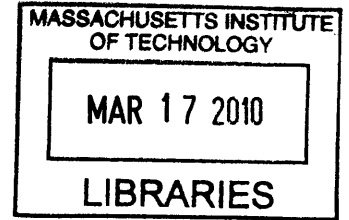# Structural elucidation of a common architecture of the nuclear pore complex and COPII vesicle coats

By

Stephen Graf Brohawn

B.S. Biochemistry
University of Delaware, 2004

Submitted to the Department of Biology in partial fulfillment of the requirements
for the degree of

ARCHIVES

Doctor of Philosophy
at the
Massachusetts Institute of Technology

June 2010

Signature of Author:_____

Department of Biology
February 10, 2010

Certified by:__

Thomas U. Schwartz
Associate Professor of Biology
Thesis Supervisor

Accepted by:___

Stephen P. Bell
Professor of Biology
Co-Chair, Biology Graduate Committee

Structural elucidation of a common architecture of the nuclear pore complex and COPII vesicle coats

By

Stephen Graf Brohawn

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology

Abstract

Nuclear pore complexes (NPCs) are massive protein assemblies that perforate the nuclear envelope and form the exclusive passageway for nucleocytoplasmic transport. NPCs play critical roles in molecular transport and a myriad of other cellular processes. Elucidation of the structure of the NPC is thus expected to provide important insight into cell biology. In this thesis, I investigate the structure of a key subcomplex of the NPC and discuss the evolutionary relationship between the NPC and COPII vesicle coats it illustrates.

The NPC is a modular assembly, with a stable structural scaffold supporting dynamically attached components. The structural scaffold is constructed from multiple copies of the Y-shaped complex and the Nic96 complex. We solved the crystal structure of the heterodimeric Nup85-Seh1 module that forms a short arm in the Y complex. Nup85 is found to contain a conserved fold, the ancestral coatomer element 1 (ACE1), also present in three other components of the NPC and in the COPII vesicle coat, providing structural evidence of coevolution from a common ancestor.

Sec31 ACE1 units interact to form edge elements in the COPII lattice. Using structural knowledge of this edge element, we identified corresponding interactions between ACE1 proteins Nup84 and Nup145C in the NPC. We solved the crystal structure of the heterotrimeric Nup84-Nup145C-Sec13 module that forms the top of the long arm in the Y complex. The heterotypic ACE1 interaction of Nup145C and Nup84 is analogous to the homotypic Sec31 edge element interaction in the COPII coat.

From these and other structures, we assemble a near complete structural model of the Y complex. Further, based on the demonstrated relationship with the COPII coatomer lattice, we propose a lattice model for the entire NPC scaffold. The common architectural principles of the edge elements in the NPC and COPII lead us to predict that Y complexes will be arranged as struts in the NPC lattice. In this manner, Nup84-Nup145C edge elements are arranged parallel to the transport axis to stabilize the positively curved nuclear envelope. From a lattice model of the NPC follow hypotheses for how other components are integrated into and function within the NPC.

Thesis Supervisor: Thomas U. Schwartz
Title: Pfizer-Laubach Career Development Associate Professor of Biology

Acknowledgements

I am fortunate enough to have had an exceptionally rewarding and fun time in graduate school and for that I have a lot of people to thank.

First, I would like to thank my advisor Thomas Schwartz for giving me the opportunity to work and learn in his lab. I feel so lucky to have been able to experience his unwavering support, infectious curiosity and excitement, and real passion for science in the lab for the past five years. He created and fosters a fantastic environment for work, science, and everything else, and for that I am very grateful. I would also like to thank Bob Sauer and Mike Yaffe for being on my thesis committee. Their insights and advice offered at every meeting have been a great source of direction and have contributed substantially to my work. I would especially like to thank them for being so accessible throughout the entire process. I also thank Cathy Drennan and Steve Harrison for joining my committee. I am greatly appreciative to these scientists that I respect so much for helping me learn to be one.

Thanks go to everyone who is in and has been in the Schwartz lab. I would especially like to thank the other graduate students in the lab; James Partridge and James Whittle, who made joining a new lab seem a lot less crazy and a lot more fun, and Nina Leksa, who has been just fantastic to work with on a lot of the projects in this thesis since she joined. Thanks to the first post docs, Thomas Boehmer and then Sandra Jeudy, for showing me a lot and having a lot of fun as well. Thanks to the rest of the current group - Silvija Bilokapic, Brian Sosa, Raven Ready, and Alexander Ulrich - and to all the other people who have left for making the lab a great place to be. Thanks also to Eric Spear for tons of help. Special thanks go to Lena Karotki, Sabin Mulepati, Glenna Wink, and Marko Gogola for working with me and in the end teaching me more than I taught them!

Lastly, thanks go to my family and friends outside the lab. Thanks to Mom and Dad for giving me the opportunity to be here and for always encouraging me to get involved in the things I was interested in. Thanks also to the rest of my family and friends, especially my brothers Dan, Dave, and Mike, for perspective and support. And finally thanks to Katie, for being with me the whole way through.

