex indicated that the single peak corresponds to a dynamic equilibrium between heterodimers

and heterotetramers (Fig 44). The analysis of the principal peak of this chimeric heterodimer by combining SEC with multi-angle light scattering (Fig 34) unambiguously confirmed that the red-colored peak in Fig 32B represents a dynamic equilibrium between heterodimers and heterotetramers.

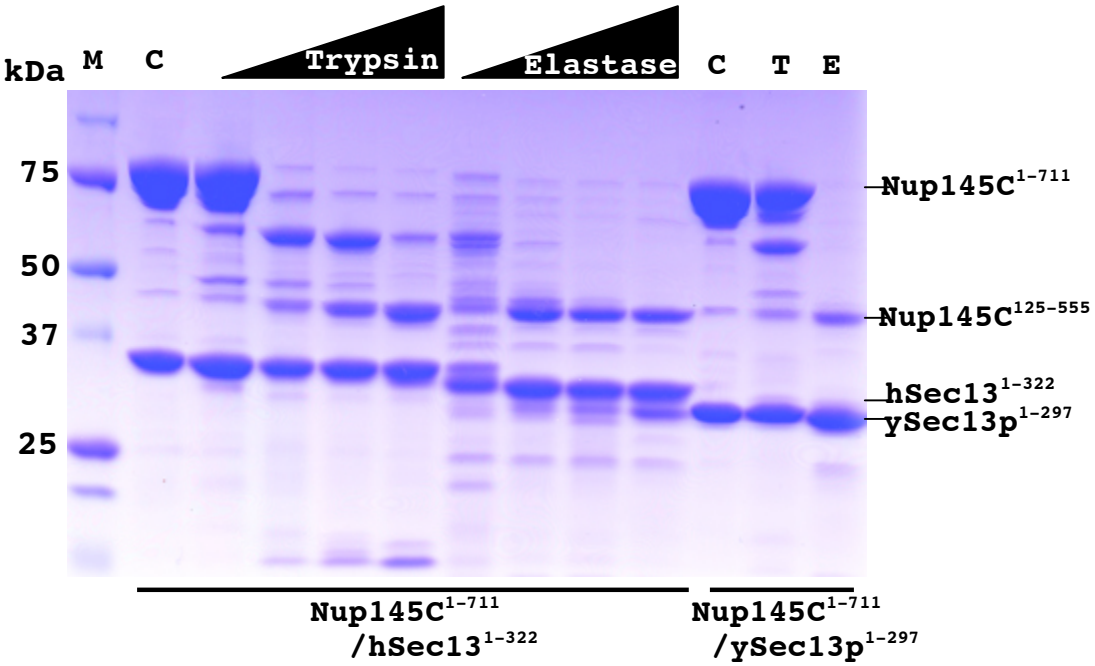


Figure 31. Limited proteolysis of full-length Nup145C/yeast Sec13 and full-length Nup145C/human Sec13

The full-length Nup145C/yeast Sec13 and full-length Nup145C/human Sec13 complexes were treated by trypsin and elastase separately and analyzed by SDS-PAGE. C, T and E indicate no protease treatment (control), trypsin treatment and elastase treatment, respectively.

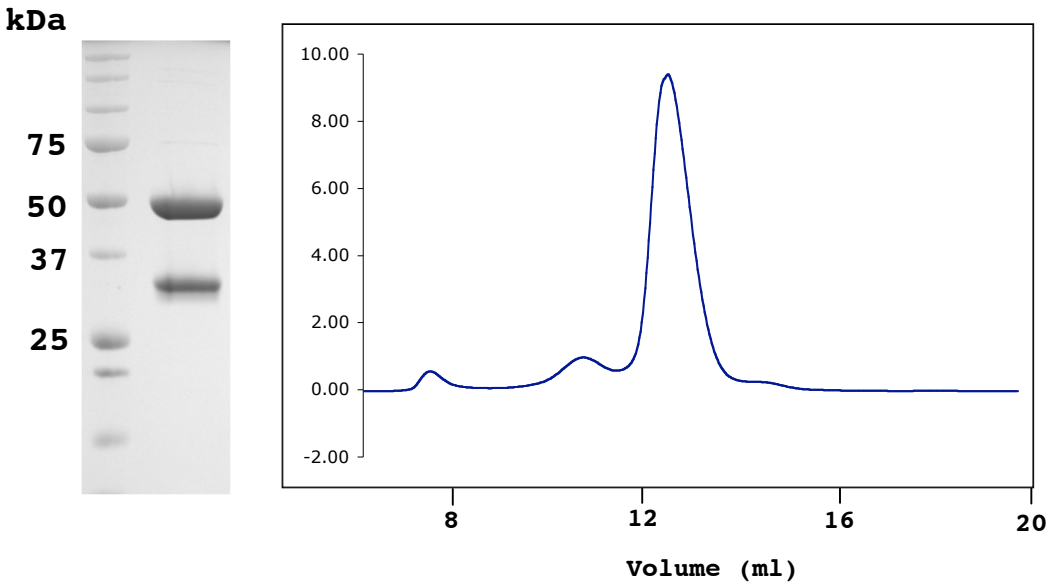


Figure 32. Purification of the truncated Nup145C$^{125-555}$/ human Sec13$^{1-322}$ chimeric complex

Gel filtration profile of the chimeric complex on a Superdex 200HR 10/300 column. The major peak fractions were pooled and analyzed by SDS-PAGE. The SDS-PAGE was stained with Coomassive brilliant blue.

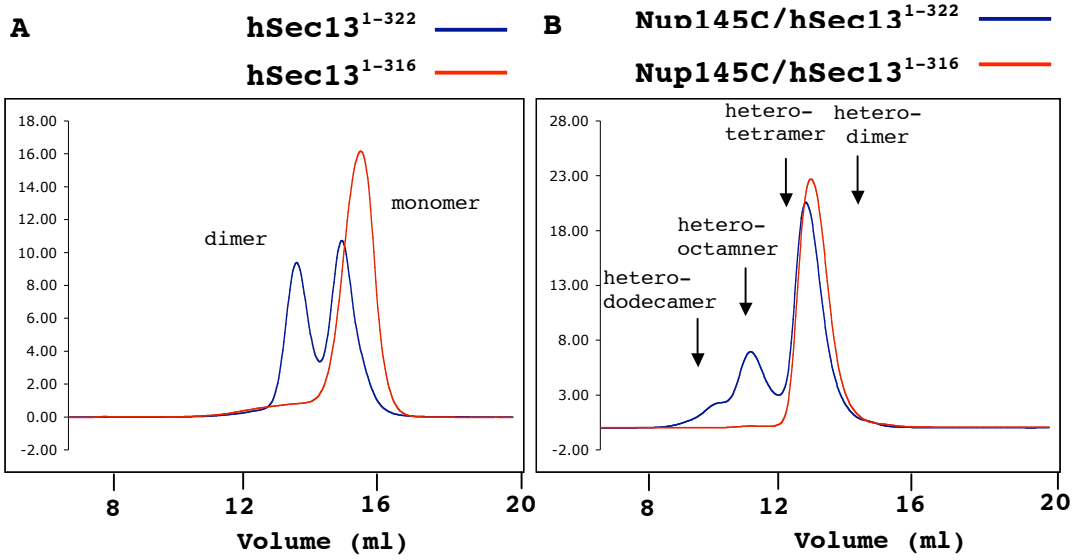


Figure 33. Analyses of oligomerization of human Sec13 and Nup145C/hSec13 complexes

(A) Gel filtration profiles of full-length human Sec13$^{1-322}$ (Blue line) and C-terminally truncated human Sec13$^{1-316}$ (Red line). Human Sec13$^{1-322}$ forms a dynamic equilibrium between monomers and dimmers in solution but human Sec13$^{1-316}$ does not. (B) Gel filtration profiles of Nup145C in complex with either hSec13$^{1-322}$ (Blue line) or hSec13$^{1-316}$ (Red line). Nup145C/hSec13$^{1-322}$ shows various oligmerization states in solution. In contrast, there are no two peaks with the majority in a single sharp peak in Nup145C/hSec13$^{1-316}$.

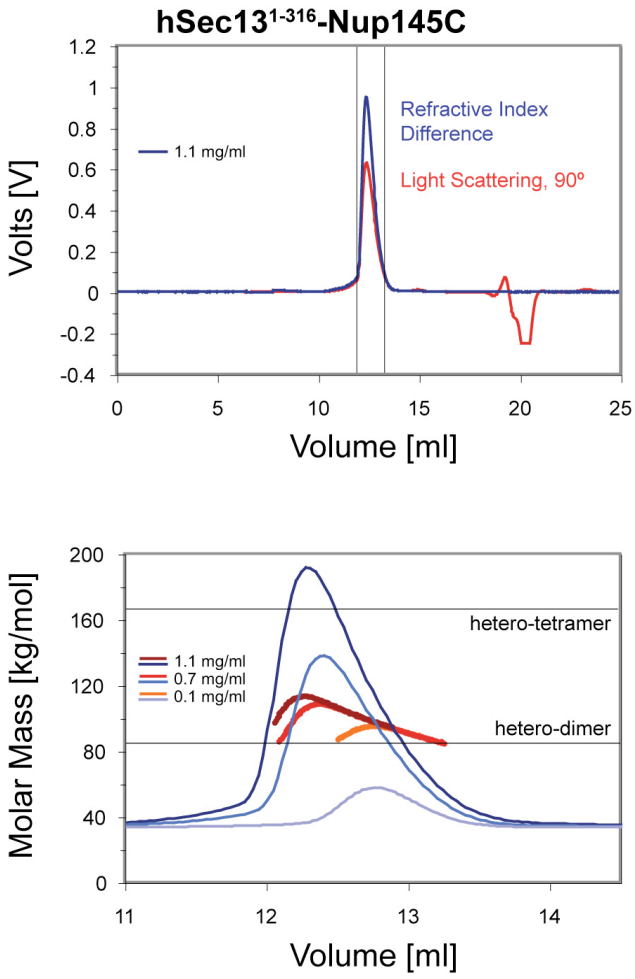


Figure 34. Multi-angle light scattering analysis of the Nup145C$^{125-555}$/hSec13$^{1-316}$ complex

The black bars on each side of the peak delimit the data points used for further analysis (upper panel). Molecular weights were determined by light scattering at the three indicated protein concentrations (lower panel). The complex forms a concentration-dependent, polydisperse dynamic equilibrium between heterodimeric and heterotetrameric assemblies in solution.

3.2.3 Crystallization and structure determination of the truncated Nup145C$^{125-555}$/hSec13$^{1-316}$ complex

The initial crystallizations were tested using various crystal screening kits and crystals were improved by using grid screens around the original conditions. The first crystal form was obtained from 0.1 M BIS-TRIS, pH 6.7, 0.2 M NaI, 16% PEG1500, and 2 mM ATP (Fig 35). Although the crystals showed a nice shape and sharp edges, the diffraction patterns were anisotropic with a weak zone of sharply decreased resolution. In addition, another severe problem was that the diffraction spots disappeared in certain rotational angles. Taken together, probably due to imperfect crystal packing, this crystal form could not provide sufficient and high-qualified diffraction data for solving the structure.

In the meantime, two other crystal forms were found. Fortunately, according to preliminary crystallographic analyses, these two crystal forms contained relatively small unit cell dimensions and also gave better quality of diffraction patterns in general (Fig 36).

The crystals of the Nup145C$^{125-555}$/hSec13$^{1-316}$ complex, based on structure determination, belong to the monoclinic space group C2 and the orthorhombic space group $C222_1$. Each crystal form contained an identical ~400 kDa hetero-octamer

in the asymmetric unit. The structures of both crystal forms were determined independently by multiple anomalous dispersion (MAD), using X-ray diffraction data obtained from seleno-L-methionine (SeMet)-labeled and several heavy-metal derivatized crystals. The final structures derived from the two space groups were refined to 3.0 and 3.15 Å resolution, respectively. Since the two structures are overall very similar, the discussion will focus on the orthorhombic crystal form in the following sections.

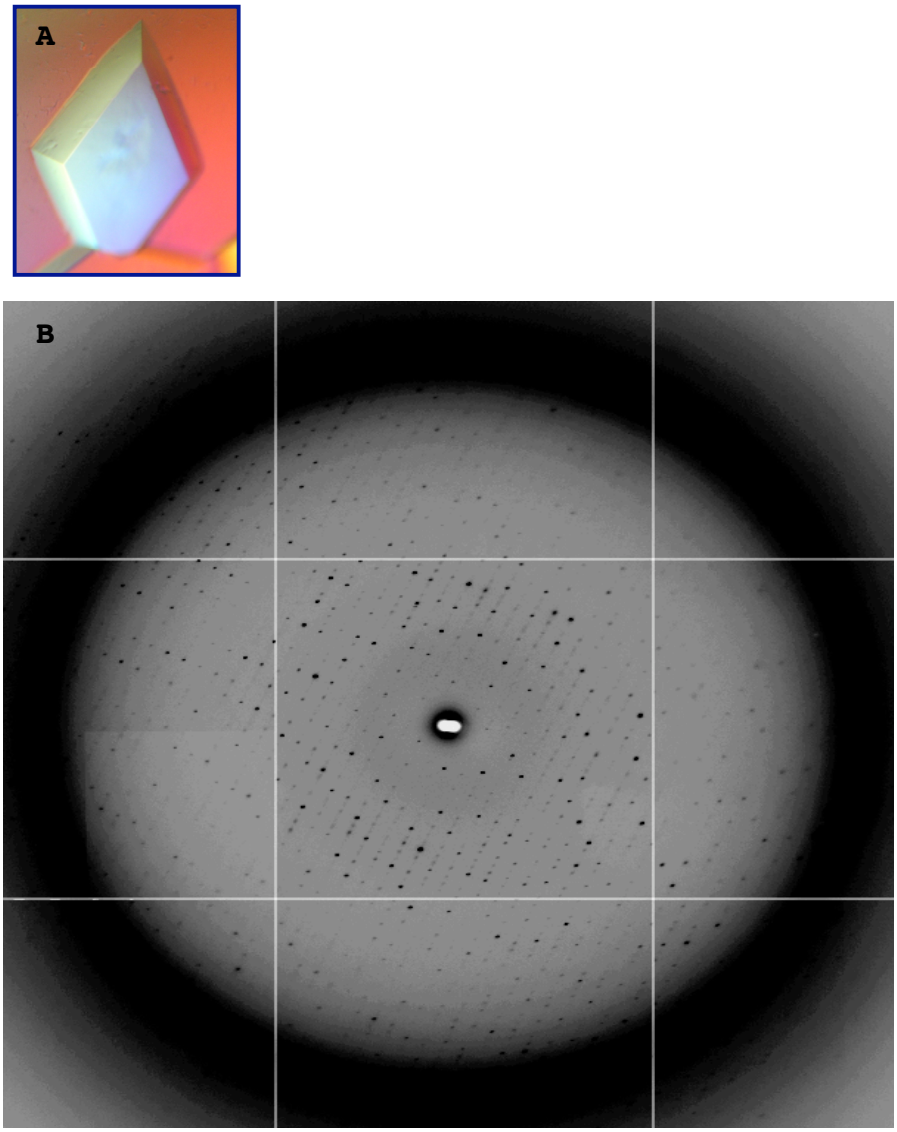


Figure 35. The truncated Nup145C$^{125-555}$/human Sec13$^{1-316}$ complex crystal and its x-ray diffraction pattern

(A) Crystal of the Nup145C$^{125-555}$/human Sec13$^{1-316}$ complex. (B) X-ray diffraction pattern of this complex obtained on beamline BL8.2.2 at ALS. The distribution of the diffraction spots was very anisotropic. Moreover, diffraction spots disappeared at certain rotation angles, probably due to disorder in the crystal packing.

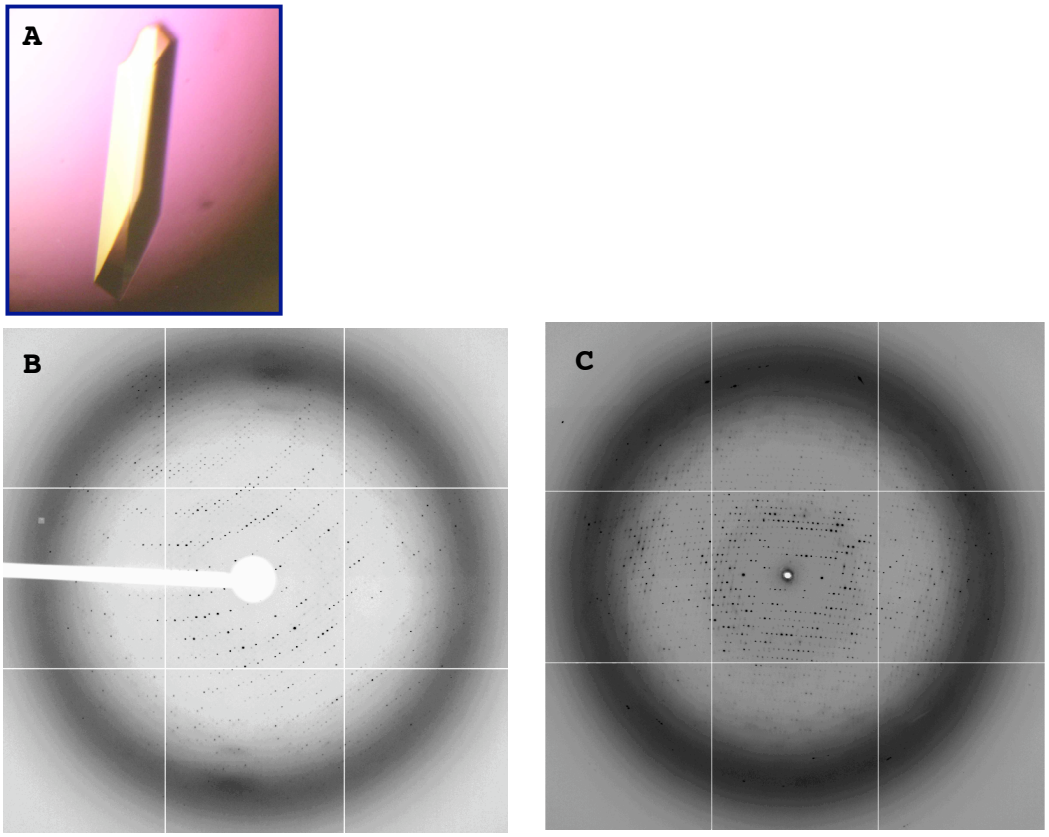


Figure 36. The truncated Nup145C$^{125-555}$/human Sec13$^{1-316}$ complex crystal and its typical X-ray diffraction patterns

(A) Crystal of the truncated Nup145C$^{125-555}$/human Sec13$^{1-316}$ complex. (B)(C) X-ray diffraction pattern of this complex obtained on beamline X29 at NLSL and on beamline BL8.2.2 at ALS, respectively. Space groups of the diffractions in (B) and (C) are monoclinic (C2) and orthorhombic ($C222_1$), respectively. Based on the Matthews coefficient calculation, there are four complexes per asymmetric unit in both space groups.

3.3 Structural features of the Nup145C$^{125-555}$/hSec13$^{1-316}$ complex

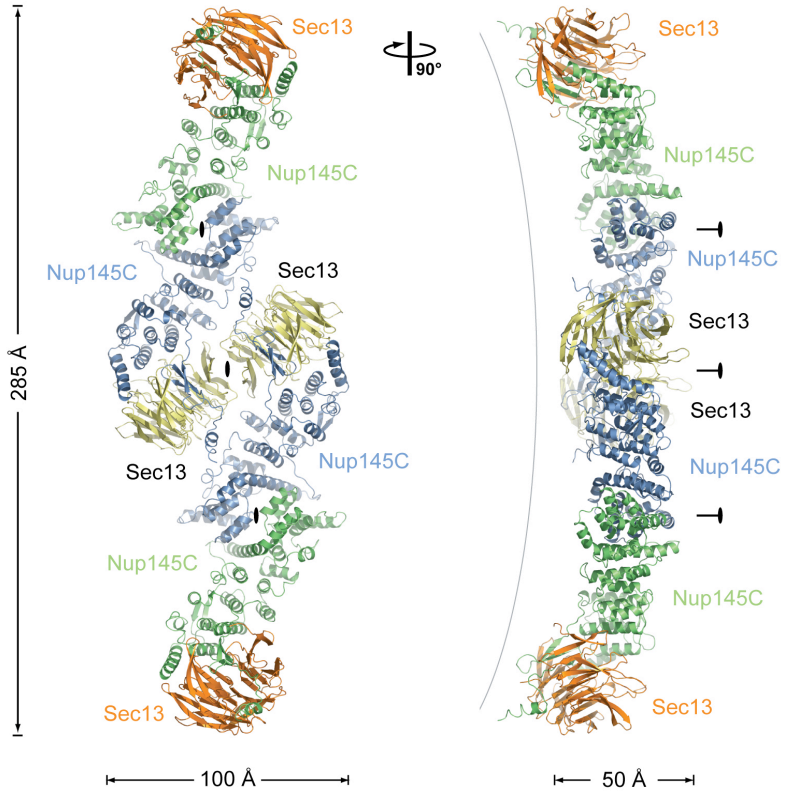
3.3.1 Architectural overview

For simplicity, Nup145C$^{125-555}$ and hSec13$^{1-316}$ are referred to as Nup145C and Sec13, respectively in the remainder of this section. The Nup145C/Sec13 hetero-octamer forms a slightly bent rod with overall dimensions of approximately 285 Å x 100 Å x 50 Å (Fig 37). The rod contains three pseudo-two-fold axes of symmetry. The central axis of symmetry passes through the interface of two heterotetramers that associate primarily, but not exclusively, via the homodimerization of Sec13. Around the adjacent two axes of symmetry, two heterodimers interact exclusively through the homodimerization of Nup145C.

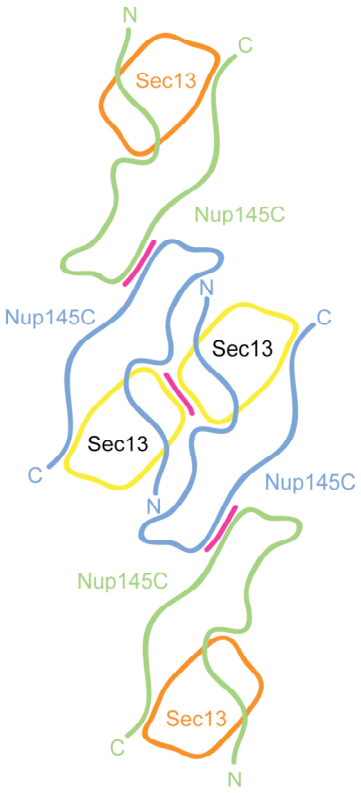
Figure 37. Overview of the structure of the Nup145C/Sec13 hetero-octamer

(A) Ribbon representation of the Nup145C/Sec13 hetero-octamer, showing Sec13 in yellow and orange and Nup145C in green and blue. A 90° rotated view is shown on the right. The three pseudo-two-fold axes (black ovals) that run through the hetero-octamer and the overall dimensions are indicated.
The Nup145C/Sec13 hetero-octamer forms a slightly bent rod. (B) Schematic representation of the Nup145C/Sec13 hetero-octamer. Magenta lines indicate interaction surfaces.

Figure 37 A


Sec13
90°
Nup145C
Nup145C
Sec13
285 Å
Sec13
Nup145C
Nup145C
Sec13
Sec13
Nup145C
Nup145C
Sec13
Nup145C
Sec13
Nup145C
Sec13
100 Å
50 Å

B


N
C
Sec13
Nup145C
Nup145C
N
C
Sec13
Sec13
C
N
Nup145C
Nup145C
Sec13
C
N

3.3.2 Structure of Nup145C

Nup145C has a novel fold in the form of a U-shaped structure. The 50-residue N-terminal region of Nup145C folds into an α helix (αA), followed by a short connecting segment (αA-β1 connector), and three anti-parallel β strands (β1–β3) (Fig 38). Together, this region represents the "domain invasion motif," or DIM, of Nup145C.