Table of Contents

List of Figures

11

List of Tables

Chapter one

The nuclear pore complex has entered the atomic age

Abstract

Nuclear pore complexes (NPCs) perforate the nuclear envelope and represent the exclusive passageway into and out of the nucleus of the eukaryotic cell. Apart from their essential transport function, components of the NPC have important, direct roles in nuclear organization and in gene regulation. Due to its central role in cell biology, it is of considerable interest to determine the NPC structure at atomic resolution. The complexity of these large, 40-60 MDa protein assemblies has for decades limited such structural studies. More recently, exploiting the intrinsic modularity of the NPC, structural biologists are making progress toward understanding this nanomachine in molecular detail. Structures of building blocks of the stable, architectural scaffold of the NPC have been solved, and distinct models for their assembly proposed. Here we review the status of the field and lay out the challenges and the next steps toward a full understanding of the NPC at atomic resolution.

Introduction

The hallmark of eukaryotic cells is an elaborate endomembrane system that creates membrane-enclosed organelles. The nucleus is the most prominent organelle, as it harbors the genetic material of the cell. NPCs are the only gateways to the nucleus and reside in circular openings in the nuclear envelope, where the inner nuclear membrane (INM) and outer nuclear membrane (ONM) of the nuclear envelope (NE) are fused. NPCs are among the largest multiprotein assemblies in the quiescent cell and were first described 50 years ago with electron microscopy (Watson, 1959). Here we provide a snapshot on the status of the structural characterization of the nuclear pore complex – the results of a truly multi-disciplinary approach. For general reviews the reader is also referred to (D'Angelo and Hetzer, 2008; Lim et al., 2008; Tran and Wente, 2006), for the mechanism of NPC assembly to (Antonin et al., 2008), and for nucleocytoplasmic transport of proteins and RNAs to (Carmody and Wente, 2009; Cook et al., 2007; Kohler and Hurt, 2007; Pemberton and Paschal, 2005; Stewart, 2007; Weis,

2003). The role of the NPC in gene regulation and nuclear organization is addressed in (Akhtar and Gasser, 2007; Heessen and Fornerod, 2007).

Overall structure

Electron microscopy has been the best technique to observe the overall structure of the NPC. A variety of cell types from different organisms have been imaged. In its internal symmetry, shape, and size the NPC seems conserved throughout evolution, though at the molecular level there are differences, as noted below. In internal symmetry, the first electron micrographs of the NPC showed that it forms an octagonal ring whose central channel is less electron dense then the eight lobes that surround it. In shape, scanning electron microscopy experiments (SEM) have recorded some of the most stunning NPC images (Figure 1.1). While the architectural core is grossly symmetric about the plane of the membrane, the peripheral components on the nuclear and cytoplasmic faces are distinct. These peripheral components recapitulate the eightfold symmetry about the transport axis exhibited by the architectural core. On the cytoplasmic side, eight knobs, thought to be attachment sides for fibrilic extensions, are visible in NPCs from multicellular species (Kiseleva et al., 2000). In yeast, these features are less pronounced, but are likely also present (Kiseleva et al., 2004). On the nucleoplasmic side, a ring termed the nuclear basket is suspended from eight filaments that join it to the NPC. Lastly, in size, the overall and central channel diameters of the NPC appear to be similar in all eukaryotes, about 90-120nm and 50nm respectively (Akey and Radermacher, 1993; Beck et al., 2007; Fahrenkrog et al., 2000; Frenkiel-Krispin et al., 2009; Hinshaw et al., 1992; Stoffler et al., 2003). However, there is still considerable uncertainty about the height, determined to be between ~30-95nm (Alber et al., 2007b; Elad et al., 2009).

A

B

side view   cytoplasmic view
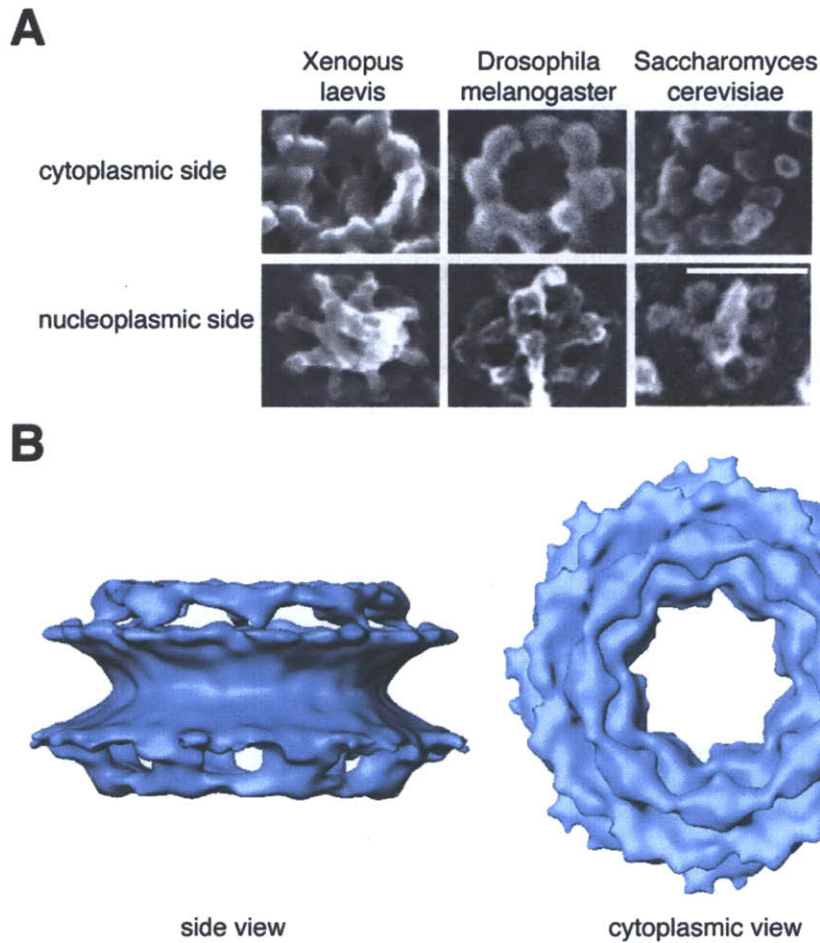
Figure 1.1 - Overall structure of the nuclear pore complex
(A) Representative micrographs of NPCs from diverse eukaryotes obtained by
scanning electron microscopy. The distinct surface features that define
cytoplasmic and nucleoplasmic faces of the NPC are conserved, so are the
overall dimensions in the plane of the nuclear envelope. Scale bar indicates
100nm.
(B) Cryo-electron tomographic (cryo-ET) reconstruction of the human NPC. The
nuclear basket structure and the cytoplasmic extensions are omitted for clarity.
The central eightfold rotational symmetry is clearly visible. A comparison
between cryo-ET reconstructions of NPCs from diverse species reveals
substantial differences in the overall height (Elad et al., 2009).

Further work has led to progressively more detailed reconstructions of the NPC.

Cryo-electron microscopy that relies on averaging images from many NPCs has

been employed to study the core NPC structure, the part that spans the distance

between the faces of INM and ONM (Akey and Radermacher, 1993; Hinshaw et

al., 1992). These studies have shown that the scaffold ring structure - the

electron dense material near the nuclear membrane - has alternating thicker and thinner regions, hence it is often called the spoke ring ((Akey and Radermacher, 1993; Hinshaw et al., 1992) Due to the confusing nature of this terminology, in this thesis, "scaffold ring" will be used and what are referred to as "spokes" in the literature will be referred to as scaffold segments). The scaffold structure appears to penetrate the pore membrane to also form a perinuclear ring structure. Using cryo-electron tomography, the best pictures of complete NPCs have been achieved, extending even to a resolution of ~ 6 nm (Beck et al., 2007; Elad et al., 2009; Frenkiel-Krispin et al., 2009). With this technique, details of the ring structures become apparent. The scaffold can be divided into three main ring elements: a central scaffold ring sandwiched between a cytoplasmic ring and a nucleoplasmic ring. The rings appear to float on top of one another or to be thinly connected, indicating that material connecting these rings is less electron-dense. Alternatively, this may be due to technical difficulties, such as the 'missing cone' problem or poor resolution in the Z-direction. Recent improvement in the resolution of a *Xenopus* NPC reconstruction has further delineated the scaffold ring into two concentric rings in the plane of the NE connected with a high density mass perpendicular to the NE (Frenkiel-Krispin et al., 2009).

The central transport cavity of the nuclear pore complex shows no distinct structural features, consistent with the perception that it is filled by an aqueous meshwork formed by natively unfolded FG-domains, which are long polypeptide sequences found in several nucleoporins that contain phenylalanine-glycine (FG) repeats but are otherwise hydrophilic. These extensions are thought to form a distinct, semi-permeable environment that prevents the diffusion of large molecules, unless they are bound to nuclear transport factors (NTRs) that facilitate entry into this central cavity. NTRs recognize export- or import-specific signal sequences on cargo molecules and interact directly with FG-nups to facilitate transport. Directional specificity of transport is accomplished by the small GTPase Ran-mediated regulation of cargo-NTR interactions in a nucleotide-bound state-dependant fashion. A gradient of Ran nucleotide binding

states is maintained in turn by the cytoplasmic localization of the Ran GTPase-activating protein (RanGAP) and nuclear localization of the Ran guanine-nucleotide exchange factor (RCC1).

In addition to the central channel, the scaffold itself likely harbors additional peripheral channels. The scaffold ring appears porous in cryo-EM/-ET structures, with gaps of ~ 9 nm diameter close to the NE membrane (Hinshaw et al., 1992; Stoffler et al., 2003). Peripheral channels have been discussed in several studies, and were postulated to transport small proteins and ions (Kramer et al., 2007). It is unclear, however, how this type of transport could be restricted to the peripheral channels, when the central channel could allow it as well. Alternatively, it also has been suggested that the peripheral channels transport membrane proteins destined for the INM. These are inserted into the ER membrane following translation and stay membrane-anchored until they reach their final destination (the ER, ONM, and INM are all contiguous). Perhaps these membrane proteins pass the NPC into the nucleus via these peripheral channels (Powell and Burke, 1990; Zuleger et al., 2008). The nucleoplasmic domains of INM proteins are limited in size to ~ 40 kDa, small enough to fit into the cavities of the observed channels.

While the general NPC architecture is well established, the cryo-EM/-ET structures do not permit the assignment of individual proteins, since their boundaries are not visible at this resolution. For this, higher resolution methods are required.