Following the DIM, there is a short segment that connects to a 400-residue C-terminal domain. This C-terminal domain is principally constructed of 18 anti-parallel α helices (αB-αS) that form a novel fold. A striking kink extending from αE to αH allows the chain to fold back onto itself. The resulting U-shaped structure is held together by extensive hydrophobic interactions between the ascending and descending arms of the "U" and by a 20 residue long loop (between αB and αC) that extends from the U's descending arm and wraps around the U's ascending arm.

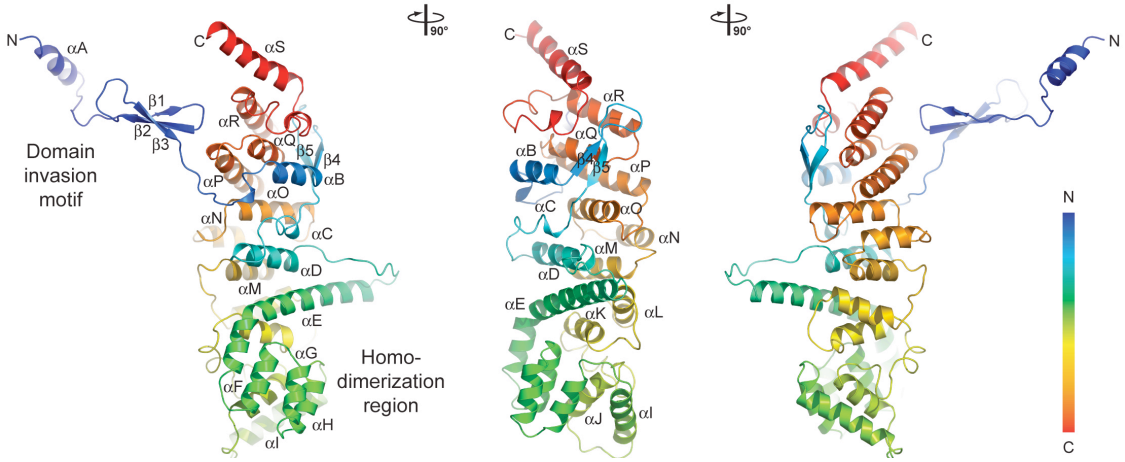


Figure 38. The structure of Nup145C

The ribbon representation of the Nup145C structure is shown in rainbow colors along the polypeptide chain from the N- to the C-terminus. The N-terminal domain invasion motif (DIM), the C-terminal α-helical domain, and their secondary structure elements are indicated.

3.3.3 Structure of the β-propeller domain of Sec13 in complex with three-stranded blade of Nup145C

The Sec13 expression construct contains 14 residues on its N-terminus that are predicted to contain an additional β strand that would typically complete the last blade in a β-propeller architecture. Like the 12 residues at the C-terminus that are predicted to contain two additional β-strands, these 26 residues were not visible in the electron density map and, therefore, were presumably disordered. Hence, the C-terminal residues are common for ySec13 and hSec13, but not including the additional C-terminal residues that are unique to hSec13.

Sec13 contains six WD40 repeats that fold into an unexpected opened six-bladed β-propeller, unlike what was predicted by bioinformatics (Devos et al., 2004). Interestingly, the N-terminal domain of Nup145C forms an additional three-stranded blade inserted into Sec13's β-propeller fold, complementing it with a seventh blade (Fig 39).

For a canonical β-propeller fold, the last blade is mostly formed by a combination of N- and C-terminal β-strands, resulting in a closed rigid ring-like structure. However, probably due to the low affinity between the additional N-terminal β-strand of human Sec13 and the DIM of

yeast Nup145C, the Sec13's N-terminal β-strand is disordered instead of binding to the last β strand of the DIM, also indicating that this N-terminal additional β-sheet probably is not crucial for the Nup145C and Sec13 recognition. As a result, the typical "Velcro" closure of a canonical β-propeller was not observed in this case.

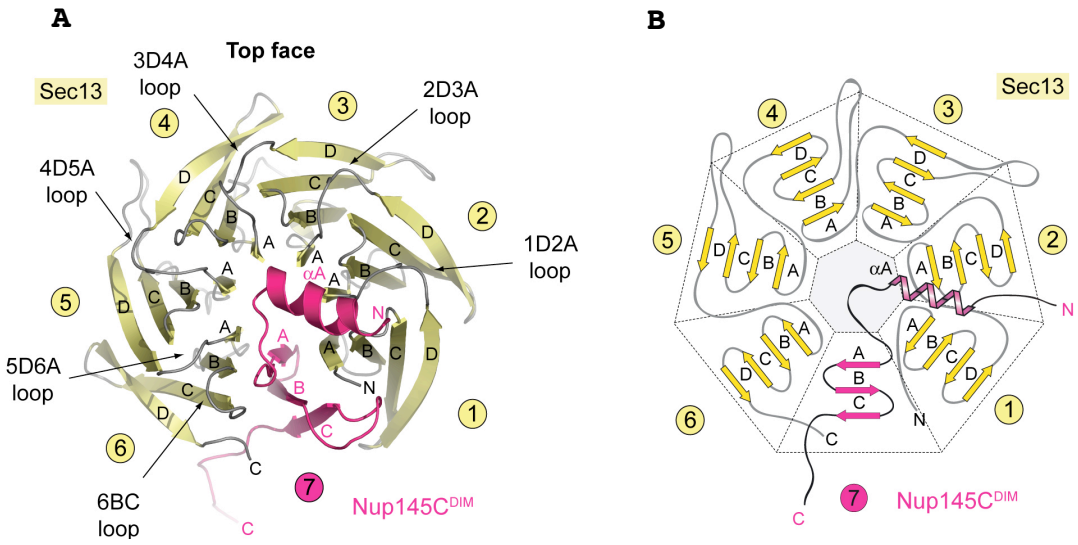


Figure 39. The structure of the Sec13 β-propeller in complex with the Nup145C^{DIM}

(A) Sec13 is shown in yellow, and the six blades are indicated. The Nup145C^{DIM} forms a three-stranded seventh blade, complementing the Sec13 β-propeller. (B) Schematic representation of the Sec13 β-propeller and its interaction with the Nup145C^{DIM}.

3.3.4 Structure of the Nup145C/Sec13 heterodimer

Sec13 and Nup145C form a heterodimer by interacting with each other via several distinct surfaces. In general, the interacting surfaces between these two proteins can be divided into three main areas.

Firstly, the DIM's α-helix αA and the αA-β1 connector segment interact with Sec13's D-A and B-C loops from all six of its blades, forming the top face of the β-propeller domain (Fig 39 and 40). These surface loops form a circular bracelet that appears to stabilize the open conformation of Sec13's β-propeller domain. Secondly, the DIM's three β-strands form a seventh blade that is inserted between blades 1 and 6 of the Sec13 β-propeller domain (Fig 39 and 40). Thirdly, the four C-terminal α-helices (αP—αS) tightly interact with almost all of the A-B and C-D loops of the six blades that connect the β-strands on the bottom face of Sec13's β-propeller domain (Fig 40). Altogether, the interaction between the two proteins buries ~5250 $Å^2$ of surface area and involves 69 and 58 residues of Sec13 and Nup145C, respectively.

The Sec13 residues participating in complex formation are highly conserved between human and yeast, in line with the observation that human Sec13 can replace yeast Sec13 in the Nup145C/Sec13 complex. Thus, the residue conservation

and the huge buried surface area for the Nup145C and Sec13 interaction strongly suggest that both proteins form a tight structural unit, highly conserved through eukaryotic evolution.

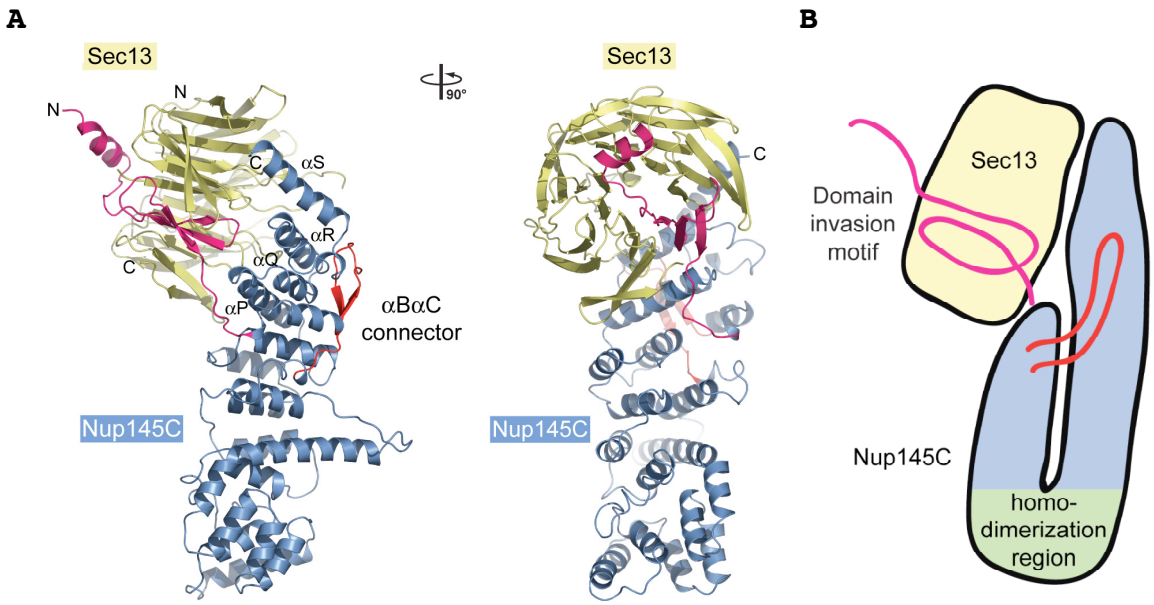


Figure 40. The structure of the Nup145C/Sec13 hetero-dimer

(A) The Nup145C^{DIM} (magenta), the Nup145C α-helical domain (blue), the Nup145C αB-αC connector segment (red), and the Sec13 β-propeller (yellow) are indicated; a 90° rotated view is shown on the right. (B) Schematic representation of the Nup145C/Sec13 interaction.

3.3.5 Structure of the Nup145C/Sec13 heterotetramer

The heterotetramer is formed by a head-to-head homodimerization of Nup145C. Four a-helices (αF, αH, αI, and αJ) that are located in the center of the domain at the sharp kink where the domain folds back onto itself form a predominantly hydrophobic interface that covers ~2700 $Å^2$ of surface area (Fig 41). A two-fold axis of symmetry runs through the dimerization interface (Fig 42A) and the residues in this interface are also highly conserved.

Figure 41. Surface properties of the Nup145C/Sec13 heterodimer

(A) Surface rendition of the Nup145C/Sec13 complex. The surface is colored according to the proteins (Sec13, yellow; Nup145C, blue) and their participation in various interactions (Sec13 from the adjacent complex, orange; Nup145C from the adjacent complex, purple; Nup145C homodimerization, green). (B) Nup145C is colored according to sequence conservation, from 40% similarity (white) to 100% identity (green). (C) Surface rendition of Nup145C, colored according to the electrostatic potential, from red (-10 k_BT/e) to blue (+10 k_BT/e).

Figure 41

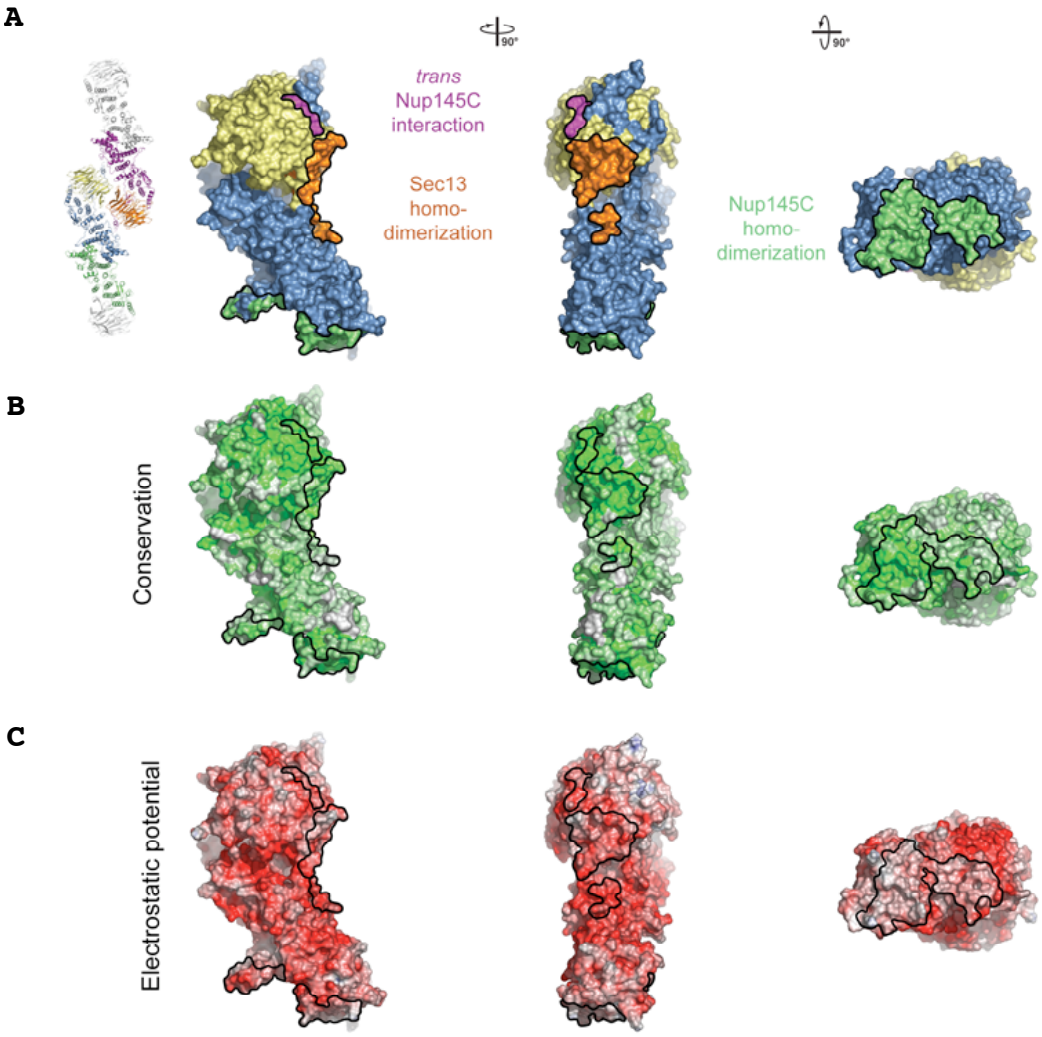
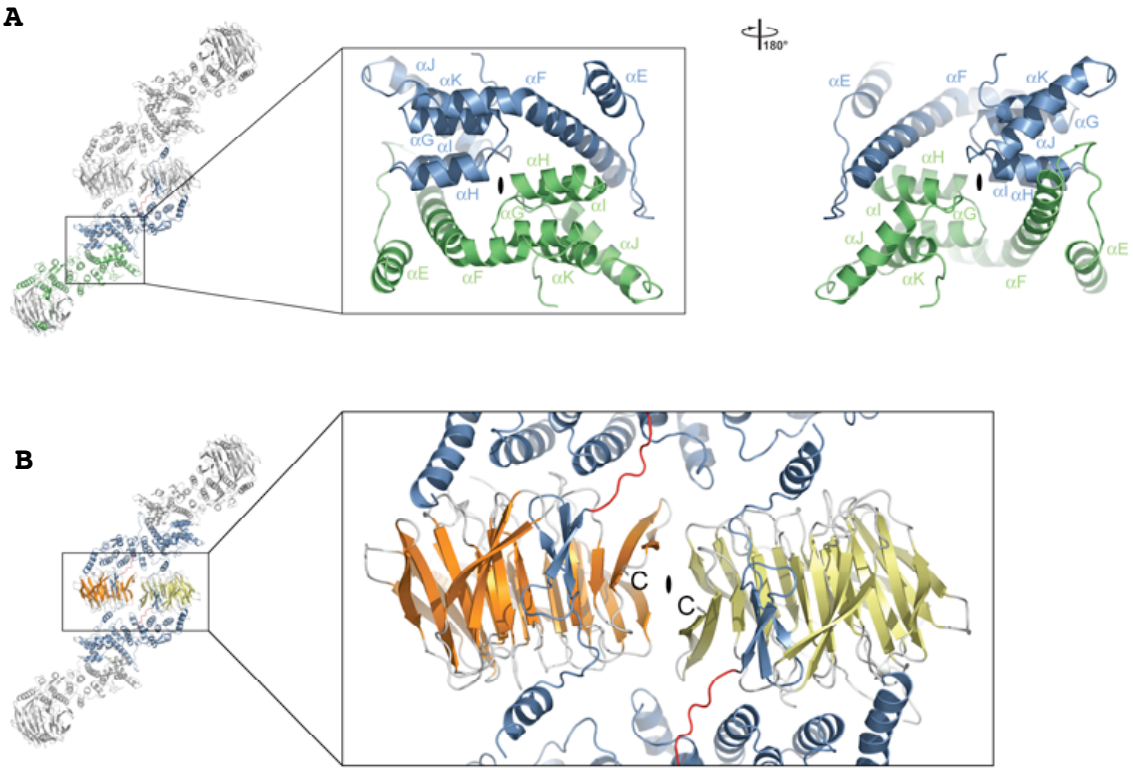

A
90°
90°
trans
Nup145C
interaction
Sec13
homo-
dimerization
Nup145C
homo-
dimerization
B
Conservation
C
Electrostatic potential

Figure 42. The assembly of the Nup145C/Sec13 hetero-octamer

(A) Ribbon representation of the dimerization of the Nup145C α-helical domain. For clarity, only the two interacting Nup145C protomers in the heterooctamer are colored (green and blue). The centrally located α-helices that facilitate the homodimerization of Nup145C are indicated. A 180° rotated view is shown on the right. (B) The dimerization interface located in the center of the Nup145C/Sec13 hetero-octamer. The two Sec13 β-propeller domains (orange and yellow) and two Nup145C molecules (blue) are colored. The two Sec13 C-termini (the last ordered residue is Val304) are indicated and are only ~10 Å apart.

Figure 42


A
αJ
αK
αF
αE
αG
αI
αH
αH
αG
αI
αJ
αE
αF
αK
180°
αE
αF
αK
αG
αJ
αH
αI
αH
αI
αG
αJ
αE
αK
αF
B
C
C

3.3.6 The curved Nup145C/Sec13 hetero-octamer rod

In both crystal forms, two identical Nup145C/Sec13 hetero-octamers were packed differently in the two crystal lattices. These two hetero-octamers superimpose with a root mean square deviation (RMSD) of 0.75 Å, strongly suggesting that the overall structure is rigid (Fig 43).

The hetero-octamer is constructed by the head-to-head interaction of two heterotetramers, involving the surfaces of both Sec13 and Nup145C. In general, this interacting interface is split into two areas that also are highly conserved in evolution and are predominantly hydrophobic (Fig 41). The first interface via the Sec13 homodimerization is mediated by the two terminal β-strands C and D of blades 5 and 6, the inter-blade groove between the two blades, as well as the connector loops 5D-6A and 6BC (Fig 42B). This interface covers ~1100 $Å^2$ of surface area buried between the two Sec13 β-propellers. The Sec13 dimerization surface is essentially invariant between the yeast and human proteins. The second interface is generated by the trans interaction of each of the two Sec13 β-propellers with an adjacent Nup145C. This interaction involves the inter-blade connector loops 3D-4A and 4D-5A of Sec13 and the αM-αN connector loop of the Nup145C α-helical domain and together covers ~600 $Å^2$. In total, ~1700 $Å^2$ of

surface area is buried between the two Nup145C/Sec13 heterotetramers.

The fact that a hetero-octamer was crystallized, although this form was present in solution only in small quantities, can readily be explained by protein concentrations that are estimated to be more than one order of magnitude higher in the crystal than the highest concentration tested in solution (Fig 44). A protein concentration similar to that in the crystal is likely to exist in the NPC in vivo.

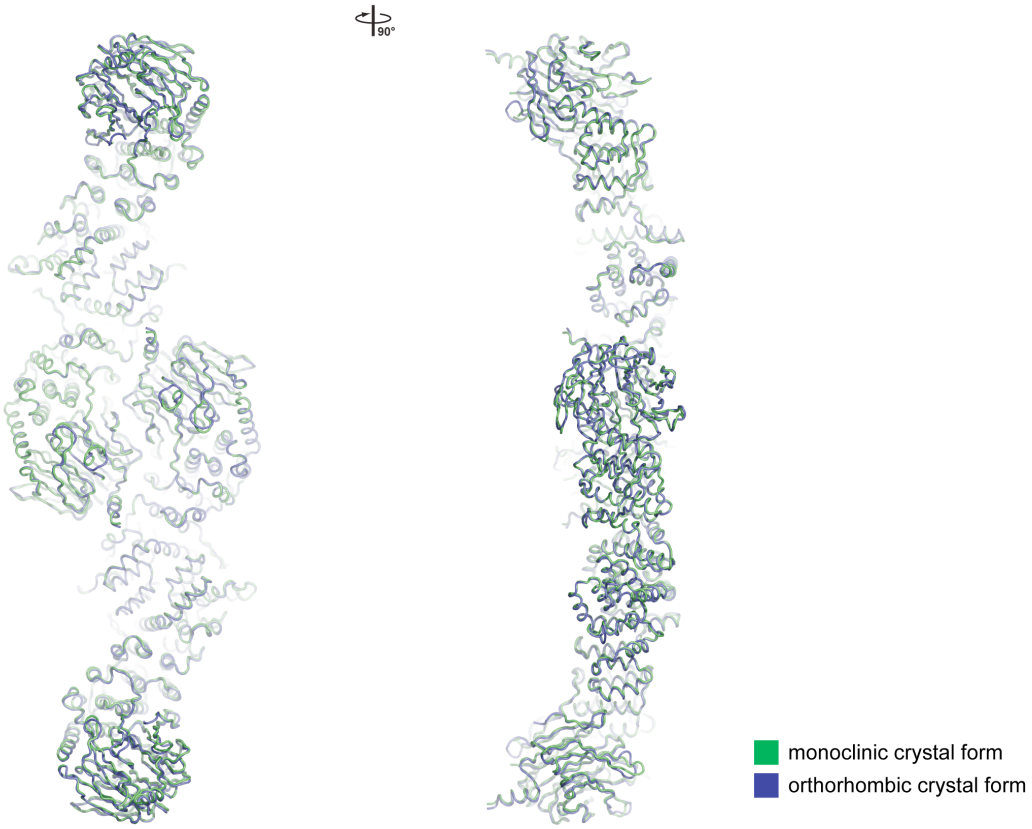


Figure 43. Comparison of the Nup145C/Sec13 structures from the two crystal forms

Coil representations show the superposition of the Nup145C/Sec13 hetero-octamers from the orthorhombic (green) and monoclinic (blue) crystal forms. A 90^0 rotated view is shown on the right.

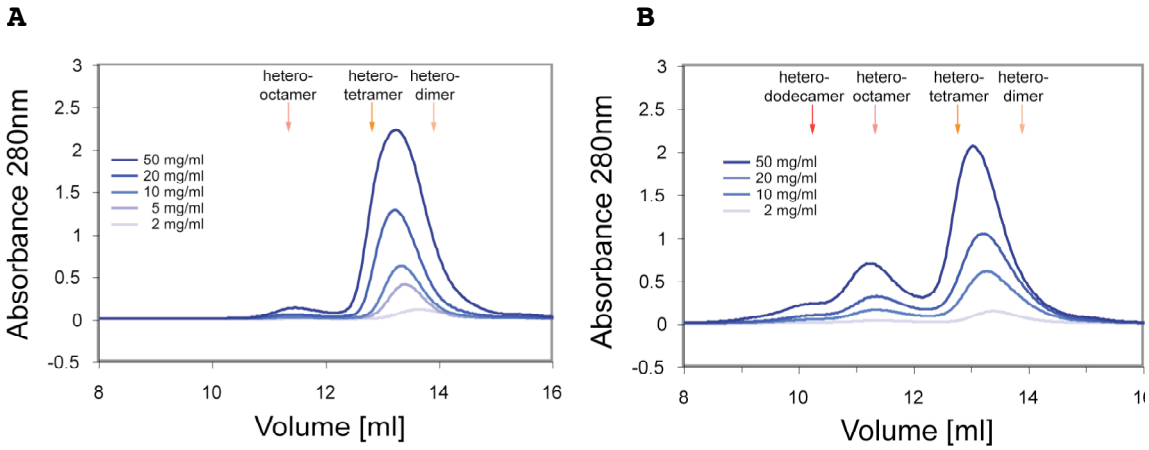


Figure 44. Dynamic behavior of Nup145C/Sec13

Gel filtration profiles of (A) the Nup145C/yeast Sec13 complex and (B) the Nup145C/human Sec13$^{1-322}$ complex are shown. All proteins were injected at the indicated concentrations. The predicted elution positions for the various assembly states of the Nup145C/Sec13 complexes are shown and have been determined using molecular weight standards.