Modularity

A characteristic of the NPC is its high degree of modularity, which manifests itself at several levels. First, the NPC is organized around a central eightfold rotational symmetry. Second, only ~30 nucleoporins, composed of a limited set of domain topologies, build the NPC. Third, nucleoporins have various dwell times at the NPC, with only a fraction being stably attached at all times. And finally, the

stably attached nucleoporins are arranged into subcomplexes, each of which assembles in multiple copies to build the entire NPC (Figure 1.2). This modularity is the basis for approaching structural determination of the assembly at atomic resolution (Schwartz, 2005).
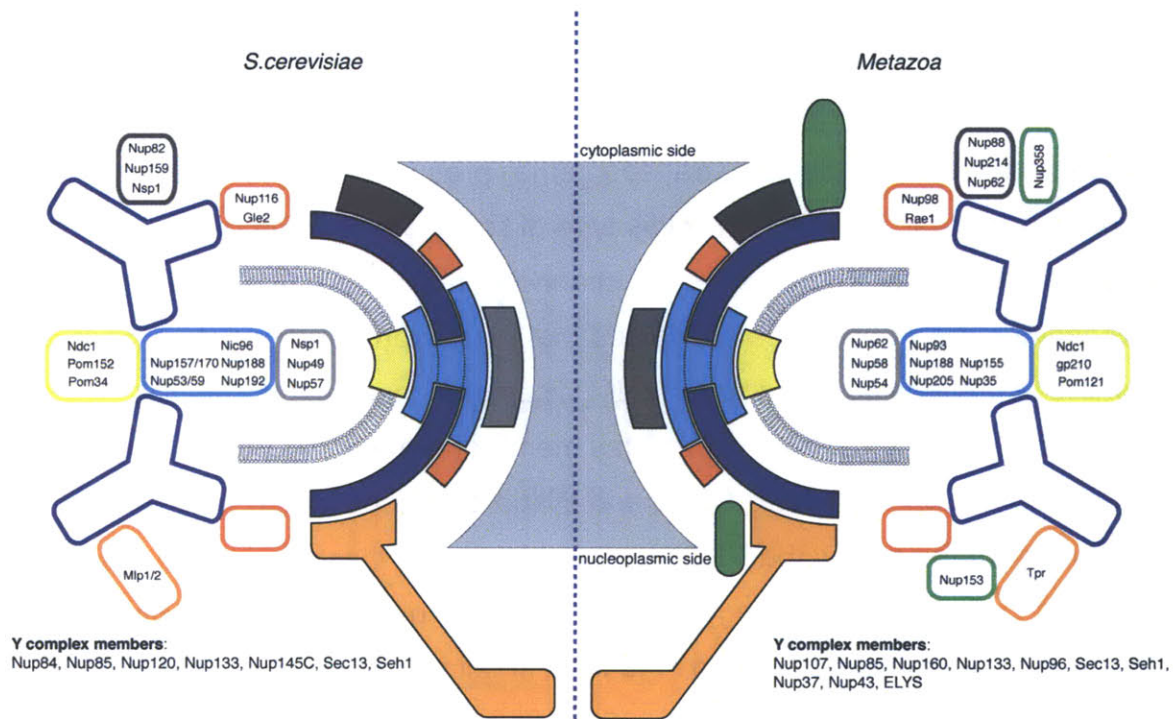
Figure 1.2 - Schematic representation of the modular assembly of the NPC
The NPC is built from ~30 nucleoporins, organized in a small set of subcomplexes. The cartoon shows the major subcomplexes that make up the lattice-like scaffold (blue colors), the membrane-attachment (yellow), and the FG-network (grey) of the NPC. *S.cerevisiae* components on the left, metazoa, with specific additional components, on the right. A few peripheral nups are left out for clarity. Simplified representation, connections are not to be taken literally and box sizes are not proportional to molecular weights.

Protein composition

Three studies, using *S. cerevisiae* (Rout et al., 2000), the trypansome *T. brucei* (DeGrasse et al., 2009), or rat hepatocytes (Cronshaw et al., 2002) as starting material, focused on determining the complete inventory of nucleoporins. In each study, cell extracts were enriched for NPCs by fractionation techniques and the purified proteins were then analyzed by mass-spectrometry. The set of proteins

found in each organism is largely identical and comprises ~ 30 different gene products. The nucleoporins can be broadly classified into three categories (Figure 1.3). ~10 contain disordered N- and/or C-terminal regions that are rich in phenylalanine-glycine (FG) repeats. These FG-repeat regions emanate into and form the transport barrier in the channel of the NPC. ~15 nucleoporins have distinct architectural functions and form the NPC scaffold structure. Three nucleoporins have transmembrane domains and anchor the NPC in the circular openings in the NE. Immunogold-labeling of all nucleoporins shows that the majority of the nups, notably scaffold nucleoporins, are symmetrically localized around a two-fold symmetry axis in the plane of the NE, perpendicular to the eightfold rotational symmetry about the main transport channel (Rout et al., 2000). Based on simple hydrodynamic and volumetric calculations the size of the NPC was estimated to range from 66 MDa in *S. cerevisiae* (Rout and Blobel, 1993) to 125 MDa in vertebrates (Reichelt et al., 1990). Calculations based on the stoichiometry of nucleoporins obtained in the proteomic studies, however, indicate that the NPC size is only 44 MDa in *S. cerevisiae* and ~ 60 MDa in rat. The discrepancy supports the conclusion that the NPC is a porous, lattice-like assembly, rather than a solid entity (Brohawn et al., 2008; Hinshaw et al., 1992), which accounts for the overestimate of mass based on volumetric analysis.
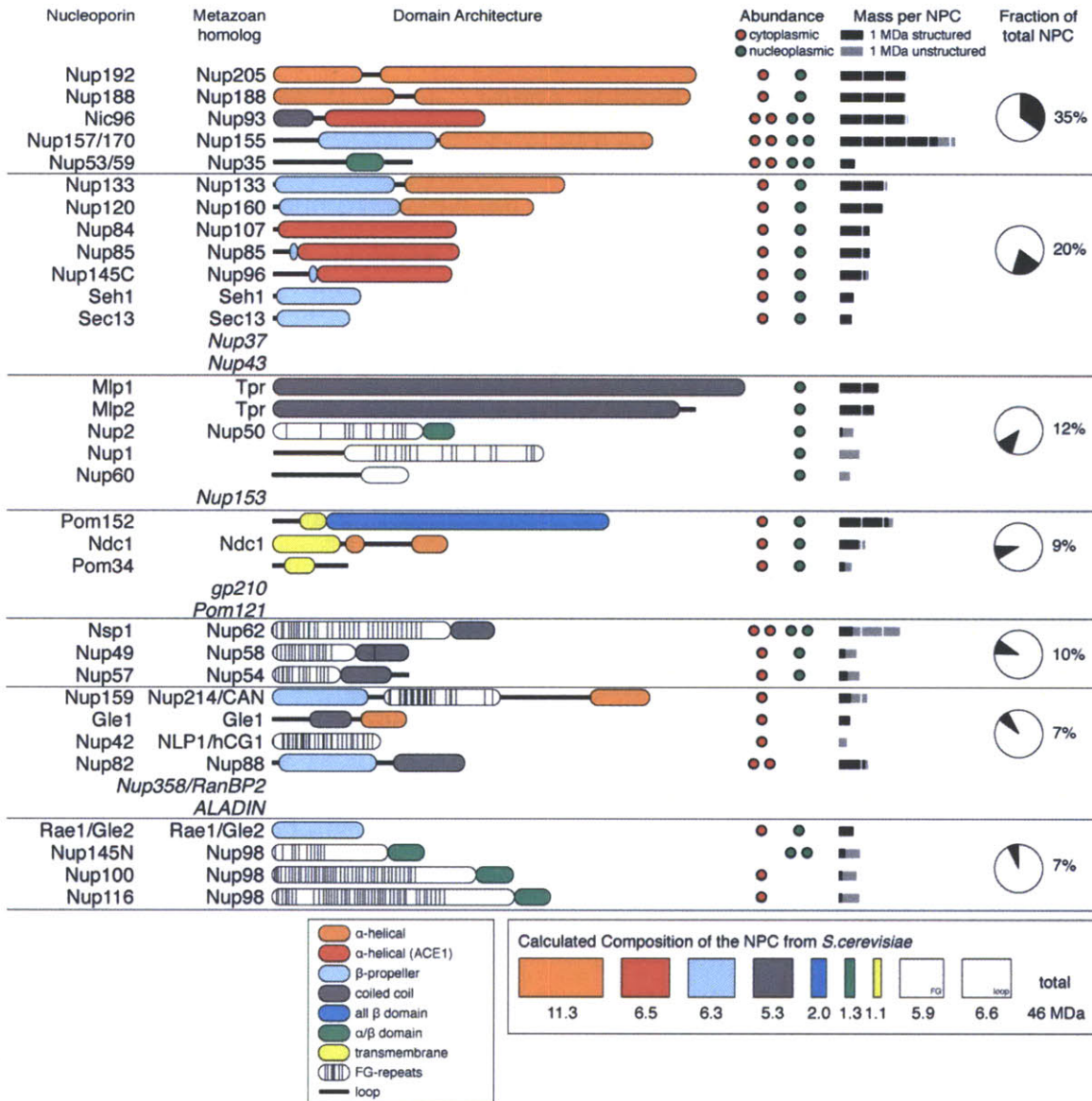
Figure 1.3 - Inventory of the NPC
Summary of the nucleoporins that make up the NPC. Domain architecture of nucleoporins from *S.cerevisiae* as determined by x-ray crystallography or prediction (where structural information is still lacking). Abundance and derived mass calculations are based on published Nup/NPC stoichiometries (Cronshaw et al., 2002; Rout et al., 2000). Nucleoporins specific to metazoa are italicized.