3.4 Binding selectivity of Sec13

Although human Sec13 has been proved to be capable of specifically forming a complex with yeast Nup145C in vitro, whether or not human Sec13 can rescue yeast Sec13's functions in vivo remains to be answered.

In order to test if human Sec13 can replace its yeast counterpart in vivo, a plasmid shuffle assay was performed. A yeast shuffle strain, which carries a disrupted ySec13 allele and a plasmid containing an additional copy of the ySec13 gene, was transformed by a second plasmid which contains either human or yeast Sec13. Surprisingly, in contrast to the viability of over-expressing ySec13 cells, yeast cells that were transformed by plasmids encoding either full-length human $Sec13^{1-322}$ or $Sec13^{1-316}$ were not viable on the 5-FOA plates (Fig 45B), indicating that human and yeast Sec13 are not interchangeable in yeast.

Although the result of the plasmid shuffle assay was surprising, further in vitro pull-down studies indicated that it was just unexpected but not unexplained. According to pull-down studies, as well as $Nup145C^{125-555}$ and $hSec13^{1-322}$, $hSec13^{1-322}$ and yeast $Sec31^{370-763}$ could only form a complex when the two proteins were co-expressed. Furthermore, once three proteins, $Nup145C^{125-555}$, $hSec13^{1-322}$ and ySec13 were co-expressed, two complexes, $Nup145C^{125-555}$/hSec13 and $Nup145C^{125-}$

^{555}/ySec13, could be pulled down by using a His tag, which is only attached to Nup145C$^{125-555}$. However, once hSec13$^{1-322}$ co-expressed with ySec31$^{370-763}$ and ySec13, probably due to lower affinity to ySec31, hSec13 was not able to compete with ySec13 and, thus, all ySec31$^{370-763}$ was occupied by ySec13 (Fig 45A). Taken together, the lower affinity between hSec13 and ySec31 perhaps results in the impairment of the COPII complex which is likely to cause the inviability of yeast observed in the plasmid shuffle assay even though hSec13 was still able to bind to Nup145C.

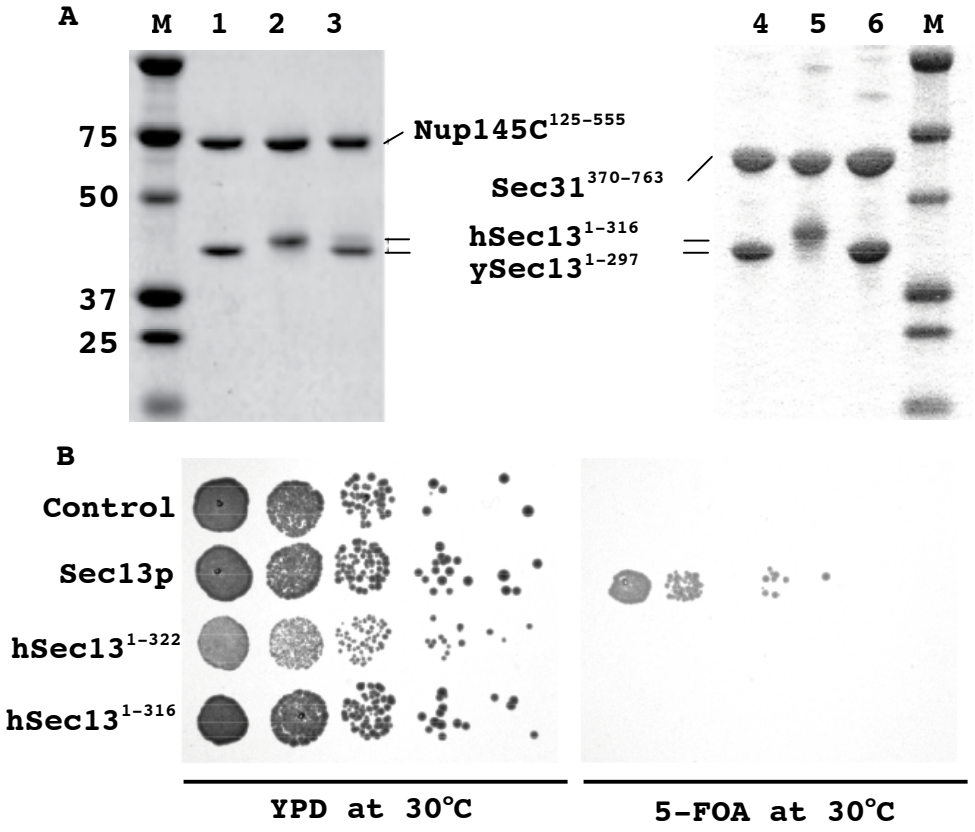


Figure 45. hSec13 is unable to rescue ySec13 function in yeast due to lower binding affinity to Sec31

(A) hSec13 and ySec13 can bind to Nup145C and Sec31 to form different complexes. (Lane 1, 2, 4 and 5) However, once hSec13 was co-expressed with ySec13/Nup145pC or ySec13/Sec31, only Nup145C could pullout some hSec13, while Sec31 was unable to do so. (Lane 3 and 6) (B) Yeast Sec13 shuffle strain (sec13::HIS3;pRS316-URA-SEC13) was transformed by control, pRS414, and pRS414 containing Sec13 variants, ySec13, $hSec13^{1-322}$ and $hSec13^{1-316}$ individually. Different transformants were spotted at 10 fold serial dilutions (from ~10^4 to 10) on the plates of YPD or minimum medium containing 5-FOA.

3.5 Reconstitution of pentameric, tetrameric and trimeric complexes in *E. coli*

3.5.1 Reconstitution of the pentameric complex

In the attempt to reconstitute larger complexes, for the first step, full-length Nup120 was purified as a single protein, whereas full-length Nup145C/Sec13 and Nup85/Seh1 complexes were isolated as pairs (Fig 46A). The pentameric complex consisting of Nup120, Nup145C, Sec13, Nup85, and Seh1 was then reconstituted by mixing these five proteins in approximately equimolar amounts. The mixture was further dialyzed against buffer containing 25 mM HEPES, pH 7.5, 300 mM NaCl, 150 mM potassium acetate, 2 mM Magnesium acetate and 3 mM DTT. After incubation for two hours, the mixture was applied to a superose 6 gel filtration column to separate the reconstituted pentameric complex from other unassembled proteins. In the gel filtration profile, the earlier peak that contains five full-length proteins in roughly stoichiometric amounts corresponds to a molecular weight of ~500 kDa (Fig 46B and C). Notably, only full-length proteins could assemble into the pentameric complex. Once the proteins were degraded, the degradation products only appeared in the later fractions (Fig 56B and C). Hence, these five proteins were capable of self-assembly to

a biochemically stable pentameric complex in vitro.

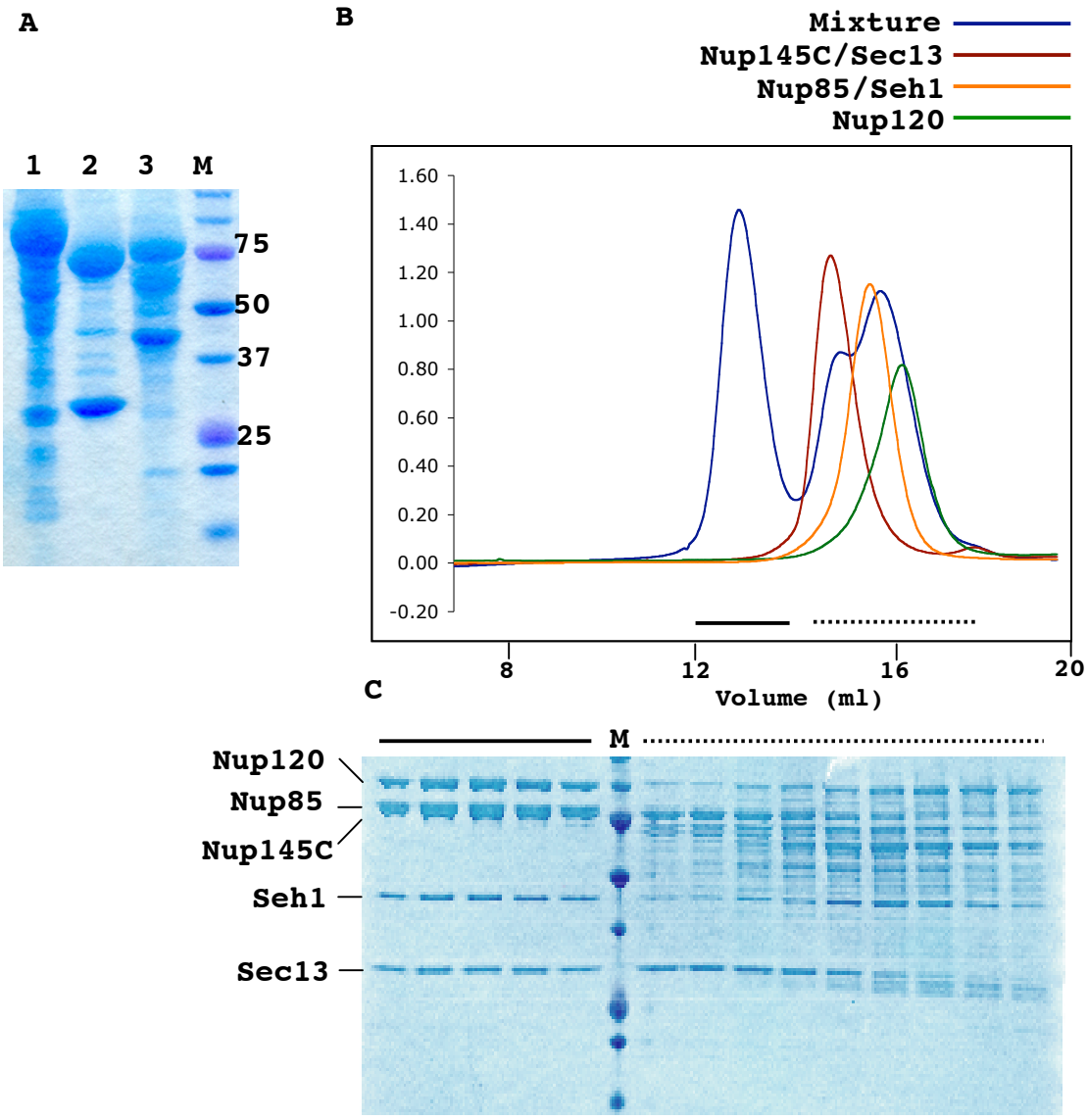


Figure 46. Reconstitution of the pentameric complex

(A) Purification of the full-length proteins of Nup120, Nup145C/Sec13 and Nup85/Seh1. (B) The gel filtration column profile of the five-protein mixture. The solid and dash lines represent the fractions that were analyzed by SDS-PAGE. (C) The fractions of the solid line showed five proteins in the SDS-PAGE.

3.5.2 Reconstitution and biochemical analyses of the trimeric complex

After the pentameric complex was obtained, we had to verify how the two pairs, Nup145C/Sec13 and Nup85/Seh1, and Nup120 interact with each other. Full-length Nup120 and Nup85/Seh1 were co-expressed and indeed could form a trimeric complex after a series of chromatographic purification steps (Fig 47A). Furthermore, the full-length trimeric complex was treated by limited proteolysis and applied to Edman sequencing and mass spectrometry analyses to identify the interacting regions between these proteins. Based on these analyses, the C-terminus of Nup85 (residues 566-744) was identified to be necessary for Nup120 binding. Moreover, since the folding of the entire Nup120 protein appears so compact that it would resist protease digestion, full-length Nup120 and Nup85$^{566-744}$ was co-expressed and purified as a protein complex (Figure 47B). Initial crystals were observed in crystal screening conditions (Figure 47C).

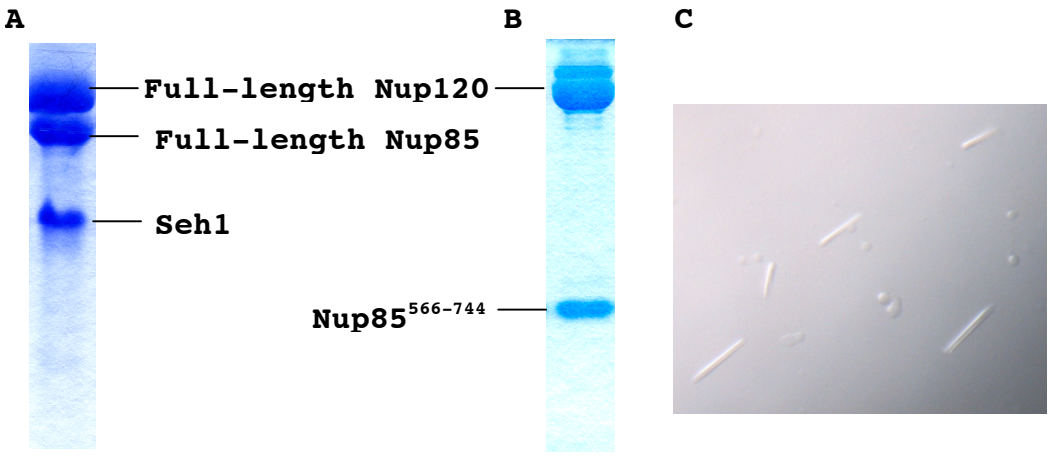


Figure 47. Purification and preliminary crystallization of the Nup120, Nup85, and Seh1 complex

(A) Purification of the trimeric comlex containing full-length proteins of Nup120, Nup85, and Seh1. (B) Purification of full-length Nup120 and $Nup85^{566-744}$ complex. (C) Initial crystals were observed within one week using crystal screening kits.

3.5.3 Reconstitution of the Nup85/Seh1/Nup145C/Sec13 complex

Since Nup120 interacts with Nup85, we determined what Nup145C's binding partner is. Thus, Nup145C/Sec13 complex was mixed with either full-length Nup120 or the Nup85/Seh1 complex and the mixtures were subjected to superose 6 gel filtration for analysis. Based on the gel filtration profiles, Nup145C/Sec13 could only bind to the Nup85/Seh1 complex instead of Nup120 (Fig 48).

Therefore, the two pairs, full-length Nup145C/Sec13 and Nup85/Seh1, were able to be purified and assembled to form a tetrameric complex (Fig 49A). Furthermore, according to limited proteolysis, Edman sequencing and mass spectrometry studies, a portion of the N-terminal unstructured region (residues 75-125) in Nup145C was identified as an interaction region for Nup85 binding. Hence, another tetrameric complex, including a short version of Nup145C, $Nup145C^{75-555}$, and three other full-length proteins was made (Fig 49B). Moreover, full-length Nup120 still could also be added to this tetrameric complex by binding to full-length Nup85 (Fig 49C).

Taken together, Nup120 binds to the C-terminal domain of Nup85 by its C-terminal region, and Nup145C interacts with Nup85 through its N-terminal residues. Hence, the C-terminal domain of Nup85, the N-terminal region of Nup145C,

and the C-terminal region of Nup120 seem to converge at one site of the Nup84 sub-complex, forming a triskelion. In addition, the β-propeller proteins, Sec13 and Seh1, and the N-terminal Nup120 domain, are not involved in the interaction of these proteins (data not shown). Moreover, based on the relative positions of the seven components in the Nup84 sub-complex, revealed by negative-stain EM (Fig 16), and the crystal structures of the Nup145/Sec13 and Nup85/Seh1 complexes, the N- and C-termini of the two U-shaped proteins, Nup85 and Nup145C, are suggested to face inward to be part of the hinge region of the pentamer, while the kink regions of the U-shaped structures might face outward. Nup85 indeed is like a bridge that links the Nup120 and Nup145C in the Nup84 sub-complex (Fig 50).

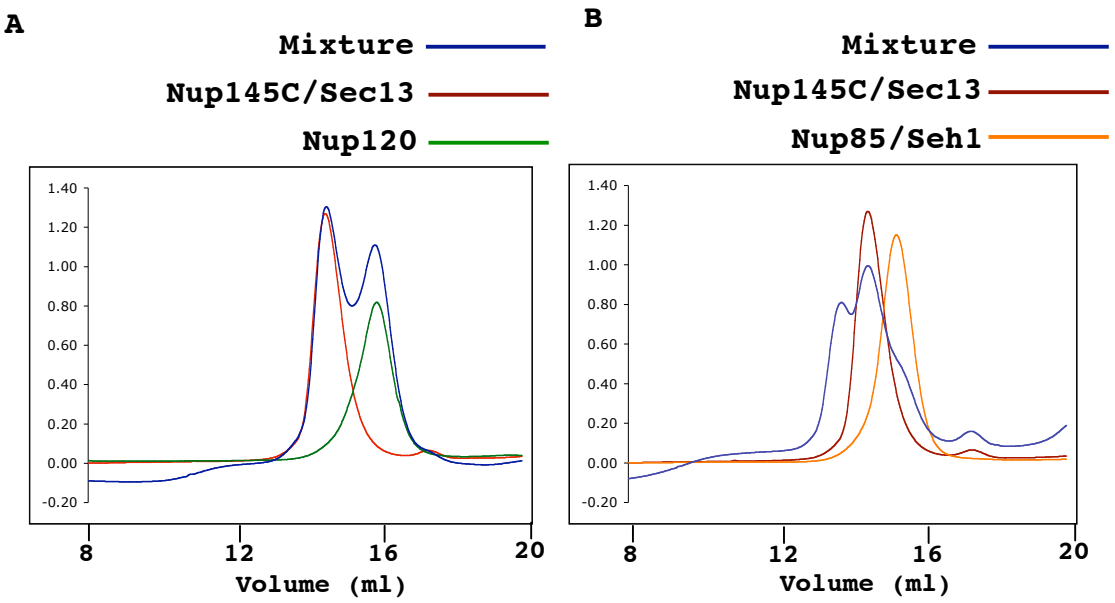


Figure 48. Identification of the binding partners of the Nup145C/Sec13 complex

(A) Full-length proteins of Nup120 and Nup145C/Sec13 complex were mixed and applied to superose 6 10/300 column. Compared with the controls (brown and green lines), Nup120 cannot interact with Nup145C/Sec13 complex. (B) Full length proteins of Nup85/Seh1 and Nup145C/Sec13 complexes were mixed and applied to a superose 6 10/300 column. According to the gel filtration profile, the additional peak indicates the interaction between the two pairs, Nup145C/Sec13 and Nup85/Seh1.

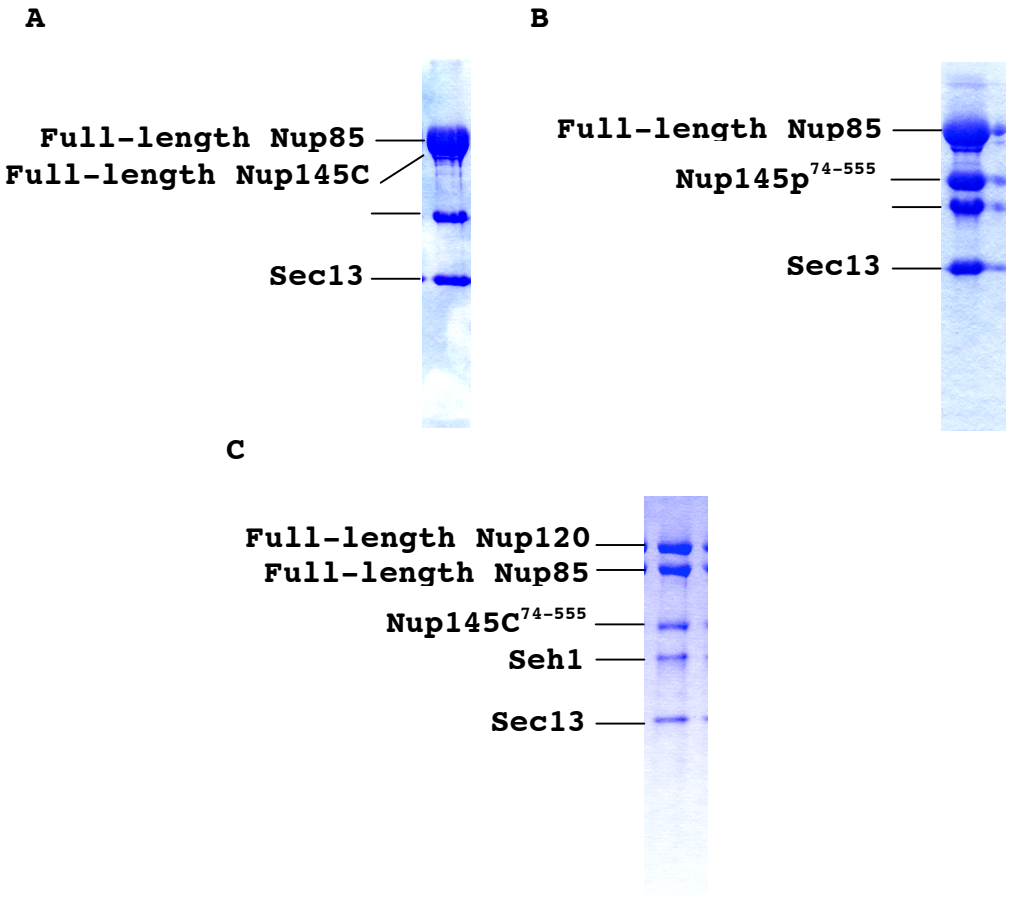
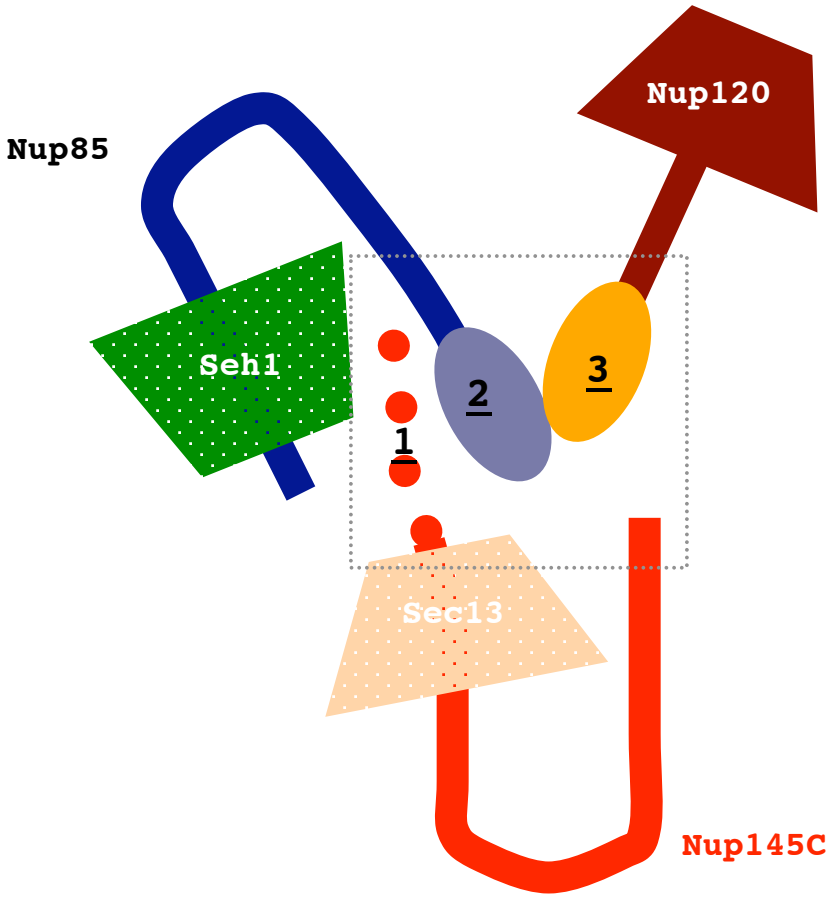


Figure 49. Purification of different complexes

(A) Purification of the tetrameric complex I containing full-length Nup85, Nup145C, Sec13, and Seh1. (B) Purification of the tetrameric complex II containing full-length Nup85, Sec13, Seh1 and Nup145C$^{74-555}$. (C) Purification of the pentameric complex II containing full-length Nup120, Nup85, Sec13, Seh1, and Nup145C$^{74-555}$.