Subcomplexes

The majority of nucleoporins are organized in discrete subcomplexes each present in multiple copies that arrange according to the symmetry elements of the NPC to form the complete structure. The subcomplexes are biochemically

defined and reflect the stable interaction of subsets of nucleoporins. Interestingly, these subcomplexes are also found as entities in mitotic extracts of higher eukaryotes, when the nuclear envelope breaks down during open mitosis (Matsuoka et al., 1999). At the end of mitosis, NPCs reassemble from these subcomplexes in a defined order (Dultz et al., 2008). Each of the eight scaffold segments arranged around the central rotational axis is composed of 5 subcomplexes (Figure 1.2). Nup82/Nup159/Nsp1 form a subcomplex localized at the cytoplasmic side of the NPC (Belgareh et al., 1998). A second pool of Nsp1 complexes with Nup57 and Nup49 and resides in the center of the NPC, forming the bulk of the central transport barrier (Grandi et al., 1993). The scaffold ring is constructed from two major subcomplexes: the heptameric Y- or Nup84-complex and the heteromeric Nic96 complex. The Y complex is the best-characterized subcomplex of the NPC and is essential for its assembly, as shown in several organisms (Boehmer et al., 2003; Fabre and Hurt, 1997; Galy et al., 2003; Harel et al., 2003; Walther et al., 2003). It has 7 universally conserved components – Nup84, Nup85, Nup120, Nup133, Nup145C, Sec13 and Seh1 – that assemble stoichiometrically and exhibit the eponymous Y-shape in electron micrographs (Kampmann and Blobel, 2009; Lutzmann et al., 2002; Siniossoglou et al., 2000). In many eukaryotes, notably excluding *S. cerevisiae*, three additional proteins, Nup37, Nup43, and ELYS/MEL-28, are considered members of the Y complex, but their architectural role is unclear (Cronshaw et al., 2002; Franz et al., 2007; Rasala et al., 2006). In most models, the Y complex is thought to symmetrically localize to the cytoplasmic and the nucleoplasmic face of the NPC sandwiching the Nic96 complex. The Nic96 complex is not as well defined as the Y complex, likely reflecting the fact that it associates less stably. However, Nic96 interacts directly with Nup53/59 (Hawryluk-Gara et al., 2005), and co-immunoprecipitation with Nup188 (Nehrbass et al., 1996) and with Nup192 have been reported (Kosova et al., 1999). Further, the Nic96 complex is the tether to the Nsp1 complex in the center of the NPC. The newest defined subcomplex contains the transmembrane Nup Ndc1, considered to be an anchor for the NPC in the pore membrane. This complex contains Nup157/170 and Nup53/59, which connect

the Ndc1 complex to the Nic96 complex (Makio et al., 2009; Onischenko et al., 2009). The other two transmembrane nups, Pom34 and Pom152, are reported to interact with Ndc1 as well, albeit less strongly. Mlp1/2 are attached to the NPC ring via Nup60 (Feuerbach et al., 2002) and likely form the nuclear basket structure (Strambio-de-Castillia et al., 1999).

Dynamics

An important aspect of the NPC is that it is not a rigidly assembled machine, but a rather dynamic entity. Inverse fluorescence recovery after photobleaching experiments using GFP-tagged nucleoporins showed that different parts of the NPC have drastically different residence times (Rabut et al., 2004). Some mobile components detach from the NPC within seconds, while other components are stable throughout the entire cell cycle. Notably, the components of the structural scaffold – the Y complex and the Nic96 complex – are stably attached, while FG-nucleoporins are more dynamic. These studies on nucleoporin dynamics are consistent with the very slow protein turnover of scaffold nucleoporins (D'Angelo et al., 2009; Daigle et al., 2001). The scaffold structure of the NPC can be viewed as a docking site for more mobile nucleoporins, which often have functional roles at sites away from the NPC (Kalverda and Fornerod, 2007; Xylourgidis and Fornerod, 2009).

Domain architecture

Until about five years ago, very little high-resolution structural information on nucleoporins was available. This was largely due to the technical difficulties of obtaining nucleoporins of sufficient quantity and quality for structural studies, a challenge particularly severe in the case of scaffold nucleoporins. Despite the scarcity of experimental evidence, structural predictions grouped nucleoporins into a small set of fold classes (Berke et al., 2004; Devos et al., 2004; Devos et al., 2006; Schwartz, 2005). First, FG-domains, the primary transport factor interaction sites, are present in about one third of all nucleoporins. Second, coiled-coil domains are present in a number of nucleoporins. Third, scaffold

nucleoporins are largely composed of β-propellers, α-helical domains, or a tandem combination of both. Using this simple classification, about 76 % of the mass of the yeast NPC was accounted for.

FG-repeats

A total of 13 % of the NPC mass is made up of FG-repeat containing peptide stretches. The repeats are found in terminal extensions of ~10 nucleoporins and make up the physical transport barrier. NTRs specifically interact with the FG-regions, which allow them to enter the central transport channel. How FG-repeat regions exactly form the transport barrier is vigorously investigated and hotly debated (Frey and Gorlich, 2007; Lim et al., 2007; Peters, 2009; Rout et al., 2003). Systematic deletion of FG-regions from different nups has shown that the total mass of these filaments is more important than any one individual FG-filament, arguing for substantial redundancy in the meshwork (Terry and Wente, 2007). The intrinsic disorder of the FG-filaments is well documented in a series of crystal structures (Bayliss et al., 2000; Fribourg et al., 2001; Grant et al., 2003; Liu and Stewart, 2005). Only short peptide stretches are orderly bound to the convex outer surface of the HEAT-repeats that build NTRs, with the phenylalanine sidechains inserting between neighboring helices. Otherwise, the filaments remain without structure. Little is known about the intervening, non-FG sequences. They are poorly conserved, but are rich in polar and charged residues, probably important for the biophysical properties of the transport barrier.