1: N-terminal region of Nup145C
2: C-terminal domain of Nup85
3: C-terminal domain of Nup120

Figure 50. A schematic model of the pentameric complex

Nup85 can bind to the C-terminal region of Nup120 through its C-terminal domain. The N-terminal region of Nup145C, which does not belong to the crystallization fragment, is responsible for the interaction with Nup85C. β-propeller proteins, Sec13 and Seh1, are not involved in these protein-protein interactions. Hence, the N- and C-termini of the two U shaped structures, Nup85 and Nup145C, face inward and the kink regions face outward. Nup85 is like a bridge linking the Nup145C and Nup120. The box with dot line indicates the hinge region of this pentameric complex.

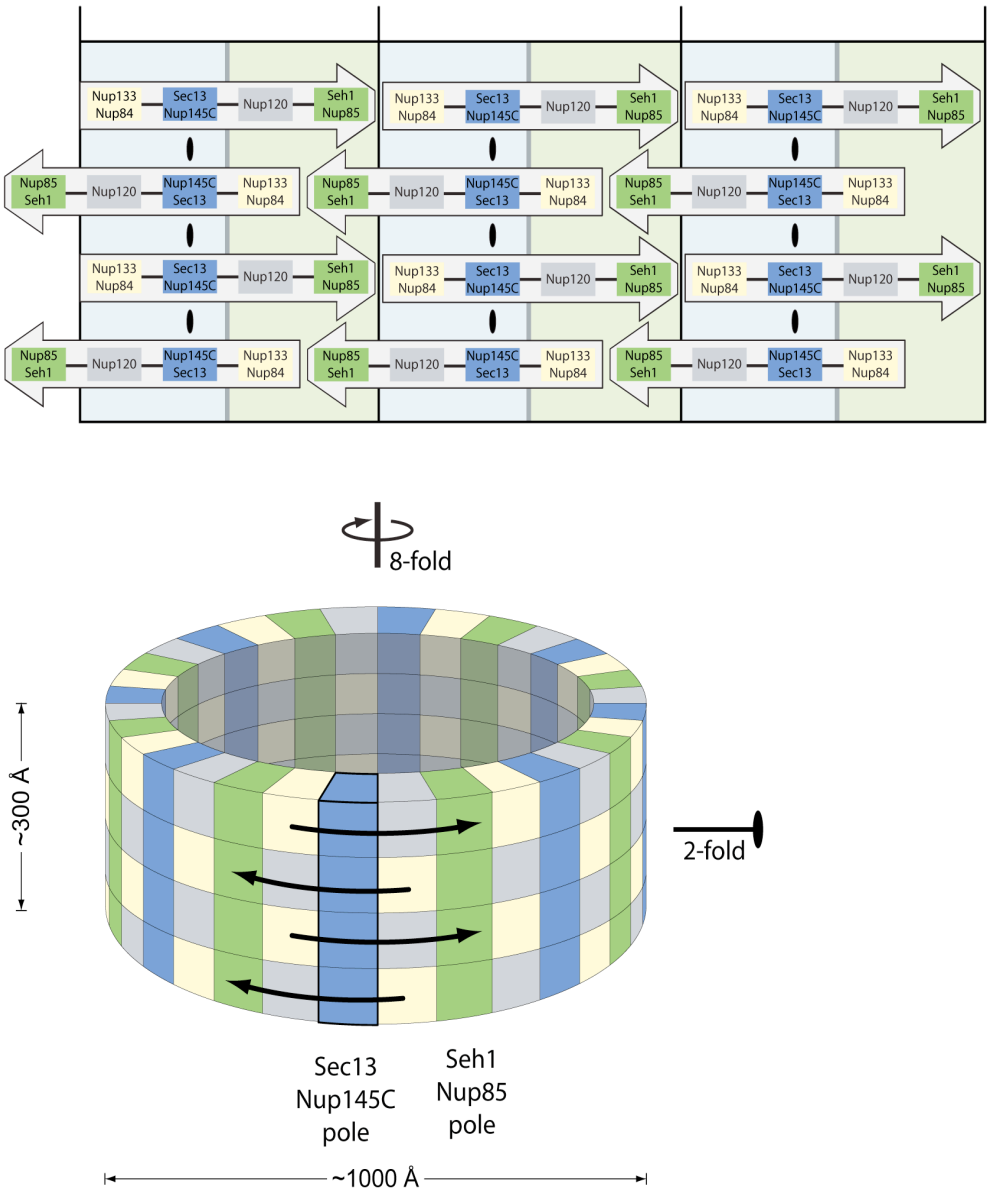
CHAPTER 4: DISCUSSION

For simplicity, the fragments hSec13$^{1-316}$ and Nup145pC$^{125-555}$ that were used for crystallization are referred to as Sec13 and Nup145C, respectively, in this section. We have determined the crystal structure of the Nup145C/Sec13 nucleoporin pair, a centrally located part of a previously characterized heptameric subcomplex of the NPC. Most strikingly, the Nup145C/Sec13 complex forms a hetero-octamer in two independent crystal forms as well as in solution. Interestingly, the behavior of the Nup145C/hSec13$^{1-322}$ and Nup145C/ySec13 complexes in solution is somewhat similar to the polymerization behavior of actin, highlighting the physiological significance of the oligmerization states of the Nup145C/Sec13 complex. Hence, based on the structural analyses, the solution behavior, and stoichiometric considerations, we proposed a model of how this hetero-octamer organizes the heptameric sub-complex into a peripheral cylindrical scaffold coating the pore membrane of the nuclear envelope (Fig 51).

Figure 51. Model for the architecture of a coat for the nuclear pore membrane

Eight heptamers are circumferentially arranged in a head-to-tail fashion in four stacked rings. The Nup145C/Sec13 and Nup85/Seh1 hetero-octamers serve as vertical poles connecting the four rings, thereby forming a scaffold. The poles are connected through their interaction with the remaining nups of the heptameric complex. Based on the two-fold axes of symmetry in the Nup145C/Sec13 hetero-octamer, the heptameric complex rings are stacked with opposite directionality.

Figure 51


Nup133
Nup84
Sec13
Nup145C
Nup120
Seh1
Nup85
Nup85
Seh1
Nup145C
Sec13
8-fold
~300 Å
2-fold
Sec13
Nup145C
pole
Seh1
Nup85
pole
~1000 Å

4.1 Controlling of Nup145C/Sec13 oligmerization states

The physiological functions of actin, a structural component of the cytoskeleton and a regulator of numerous cellular processes, relies on the existence of a dynamic equilibrium between the monomeric G-actin and the polymeric F-actin form. In order to obtain actin crystals, actin has to be kept in the monomeric state. To satisfy this requirement, actin was crystallized with its binding proteins, organic toxins, or hydrophobic probes such as tetramethyl-rhodamine, which can prevent its polymerization. More recently, Klenchin et al. showed that once the ~40 N-terminal residues of actin were removed by a specific protease, actin is non-polymerizable in the Ca-ATP bound form and could thus be crystallized in the monomeric state (Klenchin et al. 2006).

Akin to actin, Nup145C/full-length hSec13$^{1-322}$ also polymerizes in solution albeit to a substantially smaller extent. However, once the six residues from the C-terminus of hSec13 were removed and the resulting fragment co-expressed with Nup145C, the Nup145C/hSec13$^{1-316}$ complex showed primarily a single peak in the gel filtration profile, while higher-order oligomers were barely detected. The reduction of the oligomerization tendency most likely improves the quality of the crystals and lowers the

difficulty of crystallization.

In addition, a higher-order oligomer peak could also be observed for the Nup145C/ySec13 complex in solution, albeit to a relatively small amount. Hence, even though Nup145C/ySec13 crystals were obtained, they always had a large unit cell dimension instead, reflecting the smaller extent of oligomerization.

4.2 Structure of the Nup145C/Sec13 hetero-octamer

The key aspect of our model is the finding that Nup145C/Sec13 could form a hetero-octamer, with the heterodimer being the basic unit. The N-terminal portion of Nup145C invades the six-bladed Sec13p β-propeller, thereby providing a seventh blade. The remaining portion of Nup145C forms a domain composed of eighteen anti-parallel α-helices. A sharp kink in the middle of this α-helical domain results in an overall U-shaped structure.

The Nup145C/Sec13 dimer forms a heterotetramer via a head-to-head association involving the kink region of the two Nup145Cs. Finally, two heterotetramers associate into a hetero-octamer via the homodimerization of Sec13 and the interaction of Sec13 with the descending arm of the U of a "trans" Nup145C molecule. Importantly, Sec13 is able to

also form a dimer in solution, supporting the critical role that Sec13 plays in the hetero-octamerization.

The intermolecular interactions within the hetero-octamer involve large, primarily hydrophobic surfaces. Strikingly, all of these interfaces are evolutionarily highly conserved and, therefore, are likely to be physiologically relevant. However, it remains to be established whether this hetero-octamer also exists in the assembled NPC.

4.3 Model for the structural role of Nup145C/Sec13 as a pole in a pore membrane coat

The discovery of the Nup145C/Sec13 hetero-octamer and the elucidation of its structure yielded important insights into the architecture of the coat adjacent to the pore membrane of the nuclear envelope. Based on the order of the Sec13 and Nup145C molecules in the hetero-octamer and the fact that the hetero-octamer forms a bent rod of about 285 Å with three approximately equidistant axes of two-fold symmetry, we propose that the hetero-octamer forms a vertical pole that connects four horizontally arranged rings, each composed of eight heptamers (Fig 51).

In order to satisfy the symmetry of the hetero-octamer, the heptamers in each ring must be arranged in a head-to-

tail fashion and the rings must be stacked on top of each other in an anti-parallel manner (Fig 51). Moreover, because the hetero-octamer is a slightly bent rod, we suggest that the concave surface of this rod follows the bent contour of the pore membrane, whereas the convex surface faces the next inward layer of nups.

Our model satisfies the outer dimensions of about 1000 Å in diameter and 300 Å in height and symmetry requirements (eightfold axis in the nucleocytoplasmic direction and a perpendicular two-fold axis in the plane of the nuclear envelope) of the symmetric core of the yeast NPC that were established by cryoelectron microscopy (Yang et al., 1998). Since each heptamer has been determined by negative-stain electron microscopy to be about 400 Å long (Lutzmann et al., 2002), eight heptamers would add up to a ring of 3200 Å circumference with a diameter of about 1000 Å. Furthermore, the length of the Nup145C/Sec13 hetero-octamer of ~285 Å matches the height of the core of the NPC (Yang et al., 1998). With respect to symmetry, the two-fold axis of symmetry passing through the tetramer-tetramer interface and the circumferential arrangement of eight heptamers in head-to-tail order in each of the four rings (Fig 51) satisfies the requirements for two-fold and eight-fold axes of symmetry, respectively.

4.4 The Nup85/Seh1 hetero-octamer crystal structure strongly supports the model of the pore membrane coat

As a corollary of the pore membrane coat model, Nup85/Seh1 was proposed to have a similar overall fold as Nup145C/Sec13 and to form an additional pole in the pore membrane coat (Debler et al., 2008).

Indeed, crystal structures of $Nup85^{1-566}$/Seh1 (referred to as Nup85/Seh1) confirmed that the Nup85/Seh1 heterodimer adopts the same basic architecture as the Nup145C/Sec13 heterodimer and revealed that Nup85/Seh1 forms oligomers of a compact heterotetramer under three vastly different crystallization conditions in three different crystal packing environments.

4.4.1 Architectural Overview of the Nup85/Seh1 structure

The yeast Nup85 construct comprising residues 1–570 was co-expressed with full-length yeast Seh1. For simplicity, the Nup85 fragment is referred to as Nup85 (Fig 52). The Nup85/Seh1 complex crystallized in three unrelated crystal forms. The structure of Nup85/Seh1 was determined by multiple isomorphous replacement with anomalous scattering (MIRAS) from the monoclinic crystal form; the structures of the two orthorhombic space groups were solved by molecular replacement.

In all three crystal forms, higher-order structures of the Nup85/Seh1 pair were observed; while the asymmetric units of the crystal forms 1 and 2 each contained an elongated, curved Nup85/Seh1 hetero-octamer with a molecular weight of approximately 400 kDa (Fig 53). Crystal form 3 harbored an elongated heterododecamer (~600 kDa) with a slight helical twist (Fig 54). These oligomers are assembled from compact Nup85/Seh1 heterotetramers via the same Seh1/Seh1 interface (Fig 53 and 54). The heterotetramer, in turn, is formed from two Nup85/Seh1 pairs that are related by a pseudo-two-fold symmetry via a large interface (Fig 53 and 54). Furthermore, the interfaces within the oligomers are evolutionarily conserved, suggesting their physiological relevance (Fig 55).

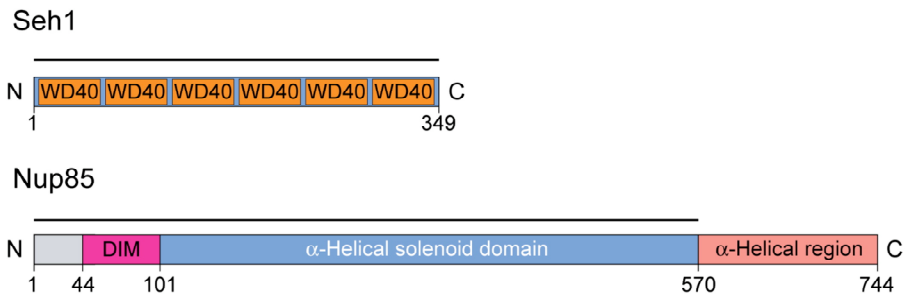
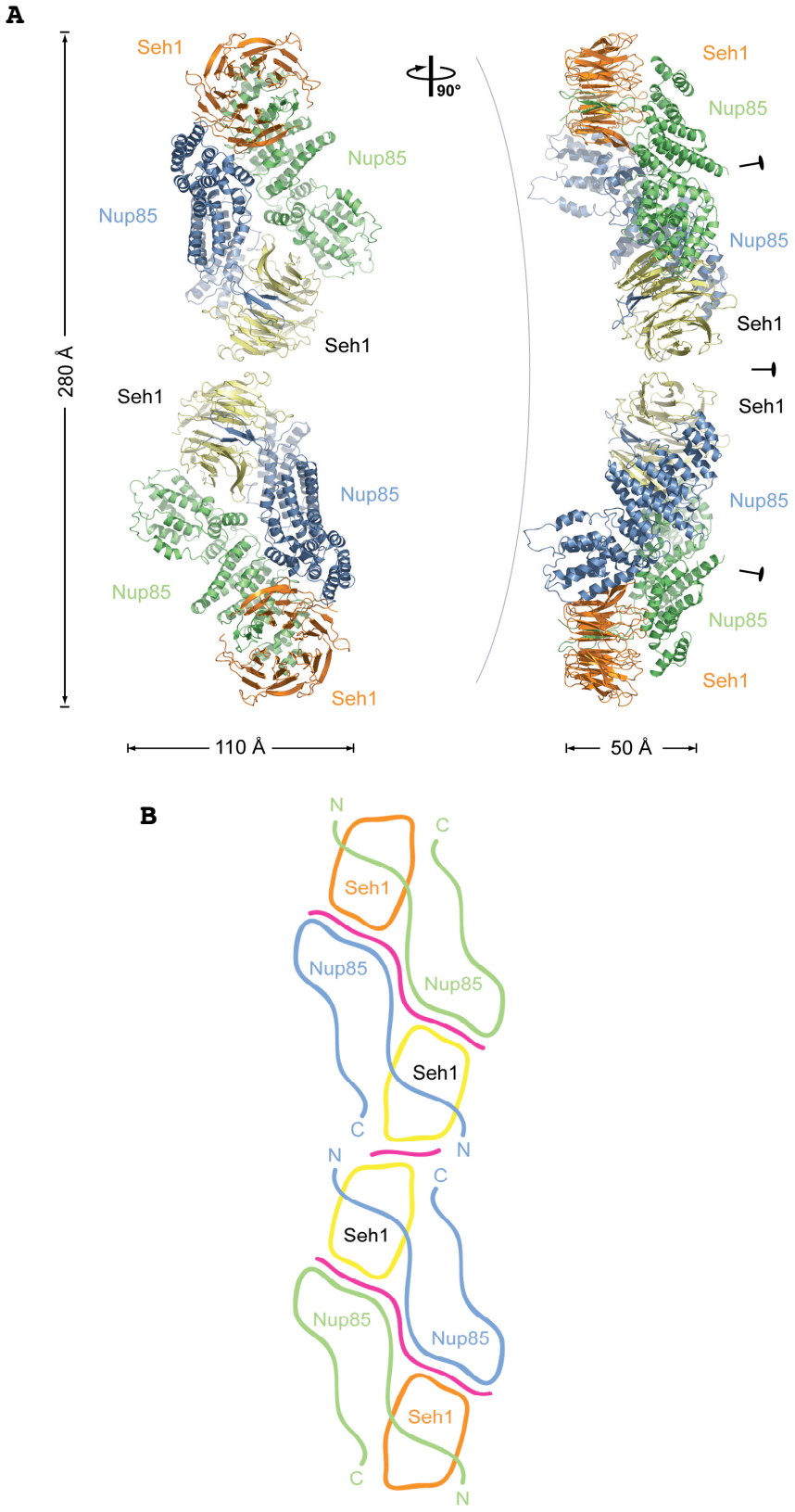


Figure 52. Domain structures of yeast Seh1 and Nup85 For Sec13, the WD40 repeats (orange) and the numbering relative to yeast Seh1 are indicated. For Nup85, the unstructured N-terminal region (gray), and the domain invasion motif (DIM) (magenta), and the α-helical domain (blue), and the C-terminal α-helical region (light pink) are indicated. The numbering is relative to yeast Nup85. The bars above the domain structures of the proteins mark the crystallized fragments. Image taken from Debler et al., 2008.

Figure 53. Overview of the structure of the Nup85/Seh1 hetero-octamer

(A) Ribbon representation of the Nup85/Seh1 hetero-octamer, showing Seh1 in yellow and orange and Nup85 in green and blue. A 90° rotated view is shown on the right. The three pseudo-two-fold axes (black ovals) and the overall dimensions are indicated. The Nup85/Seh1 hetero-octamer forms a slightly bent rod. (B) Schematic representation of the Nup85/Seh1 hetero-octamer. Magenta lines indicate interaction surfaces. Image taken from Debler et al., 2008.

Figure 53


A
Seh1
Nup85
Nup85
Seh1
Seh1
Nup85
Nup85
Seh1
280 Å
110 Å
90°
Seh1
Nup85
Nup85
Seh1
Seh1
Nup85
Nup85
Seh1
50 Å
B
N
C
Seh1
Nup85
Nup85
Seh1
C
N
N
C
Seh1
Nup85
Nup85
Seh1
C
N

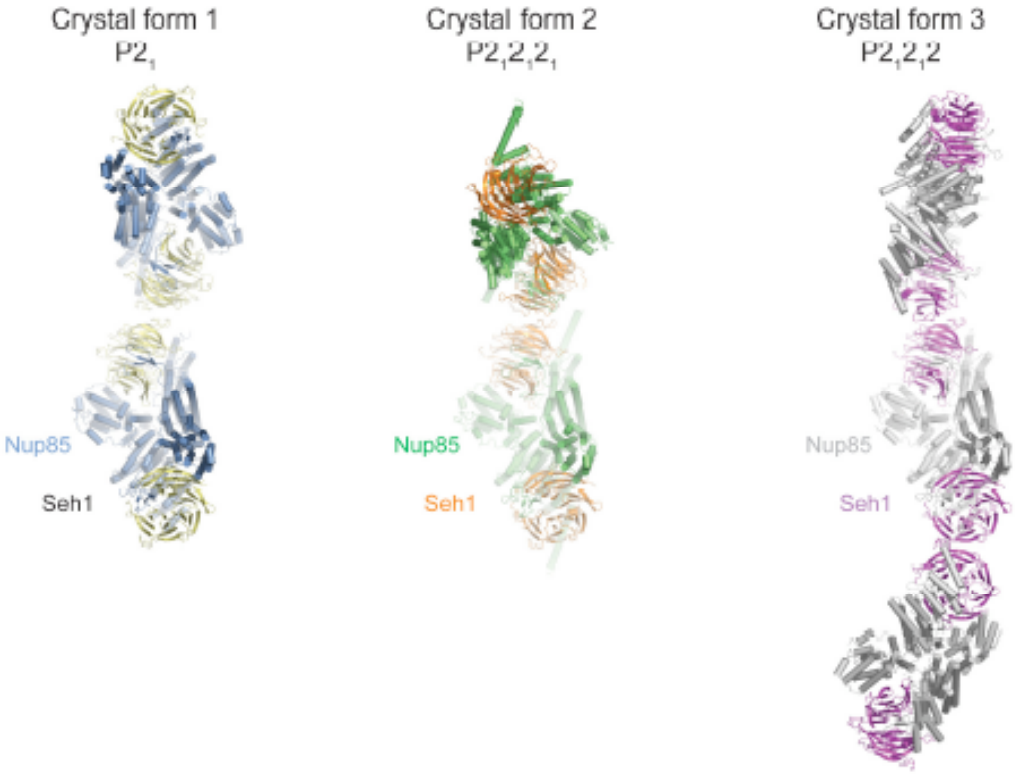


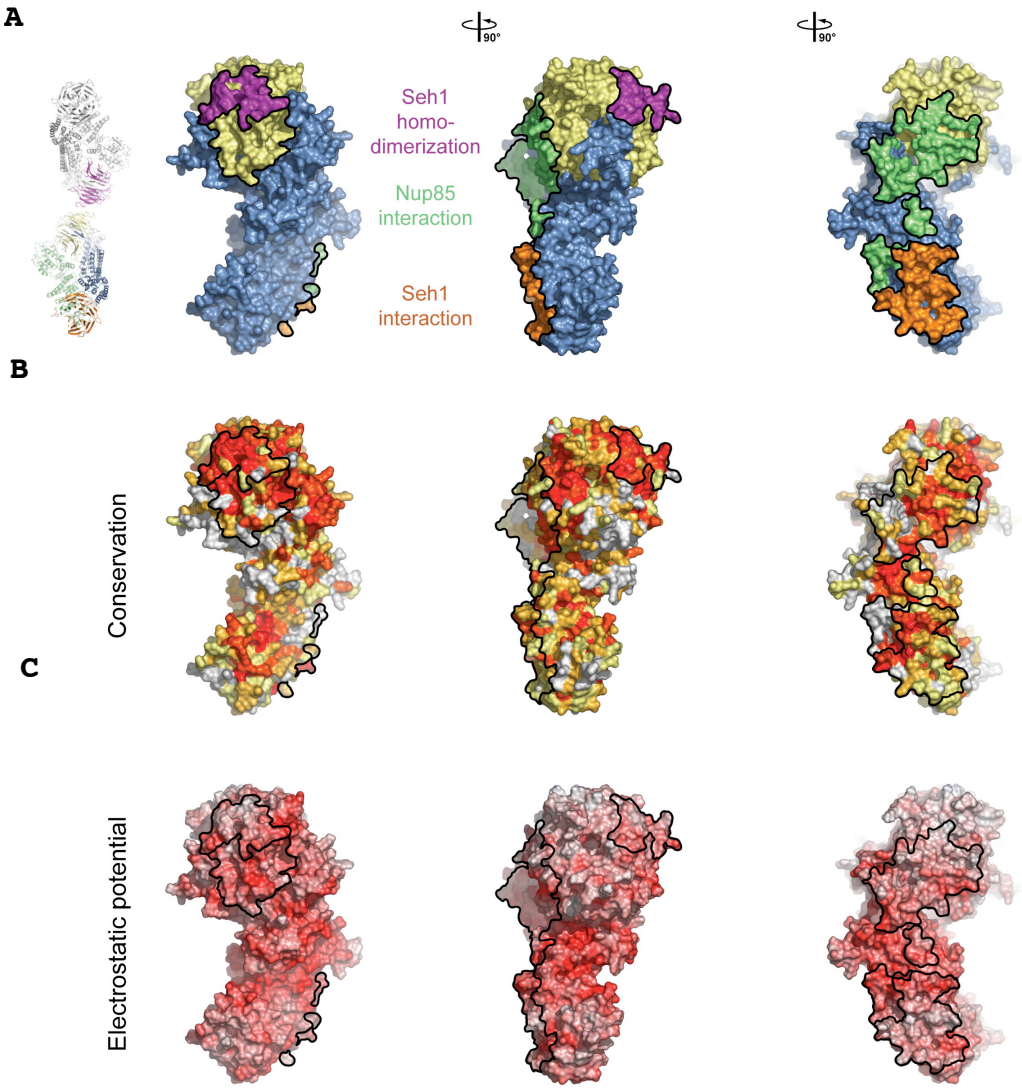
Figure 54. Three different crystal forms of the Nup85/Seh1 complex

The asymmetric units of crystal form 1 (left panel) and 2 (middle panel) each contain one hetero-octamer, while crystal form 3 (right panel) harbored one hetero-dodecamer. Image taken from Debler et al., 2008.

Figure 55. Surface properties of the Nup85/Seh1 heterodimer

(A) Surface rendition of the Nup85/Seh1 complex. The surface is colored according to the proteins (Seh1, yellow; Nup85, blue) and their participation in various interactions: with Seh1 of the adjacent heterodimer, orange; with Nup85 of the adjacent heterodimer, green; and with Seh1 of the adjacent heterotetramer, purple. (B) Nup85 is colored according to sequence conservation, from 40% similarity (yellow) to 100% identity (red). (C) Surface rendition of Nup85, colored according to the electrostatic potential, from red (-15 k_BT/e) to blue (+15 k_BT/e). Image taken from Debler et al., 2008.

Figure 55


A
90°
90°
Seh1
homo-
dimerization
Nup85
interaction
Seh1
interaction
B
Conservation
C
Electrostatic potential

4.4.2 Structure of Nup85

Nup85 adopts an α-helical fold in the form of a U-shaped solenoid structure that is preceded by a small domain invasion motif (DIM) (Fig 56). The ~470-residue solenoid domain consists of 22 α helices. Helices αA–αD form the descending part of the U, helices αE–αK pack into a six-helix bundle at the base of the U, helices αL–αQ form the lower ascending arm, while the helices αR–αV at the top of the ascending arm extensively interact with the bottom face of the Seh1 β-propeller. A salient feature is the extended 30-residue αQ–αR connector that contributes numerous contacts between two neighboring Nup85 molecules in the tetramer.

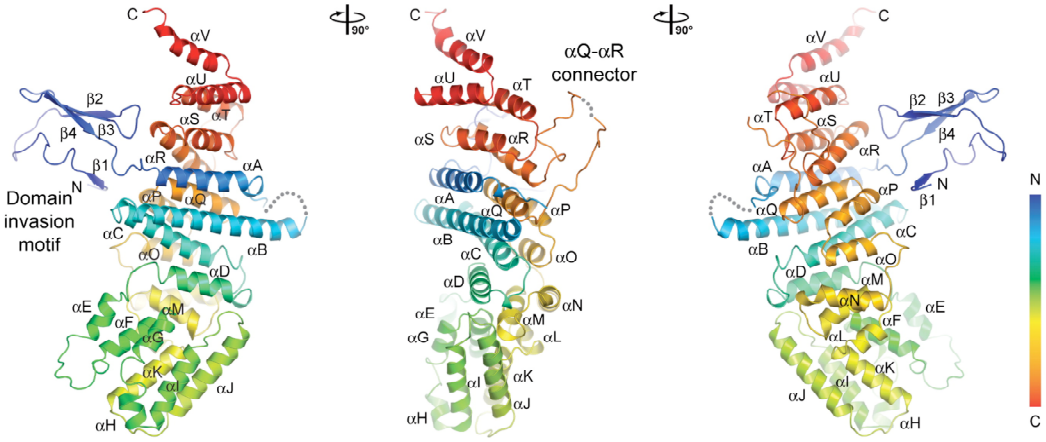


Figure 56. The structure of Nup85

Ribbon representation of the Nup85 structure is shown in rainbow colors along the polypeptide chain from the N- to the C-terminus. The N-terminal domain invasion motif (DIM), the α-helical solenoid domain, and their secondary structure elements are indicated. Image taken from Debler et al., 2008.

4.4.3 Structure of the β-propeller domain of Seh1 in complex with three-stranded blade of Nup85

The overall structure of the seven-bladed β-propeller, which is formed by Seh1 and the DIM of Nup85, conforms to the canonical β-propeller fold. The β-strands 6E of Nup85 and 7D of Seh1 each provide a Velcro closure for blades 6 and 7, respectively (Fig 57). Seh1 features two long loops, namely 2CD (the loop between strands C and D of blade 2) and 5CD. The 2CD loop mediates hetero-octamer and heterododecamer formation, respectively. By contrast, the large 5CD loop is not involved in any contacts with neighboring Seh1 or Nup85 molecules, and the major part of this surface-exposed loop is invisible in the electron density, presumably due to disorder. Altogether, 78 and 86 residues of Nup85 and Seh1, respectively, contribute to a large buried surface area of 5180 $Å^2$ at the heterodimer interface (Fig 58).

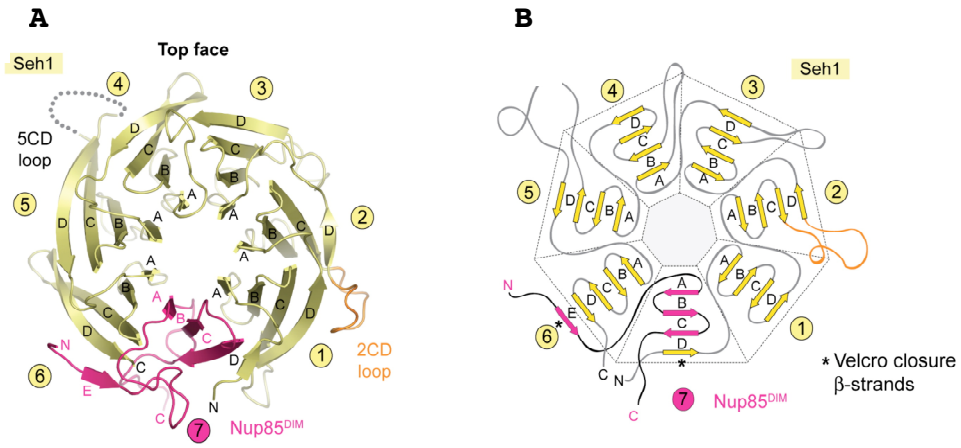


Figure 57. The structure of Seh1 β-propeller in complex with the Nup85^{DIM}

(A) The β-propeller domain of Seh1 in complex with the Nup85^{DIM}. Seh1 is shown in yellow, and the six blades are indicated. The Nup85^{DIM} contributes one strand to blade 6 and three strands to blade 7, completing the β-propeller. (B) Schematic representation of the Seh1 β-propeller and its interaction with the Nup85^{DIM}. Image taken from Debler et al., 2008.

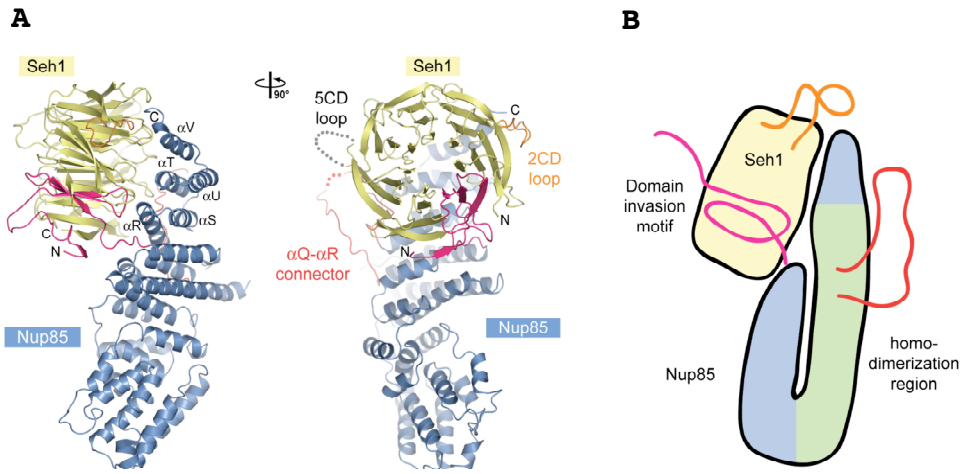


Figure 58. The structure of the Nup85/Seh1 hetero-dimer

(A) The structure of the Nup85/Seh1 heterodimer. The Nup85^{DIM} (magenta), the Nup85 α-helical solenoid domain (blue), the Nup85 αQ-αR connector (red), the Seh1 β-propeller (yellow), the disordered Seh1 5CD loop (gray dots), and the Seh1 2CD loop (orange) are indicated; a 90° rotated view is shown on the right. Dotted lines represent disordered regions. (B) Schematic representation of the Nup85/Seh1 interaction. The Seh1 2CD loop, the Nup85 αQ-αR connector, and the DIM region are highlighted in orange, red, and pink, respectively. Image taken from Debler et al., 2008.

4.4.4 Structure of the Nup85/Seh1 heterotetramer

In all three crystal forms, two Nup85/Seh1 pairs are tightly associated with each other longitudinally in an anti-parallel and complementary fashion. Superimposition of the seven crystallographically independent heterotetramers reveals that they are closely related with a root-mean square deviation of only ~1.1 Å over 1386 Cα atoms (Fig 59). Therefore, the heterotetramer structure appears to be neither influenced by the various crystal packing environments nor by the differing crystallization conditions. The dimer-dimer interface buries a huge surface area of approximately 5100 $Å^2$ that includes 101 residues on each heterodimer. Notably, the long αQ-αR connector extends to the neighboring Nup85 molecule and reinforces the interaction between the two nucleoporins pairs (Fig 60A). A large number of apolar residues impart a hydrophobic character to the upper and lower ends of the interface. Computational analysis of the heterotetrameric complex by the PISA server confirms that this assembly possesses high thermodynamic stability in solution (Krissinel and Henrick, 2007). Altogether, these findings suggest that Nup85/Seh1 heterotetramerization also occurs in vivo.

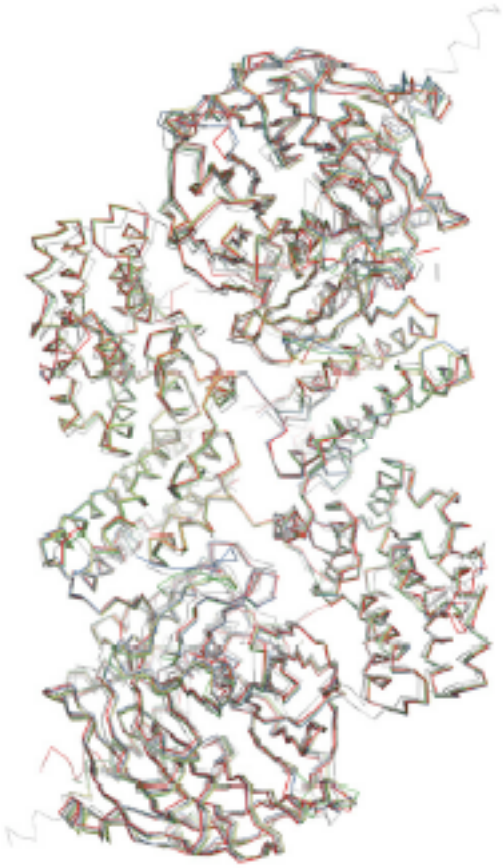


Figure 59. Superimposition of seven crystallographically independent Nup85/Seh1 heterotetramers of three crystal forms
The Nup85/Seh1 heterotetramers show little conformational plasticity and form a rigid unit. Image taken from Debler et al., 2008.

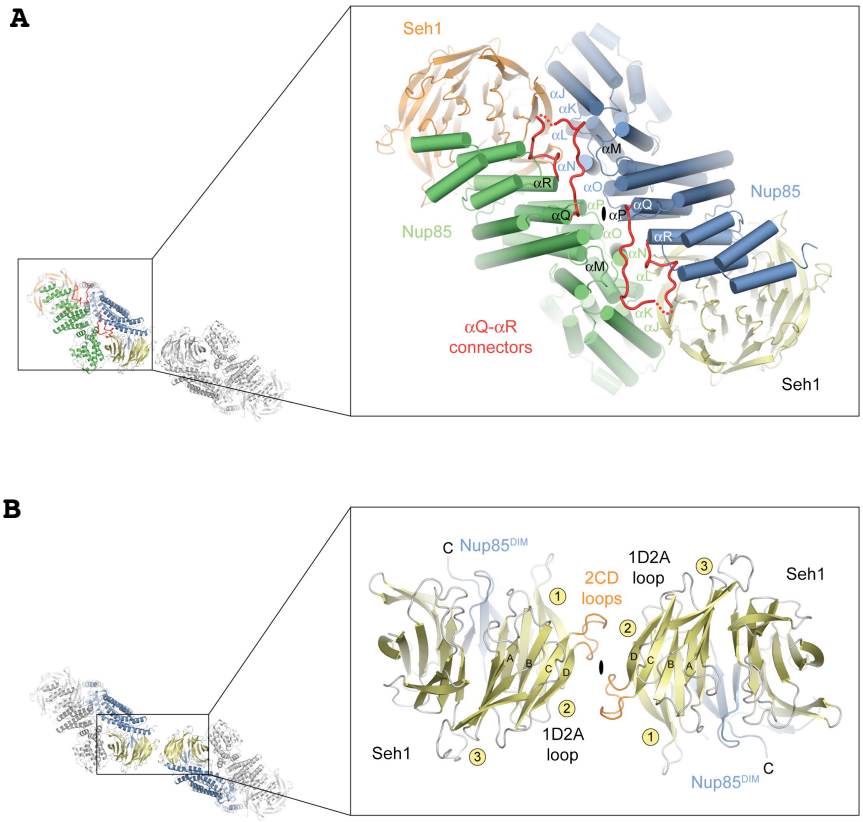


Figure 60. The interfaces of the Nup85/Seh1 hetero-octamer

(A) A large surface area of ~5100 Å^2 is buried upon heterotetramer formation and extends over almost the entire length of a Nup85/Seh1 heterodimer. α-Helices that mediate heterotetramer formation are indicated. The Nup85 αQ-αR connector segments (red) are located at the center and contribute many contacts to the interface. (B) The Seh1/Seh1 dimerization interface located at the center of the Nup85/Seh1 hetero-octamer is significantly smaller than the interface in (A). The two interacting Seh1 β-propeller domains and the two adjacent Nup85 molecules are colored yellow and blue, respectively. The locations of the pseudo-two-fold axes of symmetry that run through both interfaces are indicated (black ovals). Image taken from Debler et al., 2008.

4.4.5 Structure of the Nup85/Seh1 hetero-octamer

The hetero-octameric assembly of Nup85/Seh1 in crystal form 1 has overall dimensions of ~280 Å x ~110 Å x ~50 Å (Fig 53A). The association of two heterotetramers is exclusively mediated by the homodimerization of Seh1 (Fig 53). In detail, the 2CD loop contributes the major part to the interface, while the interblade loop 1D2A participates only to a minor extent (Fig 60B). Although only ~400 $Å^2$ of surface area are buried at this interface, the shape complementarity parameter Sc of 0.72 is high and falls into the range typically observed for protein oligomeric interfaces (0.70—0.74) (Lawrence and Colman, 1993). Importantly, 7 out of 16 Seh1 residues are charged and form several salt bridges at the interface that are likely to reinforce the hetero-octameric assembly via electrostatic interactions.

4.4.6 Comparison of the Nup145C/Sec13 and Nup85/Seh1 complexes

As suggested by the structure of Nup145C/Sec13, protein sequence alignment, and secondary structure predictions, the N-terminal portion of Nup85 inserts into Seh1, an opened six-bladed β-propeller, thereby providing the seventh blade. The rest of Nup85 likewise forms a U-shaped

structure composed of 22 helices. Although the solenoid domain in Nup85 is considerably twisted in the region most distant from the β-propeller, the overall heterodimer structures of Nup85/Seh1 and Nup145C/Sec13 complexes are highly similar (Fig 61).

Moreover, in analogy to the Nup145C/Sec13 hetero-octamer assembly, the Nup85/Seh1 heterodimer also assembles into a heterotetramer, which in turn can oligomerize into hetero-octamer and heterododecamers. (Fig 62). Detailed analysis of the dimer-dimer interface in the Nup85/Seh1 heterotetramers determined here uncovered an interface of more than 5000 $Å^2$ with characteristics typical of macromolecular complexes, further suggesting a physiological relevance of the observed oligomerization. Indeed, dimerization of the Nup85/Seh1 pair also occurs in solution, albeit to a small extent.

Furthermore, the two heterotetramers can also associate into a hetero-octamer through the dimerization of Seh1 (Fig 62). Based on the eight-fold rotational symmetry of the NPC (Hinshaw et al., 1992), experimental estimates for the copy number of Nup85 (30.4 ± 4) and Seh1 (27.2 ± 11.2) per yeast NPC (Rout et al., 2000) suggest a stoichiometry of 32 Nup85/Seh1 pairs in the NPC. A plausible higher-order structure of the Nup85/Seh1 complex in the NPC is

immediately gleaned from the Nup85/Seh1 complex arrangement in the three crystal structures, where adjacent Nup85/Seh1 heterotetramers are aligned linearly to form a heterooctamer. Strikingly, similar to the length of the Nup145C/Sec13 hetero-octamer, the length of the Nup85/Seh1 hetero-octamer is ~280 Å, in good agreement with the height of the yeast NPC of ~300 Å, determined by electron microscopy (Yang et al., 1998).

Although neither the crystalline state nor a diluted solution of the Nup85/Seh1 complex can faithfully mimic an environment of the Nup85/Seh1 complex in the NPC, the oligomeric state of the Nup85/Seh1 complex is likely to occur in the NPC, given the high local concentration of the Nup85/Seh1 complex, as well as the presence of neighboring nucleoporins that may stabilize a particular oligomerization state of the Nup85/Seh1 in the NPC. Altogether, the observation of an elongated, curved Nup85/Seh1 hetero-octamer fully supports our proposed architecture of a coat for the nuclear pore membrane.

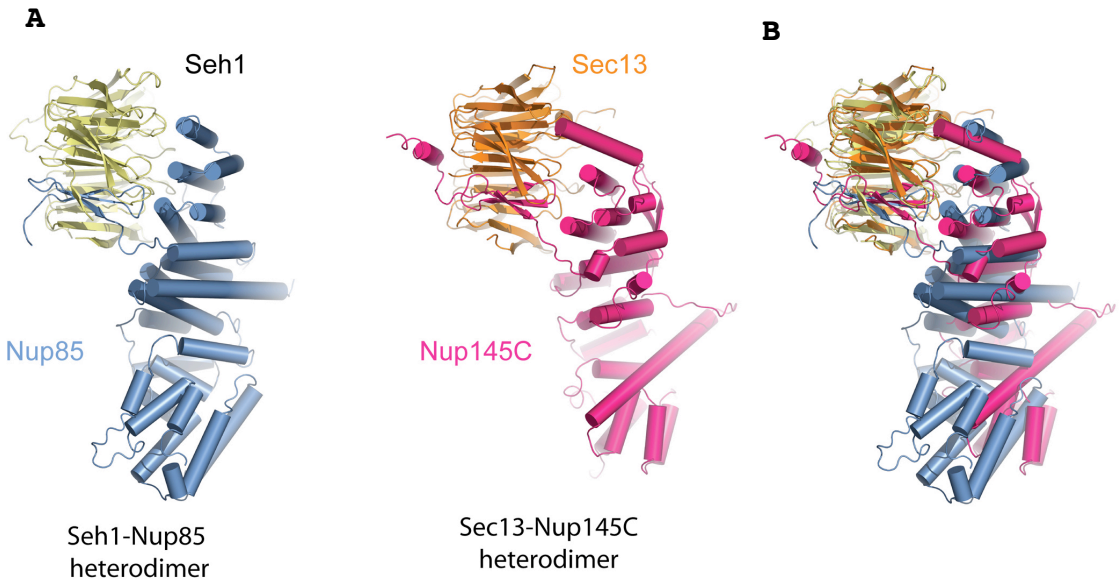


Figure 61. Superposition of Nup85/Seh1 and Nup145C/Sec13 heterodimer structures

(A) The Nup85/Seh1 complex (Left) contains similar structure features as in the Nup145C/Sec13 (Right). For instance, the DIM domain and α-helical domain also appeared in the Nup85, and the six blades Seh1 β-propeller is inserted by DIM domain from Nup85. (B) Superposition of Nup85/Seh1 and Nup145C/Sec13 heterodimer structures, based on the β-propeller. Image taken from Debler et al., 2008.

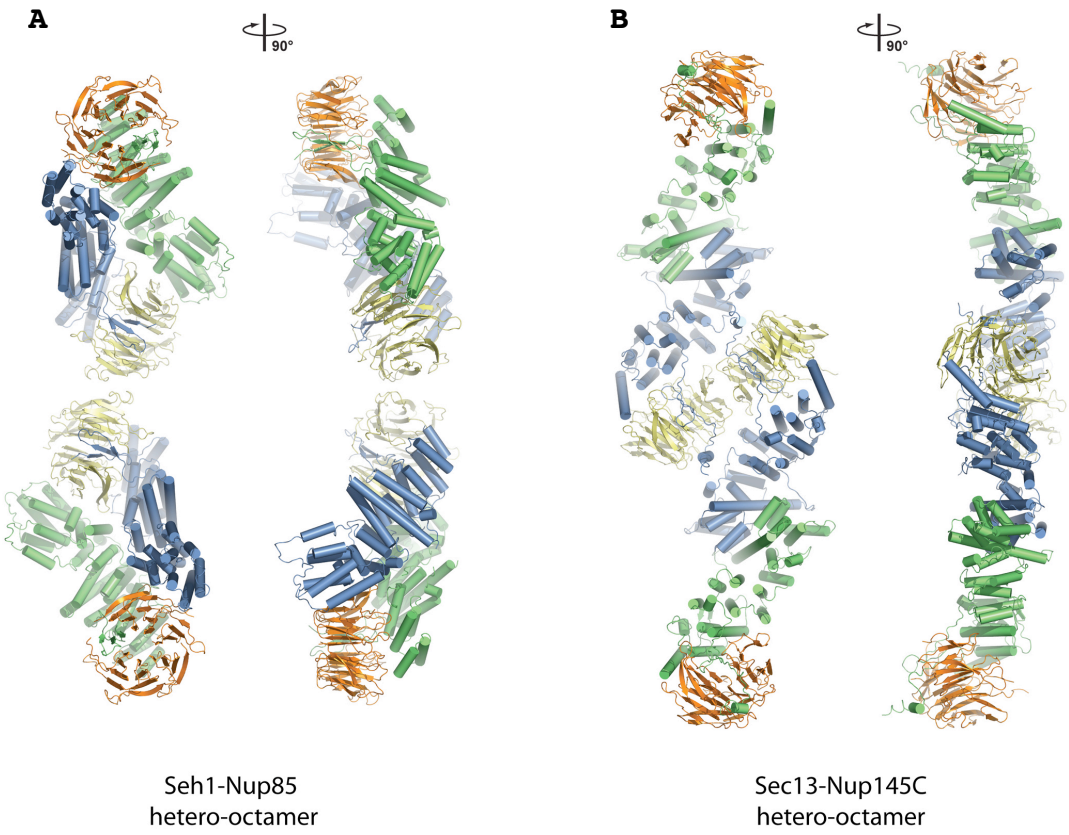


Figure 62. Overall structures of Nup145C/Sec13 and Nup85/Seh1 hetero-octamers

The Nup85/Seh1 (A) can form a hetero-octamer as well as Nup145C/Sec13 (B). Cartoon representations of both hetero-octamers, showing Seh1 and Sec13 in yellow and orange and Nup85p and Nup145C in green and blue. The Nup85/Seh1 hetero-octamer also contains three pseudo-two-fold axes and forms a slightly bent rod. Image taken from Debler et al., 2008.