Coiled-coils

Coiled coils in the NPC fulfill structural roles. The nuclear basket of the NPC is mainly constructed from the large coiled-coil proteins Mlp-1/2 in yeast and Tpr in vertebrates. Coiled-coils are often used for protein-protein interactions, thus the nuclear basket may serve as a general recruitment platform to bring accessory factors close to the NPC. The desumoylating enzyme Ulp1, for example, is stably associated with the nuclear basket (Li and Hochstrasser, 2000). Lining the

central NPC channel are 6 nucleoporins containing coiled-coil regions. The FG-Nup Nsp1 is part of two distinct entities, the Nsp1•Nup57•Nup49 complex (Grandi et al., 1993) and the Nsp1•Nup82•Nup159 complex (Bailer et al., 2001). In both, the proteins are held together by coiled-coil interactions (Bailer et al., 2001) and the Nsp1•Nup57•Nup49 complex is, in addition, tethered to the NPC scaffold via the N-terminal coiled-coil region of Nic96 (Grandi et al., 1995). So far, only a homodimerized 10 kDa fragment of Nup58 (the vertebrate ortholog to Nup57) has been structurally characterized (Melcak et al., 2007). Biochemical analysis suggests that the network involves specific rather than promiscuous interactions, arguing for a specific tethering function for the coiled-coil segments. It will be interesting to see these coiled-coil interactions in atomic detail in order to manipulate them and potentially swap the attached FG-domains within the NPC. Such experiments could provide important insight into the organization of the FG-network, if there is such.

β-propellers

A large portion of the NPC scaffold is build from β-propellers, one of the most abundant classes of proteins, especially in eukaryotes, and with diverse functions (Chaudhuri et al., 2008; Paoli, 2001). A set of Nups were initially identified as β-propellers based on sequence analysis. In yeast, only Sec13 and Seh1 contain the signature WD-40 repeat motif and were among the very first β-propellers to be recognized (Pryer et al., 1993). More nups have since been recognized as β-propellers despite the lack of signature sequence motifs. The N-terminal domain of Nup133 was the first experimentally determined β-propeller of the NPC and after this structure was solved, the additional non-canonical β-propeller domains in the NPC were identified (Berke et al., 2004). To date, 5 of the 8 universally conserved β-propellers in the NPC are structurally characterized (Figure 1.4). In Nup133, Nup120, and Nup159 (hNup214) the β-propellers are N-terminal and seven-bladed. While forming a distinct entity in Nup133 and Nup159 (Weirich et al., 2004), physically tethered but otherwise not interacting strongly with the C-

terminal part of the protein, the β-propeller in Nup120 is fully integrated with an adjacent helical domain to build one continuous oblong domain (Leksa et al., 2009). Seh1 and Sec13 are so far unique variations of β-propellers in that they are open and 6-bladed (Brohawn et al., 2008; Debler et al., 2008; Fath et al., 2007; Hsia et al., 2007). Their partner proteins insert a seventh blade into the β-propeller to complete the domain *in trans*. The function of the β-propellers is architectural and it is widely assumed that they serve as protein-protein interaction sites. Peripheral β-propellers can recruit accessory proteins, like the mRNA export factor Dbp5 (von Moeller et al., 2009), whereas those more centrally located likely are used to connect subcomplexes.
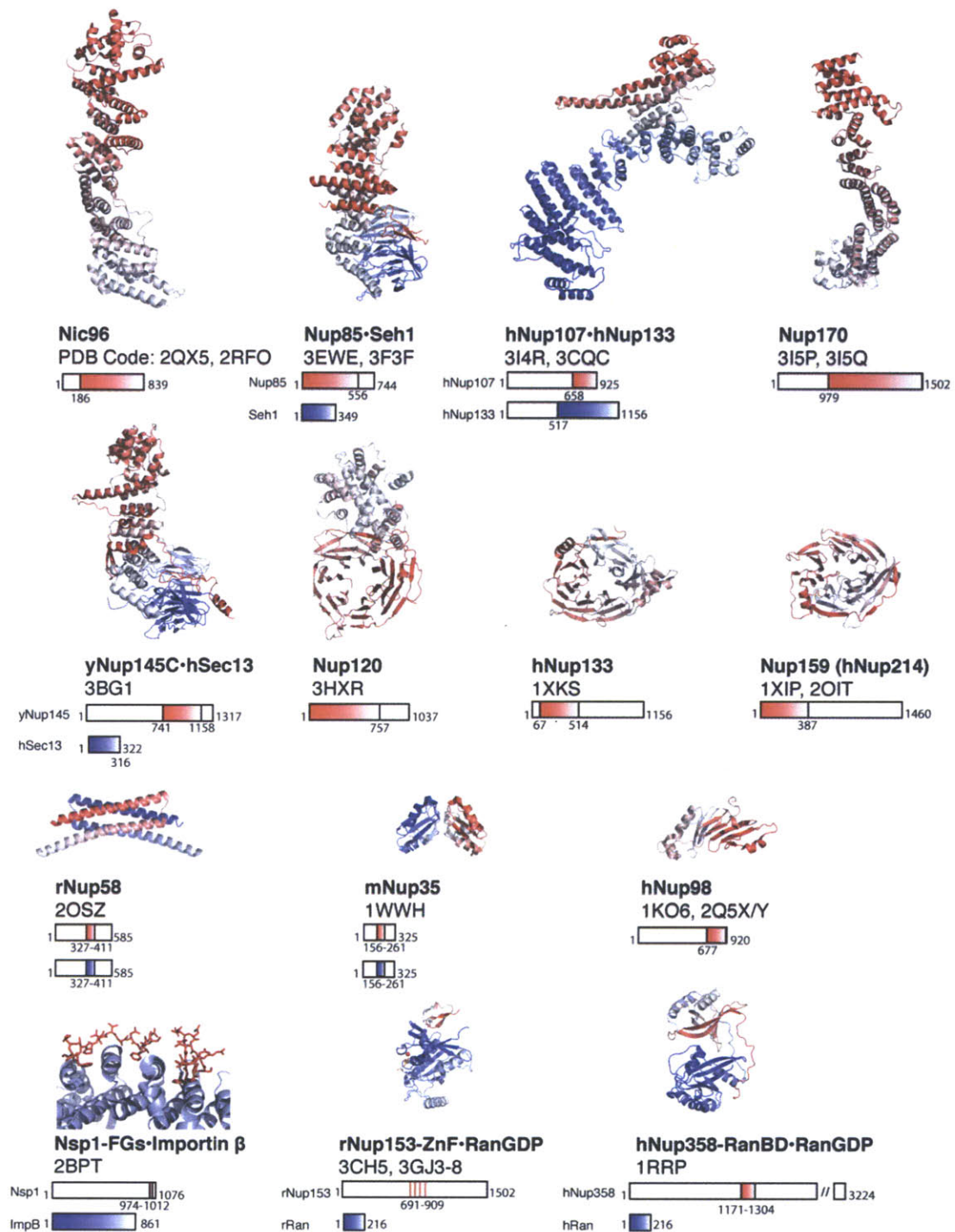
Figure 1.4 - Structures of nucleoporins