4.5 Conformational flexibility of Nup85/Seh1 oligomers

Although hetero-octamers of the Nup85/Seh1 complex have been observed in three different crystal forms, these three hetero-octamers are not as similar to each other as the Nup145C/Sec13 hetero-octamers.

The two hetero-octameric assemblies of crystal forms 1 and 2 both form elongated, curved rods that are related by a 35° hinge motion around the center of the hetero-octamer. In crystal form 2, the buried surface area of tetramer-tetramer interface is slightly decreased to 340 $Å^2$. Structural plasticity is mainly provided by several residues with long but well-defined side chains, such as arginine, lysine, and glutamate, in the loop regions of the Seh1 (Fig 63A). While adjustments of several side-chain residues take place to accommodate the related hetero-octamer conformations, the involved loop regions of the interface maintain their main-chain conformations.

In crystal form 3, the heterododecamer possesses two identical tetramer-tetramer interfaces. Comparison of crystal forms 1 and 3 reveals that two adjacent heterotetramers in the heterododecamer display the same overall curvature but that one of the heterotetramers is rotated around its long axis by 80° (Fig 63B), now burying a surface area of 950 $Å^2$. The rigid body movement of the

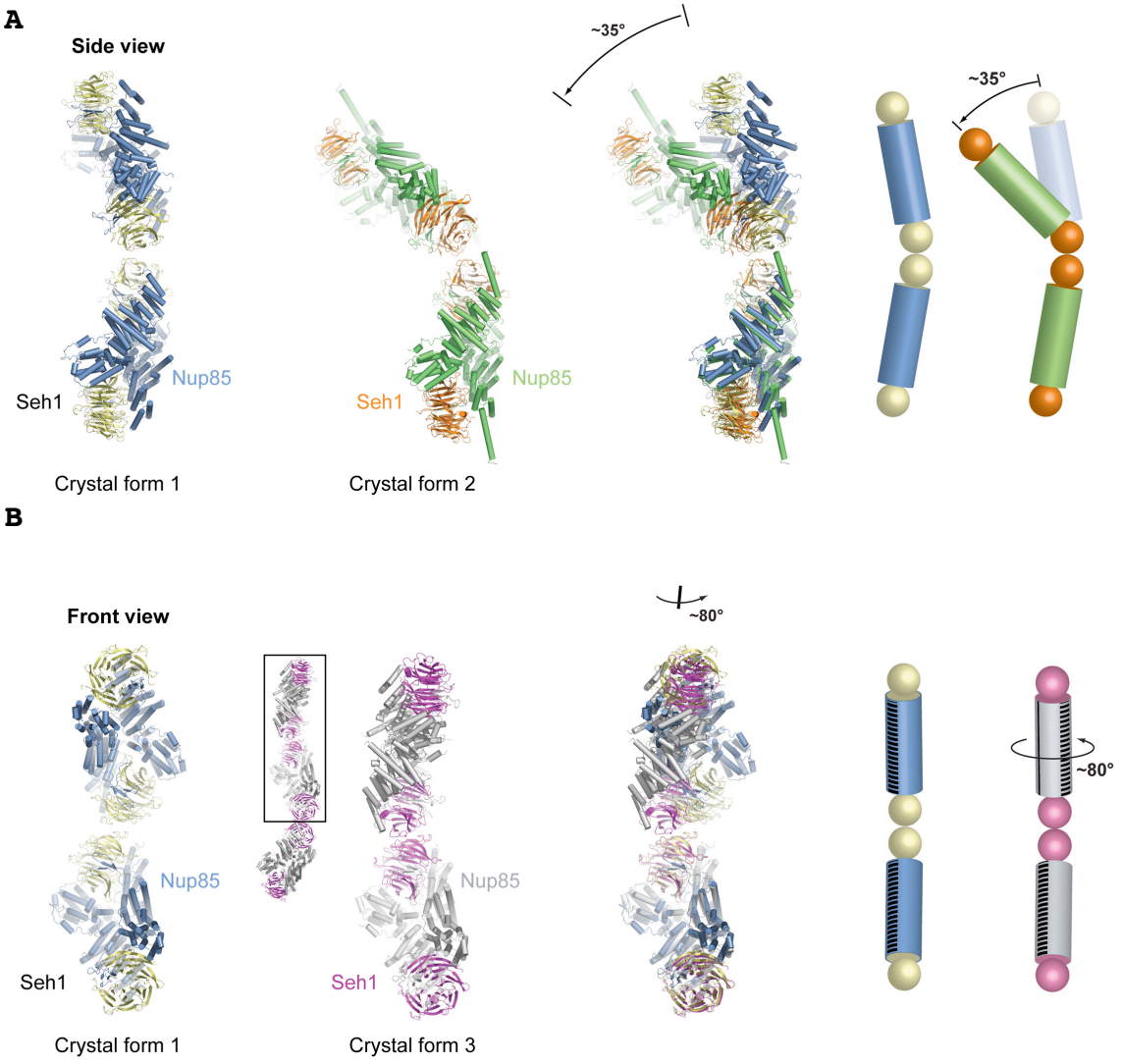
heterotetramer causes a register shift in the interface, whereby hydrogen bond donors and acceptors switch their partners, maintaining a similar hydrogen bond network in both structures. Interestingly, such intermolecular rearrangements were also observed in the Nup58/45 tetramer (Melcak et al., 2007)

Furthermore, we propose that one of two poles, Nup145C/Sec13, is relatively rigid in order to build and maintain the structure of the coat, while the other pole, Nup85/Seh1, possesses some flexibility which allows the coat to modify the shape, for example to facilitate the import of integral membrane proteins to the inner nuclear membrane (King et al., 2006). Taken together, Nup85/Seh1 forms a second vertical pole in the heptameric complex scaffold, yielding a total of 16 poles of the coat for nuclear pore membrane (Fig 51).

Figure 63. Flexibility of the Nup85/Seh1 hetero-octamer

(A) The hetero-octamers of crystal forms 1 and 2 are related by a 35° hinge motion around the center of the hetero-octamer. (B) Crystal form 3 harbors a heterododecamer in the asymmetric unit (small ribbon represention). The two interfaces between the heterotetramer are identical. In the upper two neighboring heterotetramers of crystal form 3 (boxed in the heterododecamer), one heterotetramer is rotated by ~80° around its long axis with respect to crystal form 1. Image taken from Debler et al., 2008.

Figure 63


A
Side view
Nup85
Seh1
Crystal form 1
Seh1
Nup85
Crystal form 2
~35°
~35°
B
Front view
Nup85
Seh1
Crystal form 1
Nup85
Seh1
Crystal form 3
~80°
~80°

4.6 The hybrid Nup85/Sec13 Complex implicates an architectural redundancy within the NPC

Deletion studies in yeast established that Seh1 is nonessential for cell growth and indicate that Sec13 could perform an overlapping or redundant function (Siniossoglou et al., 1996). Structure-based sequence alignment, in fact, demonstrates that most of the Seh1 residues that interact with Nup85 are either identical or conserved in Sec13, which would enable Sec13 to substitute for Seh1. In order to test this hypothesis, yeast Sec13 and Nup85 were co-expressed in *E. coli* and could be purified and characterized by analytical size exclusion chromatography (Fig 64A). In comparison with Nup85/Seh1, the significantly reduced amount of soluble protein in our bacterial expression system points toward a less stable Nup85/Sec13 complex. Consistent with this observation, the co-expression of Nup85 with Seh1 and Sec13 together exclusively resulted in Nup85/Seh1 complexes (data not shown). Finally, the co-expression of Seh1 and Nup145C did not yield a complex (data not shown), which coincides with the plasmid shuffle assay results. In the plasmid shuffle assay, over-expression of Seh1 was unable to rescue the inviability of Sec13 null mutant (Fig 64B). Hence, even though Sec13 can replace Seh1, Seh1 cannot substitute for

Sec13, probably due to the failure of Nup145C or Sec31 binding. By contrast, replacement of Seh1 with Sec13 can be tolerated on the molecular level, in agreement with the non-lethality of the genetic Seh1 deletion (Siniossoglou et al., 1996).

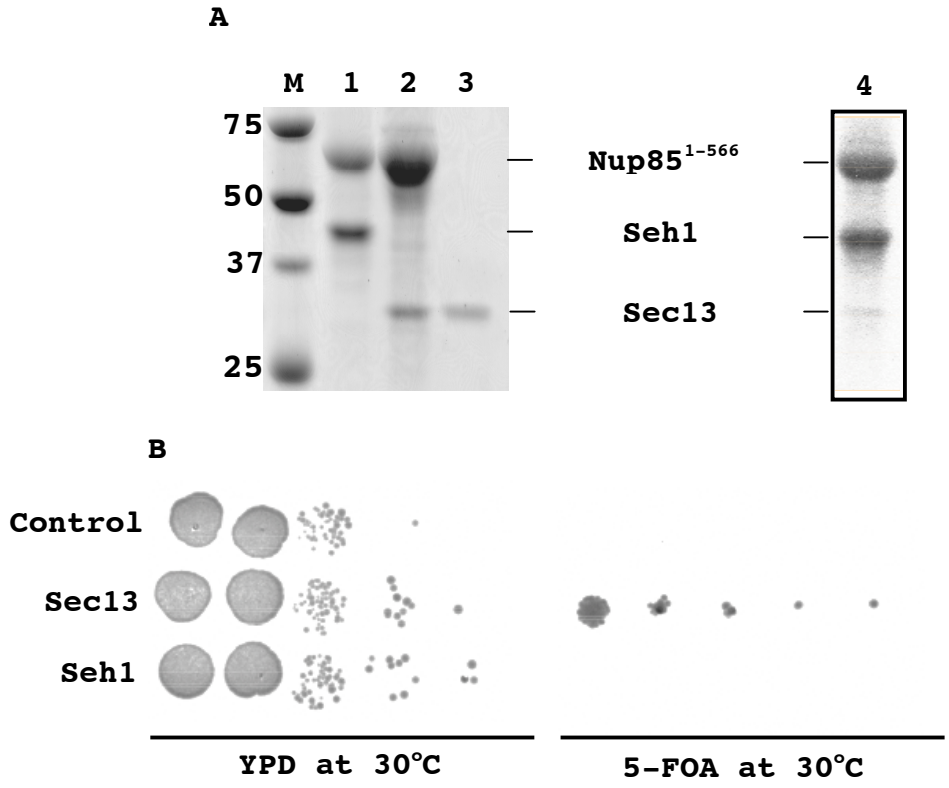


Figure 64. Sec13 can substitute Seh1 *in vivo* and *in vitro*

(A) Sec13 can form a complex with Nup85 by either co-expression with Nup85$^{1-566}$ (Lane 2) or Nup85$^{1-566}$/Seh1 (Lane 4). Nup85$^{1-566}$/Seh1 complex (Lane 1) and Sec13 (Lane 3) are control markers. (B) Yeast Sec13 shuffle strain (sec13::HIS3;pRS316-URA-SEC13) was transformed by control, pRS414, and pRS414 containing Sec13 or Seh1. Different transformants were spotted at 10-fold serial dilutions (from ~10^4 to 10) on the plates of YPD or minimum medium containing 5-FOA.

4.7 Further implications of the model

4.7.1 A second choice of arranging the Nup145C/Sec13 hetero-octamer?

An alternative arrangement of the Nup145C/Sec13 hetero-octamer has been considered, where the bent rod would be rotated: the long axis by 90°, placing it in the plane of the nuclear envelope, and the short axis by 180°, whereby the convex surface would face the center of the pore. To satisfy the symmetry and the dimension of the NPC core, the hetero-octamer must further polymerize so as to obtain a closed ring with a diameter of ~1000 Å. However, the hetero-octamer polymerization, primarily via Sec13 homodimerization (Fig 65A), yields a stretched-out spiral rather than a closed ring, with an outer diameter of only ~300 Å (Fig 65B).

In the evolutionary perspective, the vesicle-coated complexes, COPI, COPII, and clathrin and the Nup84 sub-complex all share similar structural modules for covering the curved membrane. Moreover, the vesicle is sphere-shaped, which possesses only a convex curvature in all directions. Thus, the structural modules in the vesicle coated complexes are supposed to be designed for coating the convex membrane curvature only.

In the membrane pore domain, a concave curvature also occurs in addition to the convex curvature. However, since the Nup84 sub-complex was proposed to have evolved from a common ancestor as well as vesicle-coated complexes (Devos et al., 2004), the structural modules of the Nup84 sub-complex are also designed to cover the same surface topology, convex curvature, as vesicle-coated complexes.

In addition, according to structural analyses, the Nup85/Seh1 complex also forms a hetero-octamer rod and this organization is only in agreement with the proposed coat. Altogether, this alternative arrangement is likely physiologically irrelevant.

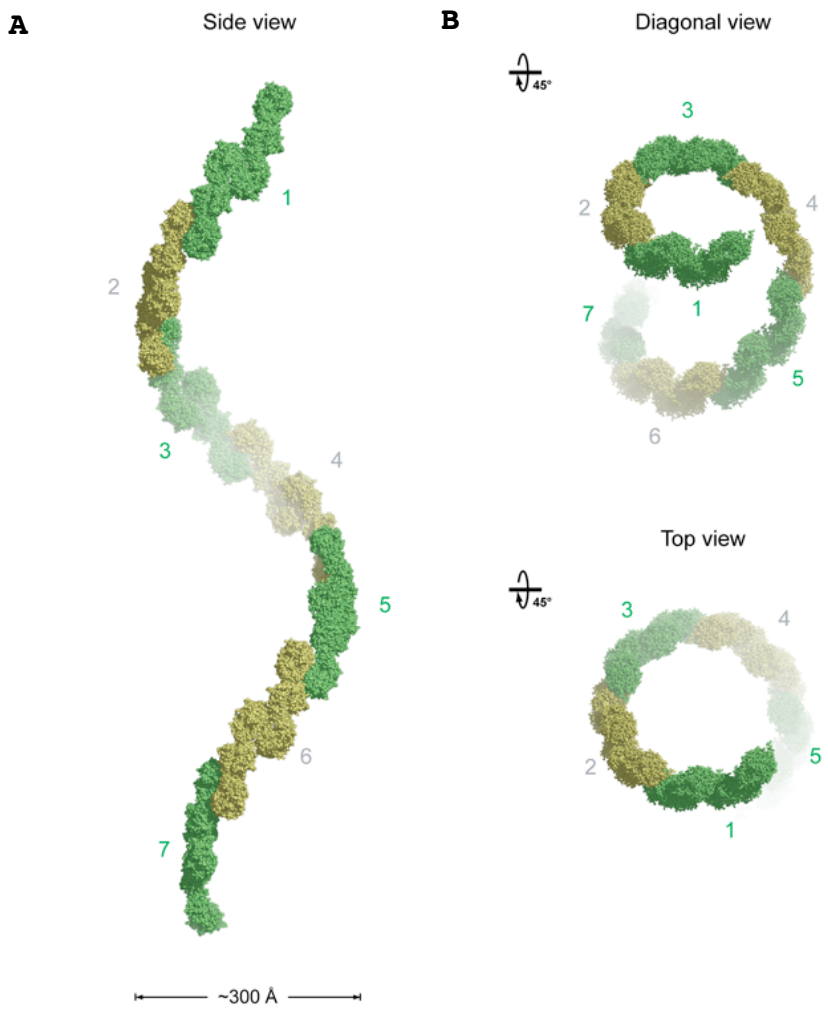


Figure 65. Model for the higher-order oligomerization of the Nup145C/Sec13

(A) The propagation of the oligomerization of the Nup145C/Sec13 hetero-octamer does not result in the formation of a closed ring, but in a spiral with an outer diameter of ~300 Å. The Nup145C/Sec13 hetero-octamers are colored in alternating yellow and green. (B) Diagonal and top views of the spiral. The model was generated by the sequential superpositioning of a terminal Sec13 protomer over a centrally located Sec13 protomer in order to generate an identical Sec13-Sec13 homo-dimerizaton interface.

4.7.2 Stoichiometric considerations of the Nup84 sub-complex complex in the NPC

In addition to the coat model we proposed, two models also have been proposed by Alber et al. and Brohawn et al. to describe how Nup84 sub-complex arranges in the NPC.

4.7.2.1 Model 1: A computational model

According to immuno-electron microscopy (EM) studies, proteomic pull-down assays, and Western blotting analysis, the Nup84 sub-complex has been proposed to locate on both sides of the NPC periphery and it was suggested to exist in 16 copies per NPC (Cronshaw et al., 2002; Rout et al., 2000). In addition, Alber et al. developed a computational method to convert these biophysical and proteomic data into the structures of the NPC assembly. Based on this computational model, the Nup84 sub-complex was suggested to form two concentric eight-membered rings, which are also referred to as outer ring, on the nucleoplasmic and cytoplasmic faces of the NPC (Alber et al., 2007) (Fig 66). These two outer rings are linked by two additional coaxial inner rings, containing Nup157, Nup188, Nup170, and Nup192. Consequently, these fours rings can form a core scaffold of the NPC and provide attachment sites for other nucleoporins, such as POMs and FG nups.

Notably, for immuno-EM studies, antibodies that carry gold particles might not be able to access the epitopes in the center of the NPC due to the size of antibodies. Hence, the distribution of the gold particles is heavily biased towards the periphery instead of the center of the NPC. Moreover, regarding the number of the Nup84 sub-complex in the NPC, it is clear that these semi-quantitative determinations, such as Western blotting, can only yield nup ratios instead of absolute numbers at best. Therefore, the stiochiometry as derived from these techniques remains approximate.

In addition, the oligomerization states of Nup145C/Sec13 and Nup85/Seh1, which could be observed in crystals and solution, do not exist this model. How Nup84 sub-complex interacts in order to form a ring structure is not shown in this model.

4.7.2.2 Model 2

Lately, another model on the arrangement of the Nup84 sub-complex in the NPC was proposed (Brohawn et al., 2008). In this model, the Nup84 sub-complex, a Y-shaped complex, was placed vertically with respect to the plane of the nuclear envelope and assembled to form two concentric

peripheral rings. The extended arm of the Y composed of Nup145C, Sec13, Nup84, and Nup133 faces outward; the pentamer, Nup85, Nup145C, Sec13, Seh1 and Nup120, forms a roughly symmetrical triskelion, resembling the vertex elements in the clathrin polygonal cage (Fig 67).

Together with the Nic96 complex, the Nup84 sub-complex would generate a lattice-like coat for the NPC similar to clathrin and COPII coats. However, if the NPC is built based on this model, the height would be significantly larger than 800 Å and the diameter of the pore would be just ~650 Å, which is completely different from experimentally determined pore dimensions. Hence, how the Nup84 sub-complex arranges in the NPC is still unknown at present.

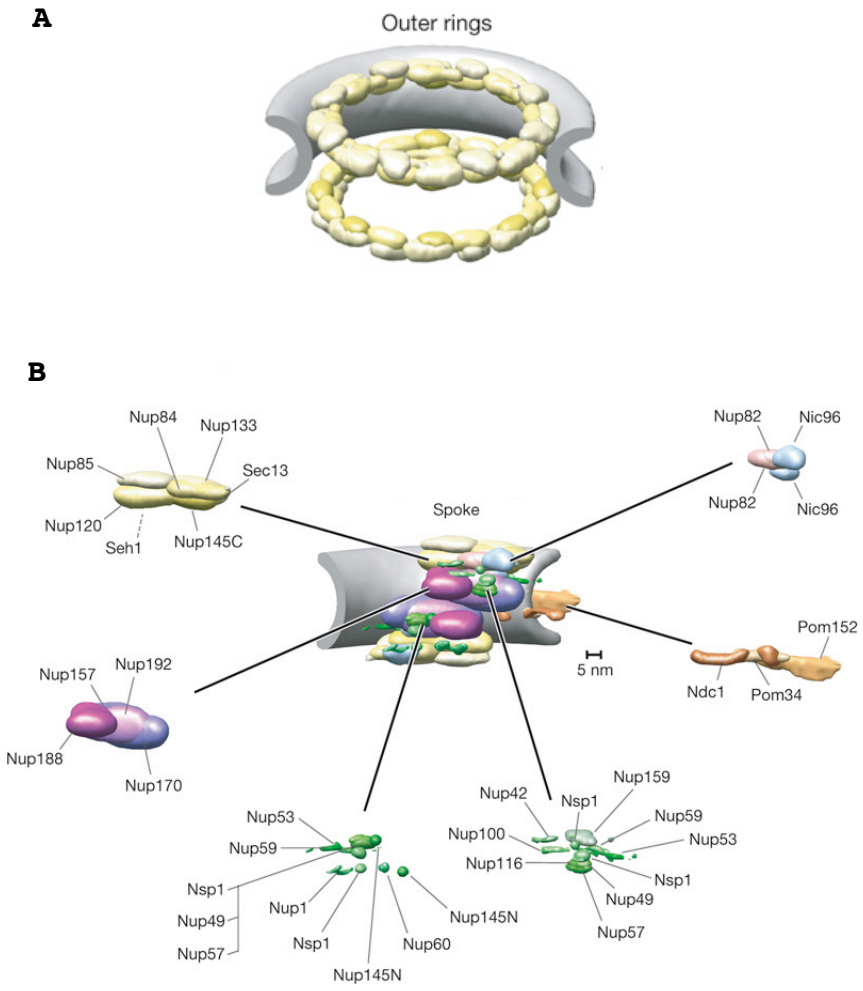


Figure 66. Localization of the Nup84 sub-complex and its components in the NPC

(A) According to a computational model, the Nup84 sub-complex is placed in two concentric eight-membered outer rings colored yellow. (B) The locations of the components of the Nup84 sub-complex (yellow): Nup85, Nup145C, Nup120, Nup84, Nup133, Seh1, and Sec13. Image taken from Alber et al., 2007.

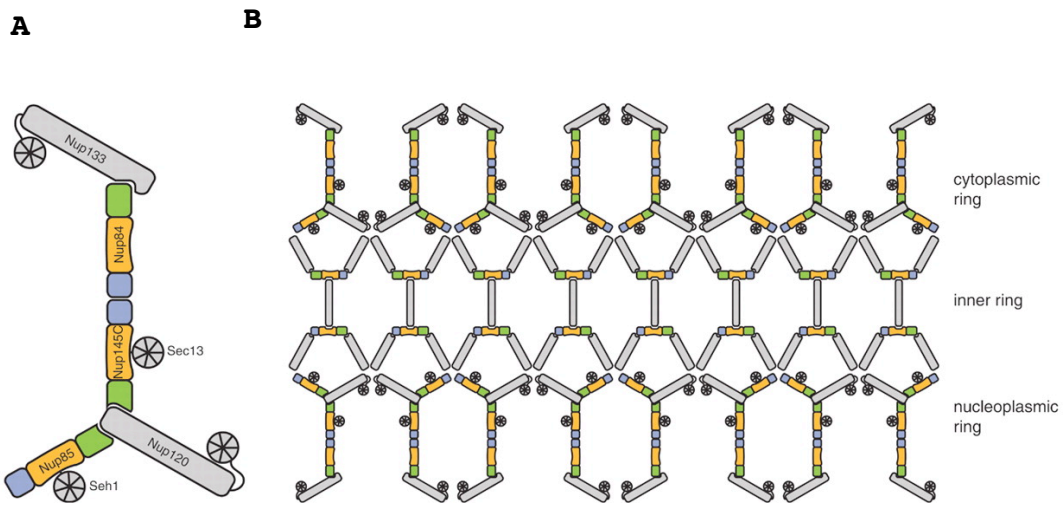


Figure 67. A lattice model of the Nup84 sub-complex in the NPC

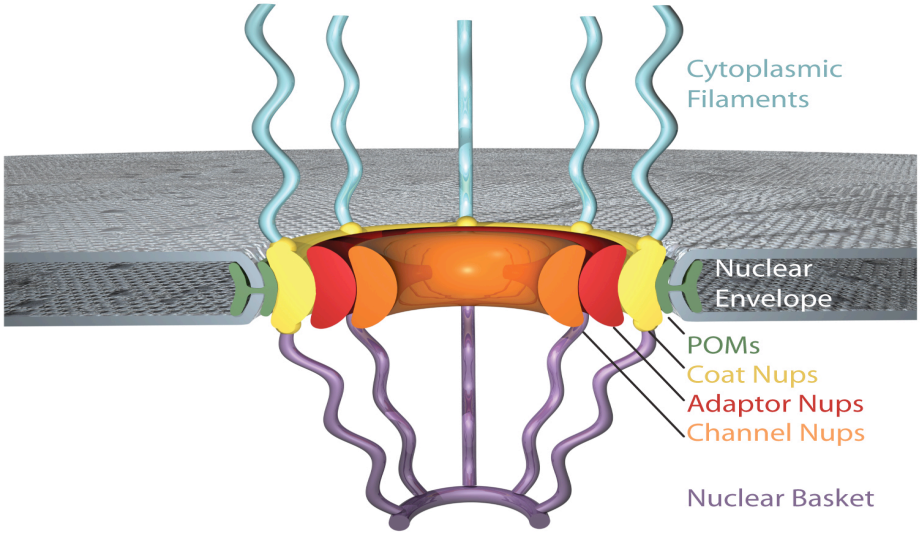
(A) Nup85, Nup145C, and Nup84 are proposed to contain a tripartite structural element of crown (blue), trunk (Orange), and tail (green), termed ancestral coatomer element 1 (ACE1). Their relative positions and interactions in the Nup84 sub-complex are shown with Sec13, Seh1, Nup133, and Nup120 colored in gray. (B) The Nup84 sub-complex is placed vertically in the plane of the nuclear envelope and forms a nuclear and a cytoplasmic ring, whereas the Nic96 complex makes up the inner ring. Image taken from Brohawn et al., 2008.

4.8 Architecture of the NPC core as a series of concentric cylinders

The symmetric core of the NPC can be considered as a series of concentric cylinders (Hsia et al., 2007). In our highly schematic model (Fig 68), the NPC core is illustrated as a series of concentric cylinders and the nucleoporins in each cylinder have distinct functions, involved in the architectural assembly of the NPC or NPC-coupled biological processes. For instance, the innermost cylinder represents the conduit for nucleocytoplasmic transport and is envisaged to be composed of the α-helical domains of channel nups. An atomic structure of two α-helical domains of two representative channel nups has recently been established (Melcak et al., 2007). These nups contain "sliding" α-helices that may allow variation of the diameter of the central transport cylinder according to cargo size. Because of the dynamic aspect of the components of the innermost transport cylinder, we further envisage that the nups that are sandwiched between the transport and the coat cylinder would function as adapters, cushioning the dynamic variations in the transport cylinder diameter.

The concept of a series of concentric cylinders does not mean that the cylinders should be considered as "solid" walls. Rather, we envisage porous layers that would allow

interdigitation and invasion of nup regions from adjacent cylinders. For example, the sixteen poles of the coat cylinder could form the vertical poles of a fence, while the three remaining nups of the coat cylinder, Nup120, Nup84, and Nup133, would form the horizontal connections, resulting in a picket fence-like coat. Domains of integral membrane proteins of the pore membrane cylinder and the adaptor nup cylinder could then protrude through the openings of the fence.



POMs	Coat Nups	Adaptor Nups	Channel Nups	Filaments	Basket
Pom152	Seh1	Nic96	Nsp1	Nup82	Nup60
Pom34	Nup85	Nup192	Nup49	Nup159	Nup2
Ndc1	Nup120	Nup188	Nup57	Nup116	Nup1
	Sec13	Nup157		Nup100	
	Nup145C	Nup170		Nup145N	
	Nup84	Nup53		Nup42	
	Nup133	Nup59			

Figure 68. Model for the architecture of the symmetric NPC core

The symmetric core of the NPC is schematically represented as a series of concentric cylinders. Each of the cylinders would contain the principal mass of the nucleoporins according to their functions and locations. For instance: integral pore membrane proteins (pom152, pom34, and Ndc1), coat nups (Seh1, Nup85, Nup120, Sec13, Nup145C, Nup84, and Nup133), adaptor nups (Nic96, Nup192, Nup188, Nup157, and Nup170), and the channel nups (Nsp1, Nup49, and Nup57). Image kindly provided by Andre Hoelz.

4.9 Comparison with the COPII complex

4.9.1 Structural comparison of Nup145C/Sec13 and Sec31/Sec13 complexes

One of the structural similarities between the Nup145C/Sec13 and the Sec31/Sec13 heterodimers is the six-bladed Sec13 β-propeller complemented by an additional blade provided by Nup145C or Sec31. Interestingly, the Sec13 β-propeller is capable of accommodating both DIMs, despite the sequence differences in the DIMS. Although both α-helical domains of Nup145C and Sec31 can dimerize, the folding and topology of the solenoids are distinct. Comparison between the two structures further revealed that the Nup145C/Sec13 is overall more compact (Fig 69A). In addition, the α-helical domain of either protein extends with different angles from the Sec13 β-propeller, resulting in an ~45° rotation with respect to their long axis (Fig 69A). Moreover, dimerization of the Sec31/Sec13 heterotetramer via Sec13, as it occurs in the Nup145C/Sec13 hetero-octamer, is sterically prevented in the Sec31/Sec13 complex (Fig 69B).

The interacting areas between Sec13 and the α-helical domains of Nup145C and Sec31 are also unrelated (Fig 70 and 71). The residues of three α-helices in Nup145C, αP, αR,

and αQ (Fig 38), interact with Sec13, whereas only one α-helix, α18, of the α-solenoid domain of Sec31 is involved in Sec13 binding (Fath et al., 2007). Since the N-terminal β-propeller domain of Sec31 is also responsible for Sec13 binding, the α-helical domain of Sec31 thus does not employ as many residues as Nup145C for Sec13 binding.

Moreover, the protein sequences of human and yeast Sec13 were aligned and the residues in ySec13 were compared with the residues of hSec13 that contact Nup145C. Interestingly, the interaction areas and residues in Sec13 appear to be fairly different (Fig 71). Besides some common areas for both Nup145C and Sec31 binding, mainly involving residues that bind to the extra blade, areas II and IV were of particular interest in Nup145C and Sec31, respectively. Unlike most of the Nup145C contact residues, which are highly conserved, the residues are more diverse in two of the Sec31 contact areas, area I and II. Hence, the binding affinity between hSec13 and Sec31 probably is lower due to the decreased conservation in these two areas, resulting in the in-viability of yeast in the Sec13 shuffle assay (Fig 45).

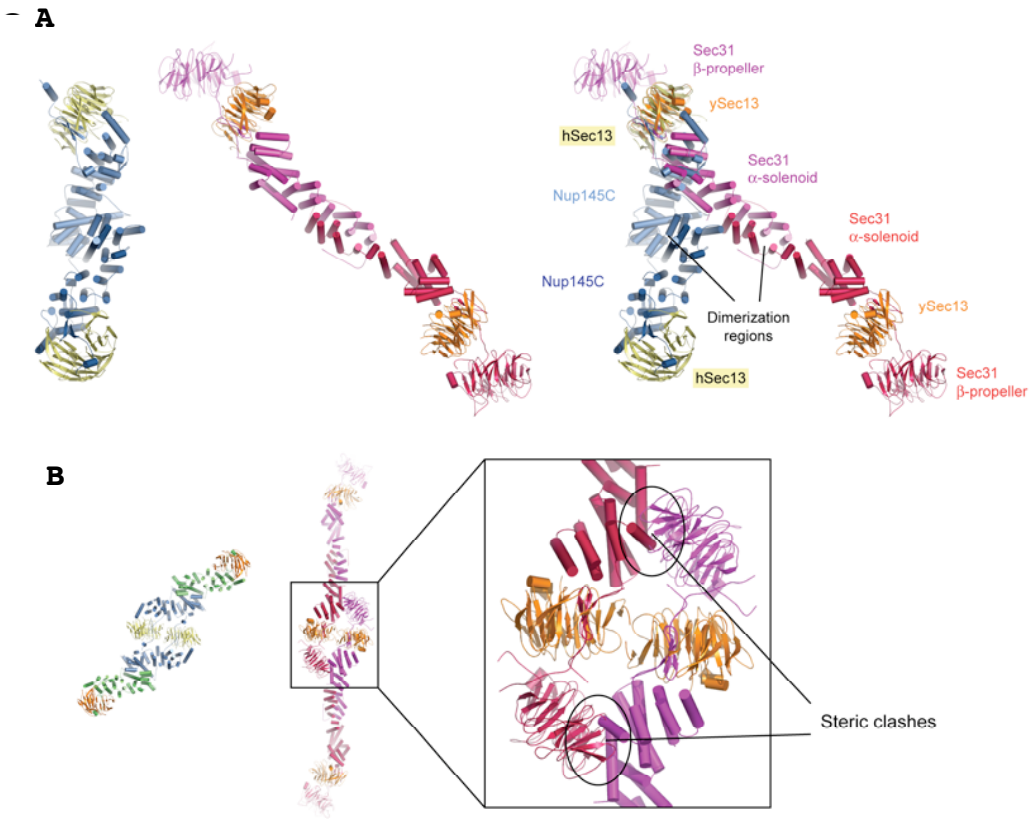


Figure 69. Structural comparison between the Nup145C/Sec13 and Sec31/Sec13 complexes

(A) Superimposition of the Nup145C/Sec13 (left) and Sec31/Sec13 (middle)(PDB code 2pm6 and 2pm9, Fath et al., 2007) hetero-tetramers based on the superimposition of the human and yeast Sec13 β-propellers, harboring the DIMs of the Nup145C and Sec31. (B) Analysis of the potential dimerization of Sec31/Sec13 hetero-tetramers into hetero-octamers via the Sec13 homo-dimerization, observed in the Nup145C/Sec13 structure. This dimerization appears to be prohibited by steric clashes between the Sec31 protomers.

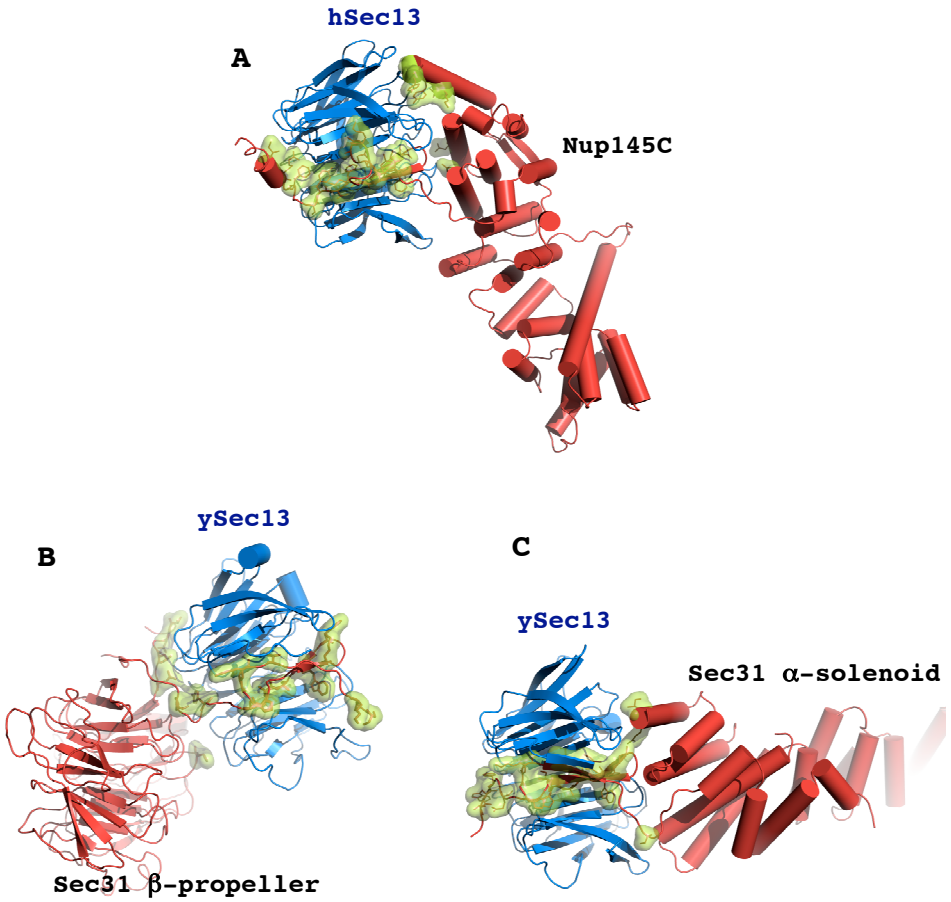


Figure 70. The Sec13 contact residues on Nup145C and Sec31

Ribbon representation of Nup145C/Sec13 (A) and Sec31/Sec13 (B)(C) complexes; Nup145C and Sec31 are shown in red and human and yeast Sec13 are shown in blue. The surface and protein side chains are shown in green and red, respectively. (A) Surface and stick rendition of the residues which interact with the human Sec13. (B) The vertex element of Sec31/Sec13 complex. Surface and stick rendition of the residues which interact with the yeast Sec13. (C) The edge element of Sec31/Sec13 complex. Surface and stick rendition of the residues which interact with the yeast Sec13.

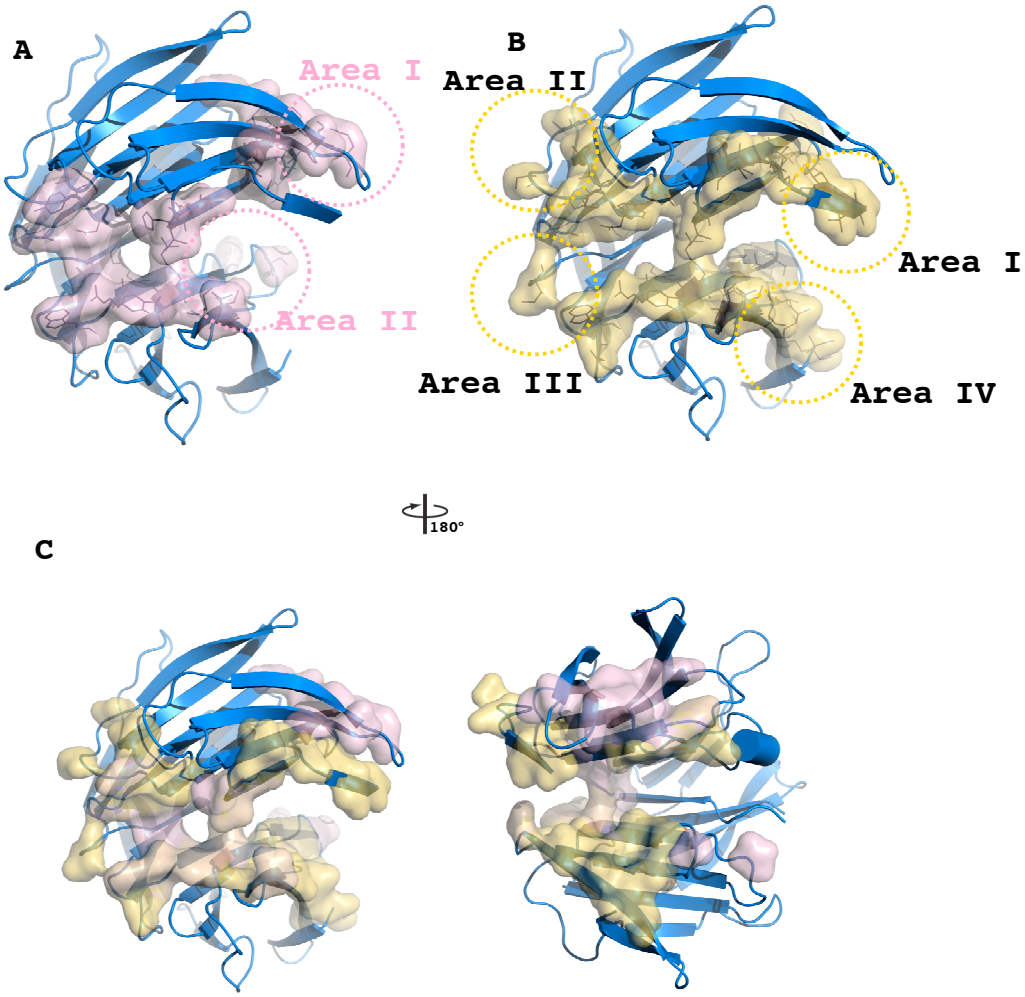


Figure 71. The areas and residues on ySec13 responsible for Nup145C and Sec31 interaction

Ribbon representation of ySec13, and surface and stick rendition of Nup145C and Sec31 interaction areas and residues. The interaction areas and residues for Nup145C binding are shown in light pink; for Sec31 binding is shown in light yellow. ySec13 is shown in blue. (A)(B) show the interaction areas and residues for Nup145C and Sec31, respectively. The areas only for Nup145C or Sec31 binding are indicated by light pink and yellow circles, respectively. (C) The areas for Nup145C and Sec31 interaction are merged. (D) A 180° rotated view is on the right.

4.9.2 Organization and function of Nup145C/Sec13 and Sec31/Sec13 complexes in the NPC and COPII

The crystallographic studies depicted that two Nup145C/Sec13 heterotetramers could assemble into a hetero-octamer primarily via the homodimerization of Sec13. In the case of COPII, a Sec31/Sec13 heterotetramer was observed in the crystal structure (Fath et al., 2007) and the higher-order oligomerization, namely hetero-octamer formation, was deduced from fitting the Sec31/Sec13 crystal structure into the cryoelectron-microscopy density maps.

In order to accommodate small cargoes, such as membrane receptors, as well as huge macromolecular cargoes, such as procollagen and chylomicron, the COPII cage has to assemble into various sizes. In addition to the cuboctahedron assembly characterized by cryo-EM and single-particle analysis (Stagg et al., 2006), a different geometry of the COPII structure, an icosidodecahedron, was recently generated by EM observation and single-particle reconstruction (Fig 72). Furthermore, according to these COPII structures, a model illustrating the molecular basis for COPII assembly and expansion was proposed (Stagg et al., 2008).

In comparing the cuboctahedron with the icosidodecahedron, two of the four angles (β angle) in the

vertex element, which is formed from the four Sec31 β-propeller domains, are variable, resulting in the formation of different geometries of the COPII cage (Fig 73). Interestingly, since Sec13 is involved in determining another two angles (α-angle), which are fixed as 60° in both COPII structures, Sec13 was suggested to function as an anchor for COPII assembly.

In analogy to the COPII cage, the NPC also has to accommodate the transport of cargo of different size and shape. As mentioned previously, the sliding capability of channel nups allows the NPC to adjust the diameter of the pore so that large cargo, such as pre-ribosomal units, can pass through the NPC properly. However, the import of integral membrane proteins has been demonstrated to proceed along another path adjacent to the pore membrane instead of the typical import pathway of the central channel (King et al., 2006). Hence, the plasticity of the membrane coat made by the Nup84 sub-complex is suggested to facilitate the import of integral membrane proteins.

The Nup145C/Sec13 hetero-octamer appears fairly rigid so that the shape of the Nup145C/Sec13 pole may not arbitrarily change, suggesting a function for building and maintaining the shape of the membrane coat. In contrast, the Seh1 in the Nup85/Seh1 hetero-octamer may additionally

function as a hinge that may impart flexibility and plasticity to the entire membrane coat.

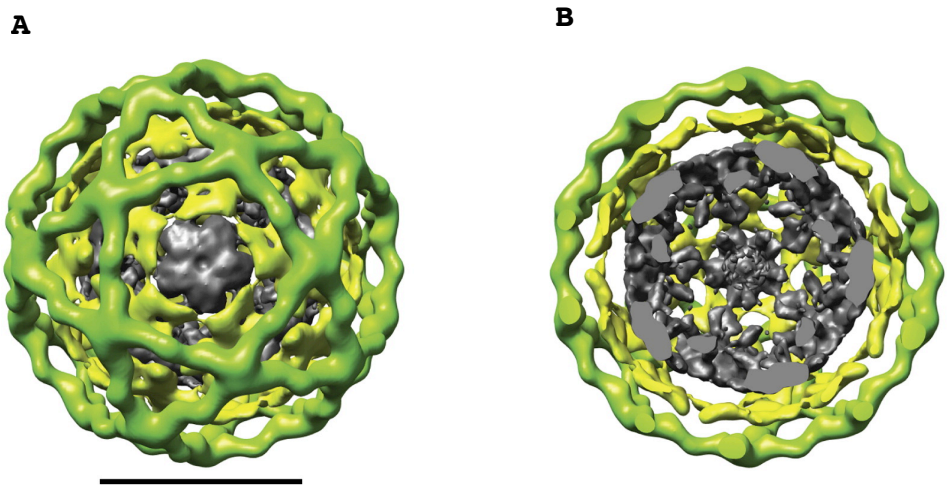


Figure 72. Structure of the COPII icosidodecahedron

(A) Single-particle reconstruction of the icosidodecahedral cage colored by radius from gray to green. Three layers of EM density are observed corresponding to Sec13-31 (green), Sec23-24 (yellow), and nonspecifically bound proteins (gray). The scale bar corresponds to 500 Å. (B) Slice through the middle of the reconstruction highlighting the three unique layers of density. Image taken from Stagg et al., 2008.

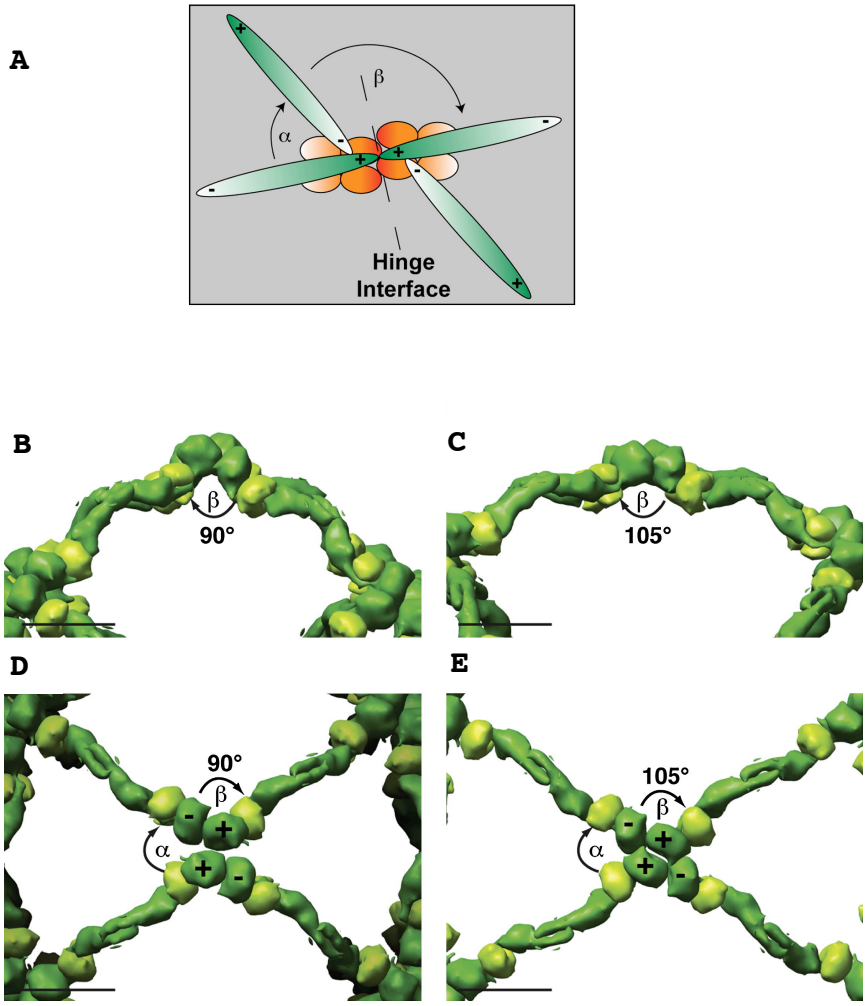


Figure 73. The angles in the vertex element influencing the assembly of the COPII Cage

(A) A diagram of the different Sec13-31 edges linked at the vertex to form the fixed α and variable β angles. A possible hinge interface for the tetramer cluster is indicated by a dashed line. The ends of the Sec13-31 edges that are closer to the vertex are labeled with a plus (+) and those farther from the vertex are labeled with a minus (-). (B) (D) The side and top views of the cuboctahedron vertex. (C) (E) Side and top views of the icosidodecahedron vertex. Image taken from Stagg et al., 2008.

4.10 The assembly of the membrane coat structure

An intriguing question is whether the membrane coat could be assembled using recombinant Nup84 sub-complex, similar to the assembly of the COPII cage.

Due to the eight-fold symmetry across the center of the pore, eight identical spokes can be observed to arrange radically around a central channel. In addition to the concentric cylinder model, the eight spokes could behave like the slices of an orange so that each spoke can be removed from or added to the NPC due to the relatively low affinity between the two spokes. Notably, seven- or nine-fold symmetry in the NPC also could be observed in isolated *Triturus* envelopes, besides the prevalent eight-fold symmetry of the NPC (Gall, 1967) (Fig 74). Consequently, individual Nup84 sub-complexes from one spoke could be pulled out from yeast cell extract instead of a ring or a coat because of the weak interactions between the Nup84 sub-complexes. (Siniossoglou et al., 2000).

Another possibility is that NPC assembly is established by numerous weak interactions among different sub-complexes. Hence, in order to form a ring or coat in vivo, multiple interactions from distinct sub-complexes are probably required or, alternatively, relatively high protein concentrations are necessary to enforce the

assembly. A single sub-complex, which is reconstituted in vitro and purified in vivo, probably is not capable of holding together to form a supra-structure.

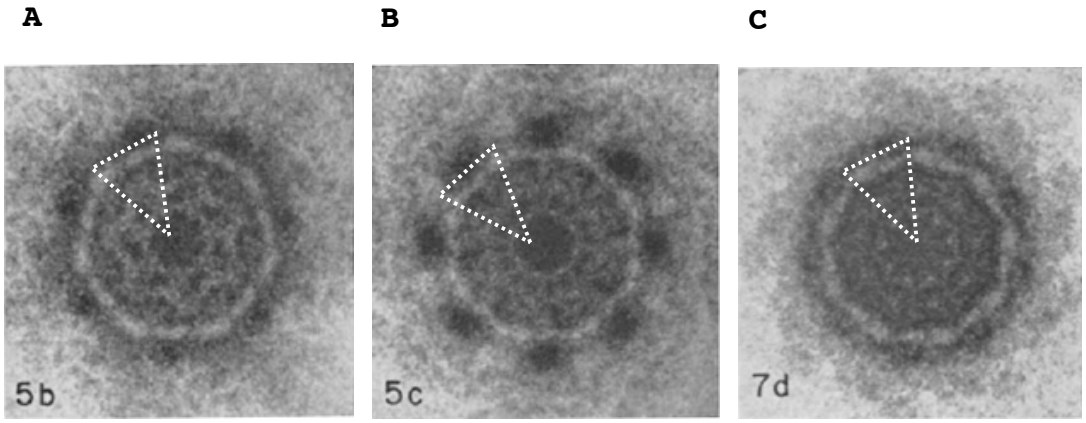


Figure 74. Different symmetries in the NPC
(A)(B)(C) Nuclear pore complexes containing 7-fold, 8-fold, and 9-fold symmetry, respectively, were observed by TEM in isolated *Triturus* envelopes. The triangle with dotted lines indicate one of the spokes in the NPC. Image taken from Gall, 1967.

4.11 Future directions

The NPC contains a significant amount of un-structured protein regions, for example FG repeats, while the majority of the remaining protein regions are expected to adopt well-defined folds, such as β-propeller, α-solenoid and coiled-coil. However, as long as atomic structures of nucleoporins are emerging, more and more unanticipated and unique features have been observed in the structures, for instance, a β-propeller invaded by an extra blade or a U-shaped α-solenoid fold. Moreover, the structural view will help us greatly to understand the function and organization of the whole NPC. Hence, solving nucleoporin structures, both on a single protein and a protein complex level, is inevitable if one wants to decipher the detailed molecular basis of nucleocytoplasmic transport and NPC-coupled processes, such as DNA repair and mRNA export.

Ideally, once nucleoporin structures are solved at the atomic level, these structures can be docked into the cryo-EM structures, like the cases of clathrin and COPII vesicles. Unfortunately, the resolution gap still precludes such fitting of atomic structures into current cryo-ET structures. Therefore, alternatively, the determination of large sub-complex structures at the atomic level appears to be the most promising route at present. In the case of the

Nup84 sub-complex, the entire sub-complex may be determined by either x-ray crystallography or small-angle X-ray scattering (SAXS), in which atomic structures of individual components can eventually be docked into the SAXS envelope.

However, the assembly of higher-order structures from distinct sub-complexes may not be as straightforward as the determination of individual sub-complexes. As mentioned previously, the interaction between each sub-complex probably may be so weak that extremely high protein concentrations or additional assembly components such as lipids or other nups may be required in order to build the supra-molecular NPC structure.

REFERENCES

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Sali, A., Rout, M.P. (2007). The molecular architecture of the nuclear pore complex. Nature *450*, 695-701.

Allen, N.P., Huang, L., Burlingame, A., Rexach, M. (2001). Proteomic analysis of nucleoporin interacting proteins. J. Biol. Chem. *276*, 29268-29274.

Antonin, W., Mattaj, I.W. (2005). Nuclear pore complexes: round the bend? Nat. Cell Biol. *7*, 10-12.

Baï, S.W., Rouquette, J., Umeda, M., Faigle, W., Loew, D., Sazer, S., Doye, V. (2004). The fission yeast Nup107-120 complex functionally interacts with the small GTPase Ran/Spi1 and is required for mRNA export, nuclear pore distribution, and proper cell division. Mol. Cell. Biol. *24*, 6379-6392.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems:

application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA *98*, 10037–10041.

Bapteste, E., Charlebois, R.L., MacLeod, D., Brochier, C. (2005). The two tempos of nuclear pore complex evolution: highly adapting proteins in an ancient frozen structure. Genome Biol. *6*, R85.

Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. *6*, 37–40.

Bayliss, R., Littlewood, T., Stewart, M. (2000). Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell *102*, 99-108.

Bayliss, R., Littlewood, T., Strawn, L.A., Wente, S.R., Stewart, M. (2002). GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. J. Biol. Chem. *277*, 50597-50606.

Beck, M., Förster, F., Ecke, M., Plitzko, J.M., Melchior, F., Gerisch, G., Baumeister, W., Medalia, O. (2004). Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. Science *306*, 1387-1390.

Beisel, H.G., Kawabata, S., Iwanaga, S., Huber, R., Bode, W. (1999). Tachylectin-2: crystal structure of a specific GlcNAc/GalNAc-binding lectin involved in the innate immunity host defense of the Japanese horseshoe crab Tachypleus tridentatus. EMBO J. *18*, 2313-2322.

Berke, I.C., Boehmer, T., Blobel, G., Schwartz, T.U. (2004). Structural and functional analysis of Nup133 domains reveals modular building blocks of the nuclear pore complex. J. Cell Biol. *167*,591-597.

Bi, X., Mancias, J.D., Goldberg, J. (2007). Insights into COPII coat nucleation from the structure of Sec23.Sar1 complexed with the active fragment of Sec31. Dev. Cell *13*, 635-645.

Blobel, G. (1980). Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA *77*, 1496-1500.

Blobel, G. (1985). Gene gating: a hypothesis. Proc. Natl. Acad. Sci. USA *82*, 8527-8529.

Boehmer, T., Enninga, J., Dales, S., Blobel, G., Zhong, H. (2003). Depletion of a single nucleoporin, Nup107, prevents the assembly of a subset of nucleoporins into the nuclear pore complex. Proc. Natl. Acad. Sci. USA *100*, 981-985.

Brohawn, S.G., Leksa, N.C., Spear, E.D., Rajashankar, K.R., Schwartz, T.U. (2008). Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. Science *322*, 1369-1373.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905—921.

CCP4 (1994). Collaborative Computational Project, Number 4, The CCP4 Suite: Programs for Protein Crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760—763.

Chial, H.J., Rout, M.P., Giddings, T.H., Winey, M. (1998). Saccharomyces cerevisiae Ndc1p is a shared component of

nuclear pore complexes and spindle pole bodies. J. Cell Biol. *143*, 1789-1800.

Cingolani, G., Petosa, C., Weis, K., Müller, C.W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. Nature *399*, 221-229.

Chook, Y.M., and Blobel, G. (1999). Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. Nature *399*, 230-237.

Chook, Y.M., and Blobel, G. (2001). Karyopherins and nuclear import. Curr. Opin. Struct. Biol. *11*, 703-715.

Cook, A., Bono, F., Jinek, M., Conti, E. (2007). Structural biology of nucleocytoplasmic transport. Annu. Rev. Biochem. *76*, 647-671.

Cordes, V.C., Reidenbach, S., 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., Matunis, M.J. (2002). Proteomic analysis of the mammalian nuclear pore complex. J. Cell Biol. *158*, 915-927.

Debler, E.W., Ma, Y., Seo, H.S., Hsia, K.C., Noriega, T.R., Blobel, G., Hoelz, A. (2008) A fence-like coat for the nuclear pore membrane. Mol. Cell *32*, 815-826.

Devos, D., Dokudovskaya, S., Alber, F., Williams, R., Chait, B.T., Sali, A., Rout, M.P. (2004). Components of coated vesicles and nuclear pore complexes share a common molecular architecture. PLoS Biol. *2*, e380.

Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., 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., Sali, A. (2006). Simple fold composition and modular architecture of the nuclear pore complex. Proc. Natl. Acad. Sci. USA *103*, 2172-2177.

Denning, D.P., and Rexach, M.F. (2007). Rapid evolution exposes the boundaries of domain structure and function in natively unfolded FG nucleoporins. Mol. Cell. Proteomics *6*, 272-282.

Dockendorff, T.C., Heath, C.V., Goldstein, A.L., Snay, C.A., Cole, C.N. (1997). C-terminal truncations of the yeast nucleoporin Nup145p produce a rapid temperature-conditional mRNA export defect and alterations to nuclear structure. Mol. Cell. Biol. *17*, 906-920.

Drin, G., Casella, J.F., Gautier, R., Boehmer, T., Schwartz, T.U., Antonny, B. (2007). A general amphipathic alpha-helical motif for sensing membrane curvature. Nat. Struct. Mol. Biol. *14*, 138-146.

Faber, H.R., Groom, C.R., Baker, H.M., Morgan, W.T., Smith, A., Baker, E.N. (1995). 1.8 A crystal structure of the C-terminal domain of rabbit serum haemopexin. Structure *3*, 551-559.

Fahrenkrog, B., Köser, J., Aebi, U. (2004). The nuclear pore complex: a jack of all trades? Trends Biochem. Sci. *29*, 175-182.

Fath, S., Mancias, J.D., Bi, X., Goldberg, J. (2007). Structure and organization of coat proteins in the COPII cage. Cell *129*, 1325-1336.

Fontes, M.R.M., Teh, T., Toth, G., John, A., Pavo, I., Jans, D.A., Kobe, B. (2003). Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-alpha. Biochem. J. *375*, 339-349.

Fontoura, B.M., Blobel, G., Matunis, M.J. (1999). A conserved biogenesis pathway for nucleoporins: proteolytic processing of a 186-kilodalton precursor generates Nup98 and the novel nucleoporin, Nup96. J. Cell Biol. *144*, 1097-1112.

Fotin, A., Cheng, Y., Sliz, P., Grigorieff, N., Harrison, S.C., Kirchhausen, T., Walz, T. (2004). Molecular model for a complete clathrin lattice from electron cryomicroscopy. Nature *432*, 573-579.

Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Mini, T., Jenö, P., Engel, J., Kammerer, R.A. (2002).

Characterization of the matrilin coiled-coil domains reveals seven novel isoforms. J. Biol. Chem. *277*, 19071-19079.

Frey, S., Richter, R.P., Görlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science *314*, 815-817.

Gall, J.G. (1967). Octagonal nuclear pores. J. Cell Biol. *32*, 391-399.

Galy, V., Olivo-Marin, J.C., Scherthan, H., Doye, V., Rascalou, N., Nehrbass, U. (2000). Nuclear pore complexes in the organization of silent telomeric chromatin. Nature *403*, 108-112.

Garcia-Higuera, I., Fenoglio, J., Li, Y., Lewis, C., Panchenko, M.P., Reiner, O., Smith, T.F., Neer, E.J. (1996). Folding of proteins with WD-repeats: comparison of six members of the WD-repeat superfamily to the G protein beta subunit. Biochemistry *35*, 15215-15221.

Goldstein, A.L., Snay, C.A., Heath, C.V., Cole, C.N. (1996). Pleiotropic nuclear defects associated with a conditional allele of the novel nucleoporin Rat9p/Nup85p. Mol. Biol. Cell *7*, 917-934.

Gotta, M., Laroche, T., Formenton, A., Maillet, L., Scherthan, H., Gasser, S.M. (1996). The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type Saccharomyces cerevisiae. J. Cell Biol. *134*, 1349-1363.

Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A., Barford, D. (1999) The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. Cell *96*, 99-110.

Hinshaw, J.E., Carragher, B.O., Milligan, R.A. (1992). Architecture and design of the nuclear pore complex. Cell *69*, 1133-1141.

Hoelz, A., and Blobel, G. (2004). Cell biology: popping out of the nucleus. Nature *432*, 815-816.

Hoelz, A., Nairn, A.C., Kuriyan, J. (2003). Crystal structure of a tetradecameric assembly of the association domain of Ca2+/calmodulin-dependent kinase II. Mol. Cell *11*, 1241-1251.

Hsia, K.C., Stavropoulos, P., Blobel, G., Hoelz, A. (2007). Architecture of a coat for the nuclear pore membrane. Cell *131*, 1313-1326.

Iovine, M.K., and Wente, S.R. (1997). A nuclear export signal in Kap95p is required for both recycling the import factor and interaction with the nucleoporin GLFG repeat regions of Nup116p and Nup100p. J. Cell Biol. *137*, 797-811.

Jawad, Z., and Paoli, M. (2002). Novel sequences propel familiar folds. Structure *10*, 447-454.

Jeanmougin, F., Thompson, J.D., Guoy, M., Higgins, D.G., and Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. Trends Biochem. Sci. *23*, 403—405.

Jeon, H., Meng, W., Takagi, J., Eck, M.J., Springer, T.A., Blacklow, S.C. (2001). Implications for familial hypercholesterolemia from the structure of the LDL receptor

YWTD-EGF domain pair. Nat. Struct. Biol. *8*, 499-504

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crysallogr. A Found. Crystallogr. *47*, 110—119.

Kaiser, C.A., Schekman, R. (1990). Distinct sets of SEC genes govern transport vesicle formation and fusion early in the secretory pathway. Cell *61*, 723-733.

Keminer, O. and Peters, R. (1999). Permeability of single nuclear pores. Biophys. J. *77*, 217-228.

King, M.C., Lusk, C.P., Blobel, G. (2006). Karyopherin-mediated import of integral inner nuclear membrane proteins. Nature *442*, 1003-1007.

Kirchhausen, T. (2000). Three ways to make a vesicle. Nat. Rev. Mol. Cell Biol. *1*, 187-198.

Klenchin, V.A., Khaitlina, S.Y., Rayment, I. (2006). Crystal structure of polymerization-competent actin. J. Mol. Biol. *362*, 140-150.

Kiseleva, E., Allen, T.D., Rutherford, S.A., Murray, S., Morozova, K., Gardiner, F., Goldberg, M.W., Drummond, S.P. (2007). A protocol for isolation and visualization of yeast nuclei by scanning electron microscopy (SEM). Nat. Protoc. *2*, 1943-1953.

Kobe, B., and Kajava, A.V. (2000). When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends Biochem. Sci. *25*, 509-515.

Krull, S., Thyberg, J., Björkroth, B., Rackwitz, H.R., Cordes, V.C. (2004). Nucleoporins as components of the nuclear pore complex core structure and Tpr as the architectural element of the nuclear basket. Mol. Biol. Cell *15*, 4261-4277.

Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E., Sigler, P.B. (1996). The 2.0 A crystal structure of a heterotrimeric G protein. Nature *379*, 311-319

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl.

Cryst. *26*, 283–291.

Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J. Mol. Biol. *234*, 946-950.

Lee, S.J., Matsuura, Y., Liu, S.M., Stewart, M. (2005). Structural basis for nuclear import complex dissociation by RanGTP. Nature *435*, 693-696.

Lim, R.Y., Huang, N.P., Köser, J., Deng, J., Lau, K.H., Schwarz-Herion, K., Fahrenkrog, B., Aebi, U. (2006). Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. Proc. Natl. Acad. Sci. USA *103*, 9512-9517.

Loïodice, I., Alves, A., Rabut, G., Van Overbeek, M., Ellenberg, J., Sibarita, J.B., Doye, V. (2004). The entire Nup107-160 complex, including three new members, is targeted as one entity to kinetochores in mitosis. Mol. Biol. Cell *15*, 3333-3344.

Lupas, A. (1997). Predicting coiled-coil regions in proteins. Curr. Opin. Struct. Biol. *7*, 388-393.

Lutzmann, M., Kunze, R., Buerer, A., Aebi, U., Hurt, E. (2002). Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins. EMBO J. *2*, 387-397.

Macara I.G. (2001). Transport into and out of the nucleus. Microbiol. Mol Biol. Rev. *65*, 570-594.

Matsuoka, K., Schekman, R., Orci, L., Heuser, J.E. (2001). Surface structure of the COPII-coated vesicle. Proc. Natl. Acad. Sci. USA *98*, 13705-13709.

Maul, G.G., Maul, H.M., Scogna, J.E., Lieberman, M.W., Stein, G.S., Hsu, B.Y., Borun, T.W. (1972). Time sequence of nuclear pore formation in phytohemagglutinin-stimulated lymphocytes and in HeLa cells during the cell cycle. J. Cell Biol. *55*, 433-447.

Melcák, I., Hoelz, A., Blobel, G. (2007). Structure of Nup58/45 suggests flexible nuclear pore diameter by intermolecular sliding. Science *315*, 1729-1732.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods

Enzymol. *276*, 307—326.

Panté, N., Kann, M. (2002). Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. Mol. Biol. Cell *13*, 425-434.

Paoli, M., Anderson, B.F., Baker, H.M., Morgan, W.T., Smith, A., Baker, E.N. (1999). Crystal structure of hemopexin reveals a novel high-affinity heme site formed between two beta-propeller domains. Nat. Struct. Biol. *6*, 926-931.

Patel, S.S., Belmont, B.J., Sante, J.M., Rexach, M.F. (2007). Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. Cell *129*, 83-96.

Richards, S.A., Lounsbury, K.M., Macara, I.G. (1995). The C terminus of the nuclear RAN/TC4 GTPase stabilizes the GDP-bound state and mediates interactions with RCC1, RAN-GAP, and HTF9A/RANBP1. J. Biol. Chem. *270*, 14405-14411.

Ron, D., Chen, C.H., Caldwell, J., Jamieson, L., Orr, E., Mochly-Rosen, D. (1994). Cloning of an intracellular

receptor for protein kinase C: a homolog of the beta subunit of G proteins. Proc. Natl. Acad. Sci. USA *91*, 839-843.

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., 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.

Rout, M.P., and Wente, S.R. (1994). Pores for thought: nuclear pore complex proteins. Trends Cell Biol. *4*, 357-365.

Salama, N.R., Yeung, T., Schekman, R.W. (1993). The Sec13p complex and reconstitution of vesicle budding from the ER with purified cytosolic proteins. EMBO J. *12*, 4073-4082.

Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers *38*, 305—320.

Saxena, K., Gaitatzes, C., Walsh, M.T., Eck, M., Neer, E.J., Smith, T.F. (1996). Analysis of the physical properties and molecular modeling of Sec13: A WD repeat protein involved in vesicular traffic. Biochemistry. *35*, 15215-15221.

Schwartz, T.U. (2005). Modularity within the architecture of the nuclear pore complex. Curr. Opin. Struct. Biol. *15*, 221-226.

Shulga, N., and Goldfarb, D.S. (2003). Binding dynamics of structural nucleoporins govern nuclear pore complex permeability and may mediate channel gating. Mol. Cell. Biol. *23*, 534-542.

Siniossoglou, S., Lutzmann, M., Santos-Rosa, H., Leonard, K., Mueller, S., Aebi, U., Hurt, E. (2000). Structure and assembly of the Nup84p complex. J. Cell Biol. *149*, 41-54.

Siniossoglou, S., Wimmer, C., Rieger, M., Doye, V., Tekotte, H., Weise, C., Emig, S., Segref, A., Hurt, E.C. (1996). A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. Cell *84*, 265-275.

Stagg, S.M., Gürkan, C., Fowler, DM., LaPointe, P., Foss, T.R., Potter, C.S., Carragher, B., Balch, W.E. (2006). Structure of the Sec13/31 COPII coat cage. Nature *439*, 234-238.

Stagg, S.M., LaPointe, P., Razvi, A., Gürkan, C., Potter, C.S., Carragher, B., Balch, W.E. (2008). Structural basis for cargo regulation of COPII coat assembly. Cell *134*, 474-484.

Stewart, M., Kent, H.M., McCoy, A.J. (1998). The structure of the Q69L mutant of GDP-Ran shows a major conformational change in the switch II loop that accounts for its failure to bind nuclear transport factor 2 (NTF2). J. Mol. Biol. *284*, 1517-1527.

Strawn, L.A., Shen, T., Shulga, N., Goldfarb, D.S., Wente, S.R. (2004). Minimal nuclear pore complexes define FG repeat domains essential for transport. Nat. Cell Biol. *6*, 197-206.

Suntharalingam, M., and Wente, S.R. (2003) Peering through the pore: nuclear pore complex structure, assembly, and function. Dev. Cell. *4*, 775-789.

Teixeira, M.T., Fabre, E., Dujon, B. (1999). Self-catalyzed cleavage of the yeast nucleoporin Nup145p precursor. J. Biol. Chem. *274*, 32439-32444.

Teixeira, M.T., Siniossoglou, S., Podtelejnikov, S., Bénichou, J.C., Mann, M., Dujon, B., Hurt, E., Fabre, E. (1997). Two functionally distinct domains generated by in vivo cleavage of Nup145p: a novel biogenesis pathway for nucleoporins. EMBO J. *16*, 5086-5097.

Vasu, S.K., and Forbes, D.J. (2001). Nuclear pores and nuclear assembly. Curr. Opin. Cell Biol. *13*, 363-375.

Vetter, I.R., Arndt, A., Kutay, U., Gorlich, D., Wittinghofer, A. (1999). Structural view of the Ran-Importin beta interaction at 2.3 A resolution. Cell *97*, 635-646.

Vetter, I.R., Nowak, C., Nishimoto, T., Kuhlmann, J., Wittinghofer, A. (1999). Structure of a Ran-binding domain

complexed with Ran bound to a GTP analogue: implications for nuclear transport. Nature *398*, 39-46.

Vinson, C.R., Sigler, P.B., McKnight, S.L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science *246*, 911-916.

Watson, M.L. (1954). Pores in the mammalian nuclear membrane. Biochim. Biophys. Acta. *15*, 475-479.

Weis, K. (2007). The nuclear pore complex: oily spaghetti or gummy bear? Cell *130*, 405-407.

Winey, M., Yarar, D., Giddings, T.H. Jr., Mastronarde, D.N. (1997). Nuclear pore complex number and distribution throughout the Saccharomyces cerevisiae cell cycle by three-dimensional reconstruction from electron micrographs of nuclear envelopes. Mol. Biol. Cell *8*, 2119-2132.

Wolf, E., Kim, P.S., Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. Protein Sci. *6*, 1179-1189.

Wozniak, R.W., Blobel, G., Rout, M.P. (1994). POM152 is an integral protein of the pore membrane domain of the yeast nuclear envelope. J. Cell Biol. *125*, 31-42.

Xia, Z., Dai, W., Zhang, Y., White, S.A., Boyd, G.D., Mathews, F.S. (1996). Determination of the gene sequence and the three-dimensional structure at 2.4 angstroms resolution of methanol dehydrogenase from Methylophilus W3A1. J. Mol. Biol. *259*, 480-501.

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

Zhao, J., Jin, S.B., Björkroth, B., Wieslander, L., Daneholt, B. (2002). The mRNA export factor Dbp5 is associated with Balbiani ring mRNP from gene to cytoplasm. EMBO J. *21*, 1177-1187.