Comprehensive list of all representative nucleoporin structures published to date. PDB accession codes are indicated. Structures are gradient-colored red- or blue-to-white from N to C terminus. Residue information for each crystallized fragment is given below the structure. Structures are shown in the assembly state that is supported by crystallographic and biochemical evidence. Structures are from *S.cerevisiae* unless noted otherwise (h, human; m,mouse; r, rat). 2QX5(Jeudy

28

and Schwartz, 2007); 2RFO(Schrader et al., 2008b); 3EWE(Brohawn et al., 2008); 3F3F(Debler et al., 2008); 3CQC(Boehmer et al., 2008) 3I4R, 3I5P, 3I5Q(Whittle and Schwartz, 2009), 3BG1(Hsia et al., 2007); 3HXR(Leksa et al., 2009); 1XKS(Berke et al., 2004); 1XIP(Weirich et al., 2004); 2OIT(Napetschnig et al., 2007); 2OSZ(Melcak et al., 2007); 1WWH(Handa et al., 2006); 1KO6(Hodel et al., 2002); 2Q5X/Y(Sun and Guo, 2008); 2BPT(Liu and Stewart, 2005); 3CH5(Schrader et al., 2008a); 3GJ3-8(Partridge and Schwartz, 2009); 1RRP(Vetter et al., 1999).

α-helical domains

α-helical domains make up more than half of the mass of the NPC scaffold. Structural prediction classified the non-coiled-coil α-helical domains into a strongly related group of α-helical solenoids (Devos et al., 2006). α-solenoids are characterized by a two or three helix unit that is repeatedly stacked to form an elongated, often superhelical domain with N and C terminus at opposite ends of the molecule (Kobe and Kajava, 2000). Such regular, α-helical repeat structures are, often in combination with β-propellers, common scaffolds in large protein assemblies such as the clathrin vesicle coat (Edeling et al., 2006), the protein phosphatase 2A holoenzyme (Xu et al., 2006) and the anaphase promoting complex (Herzog et al., 2009), to name a few. Surprisingly, structural characterization of α-helical domain containing Nups has revealed three different α-helical folds, each distinct from a regular α-solenoid arrangement (Boehmer et al., 2008; Jeudy and Schwartz, 2007; Leksa et al., 2009; Schrader et al., 2008b; Whittle and Schwartz, 2009). Nic96 was the first experimentally determined α-helical structure of a scaffold nucleoporin and it showed an unexpected, atypical α-helical topology (Jeudy and Schwartz, 2007; Schrader et al., 2008b). The 30 helices of the ~65 kDa domain, excluding the ~200 residue N-terminal coiled-coil domain, are arranged in a J-like topology, forming an oblong domain. The chain starts in the middle of the elongated domain, zigzags up on one side of the molecule, folds back over a stretch of 7 helices and then continues past the N terminus to the other end of the molecule (Figures 1.4, 1.5). Three other α-helical scaffold nucleoporins (Nup84, Nup85 and Nup145C) have since been structurally characterized and shown to adopt the same fold as Nic96, pointing to a common ancestor (Brohawn et al., 2008) (see below). A second, distinct α-helical fold has

29

been identified in structures of Nup133 and Nup170, which are more distantly related, but share an extended and stretched α-helical stack (Boehmer et al., 2008; Whittle and Schwartz, 2009), substantially different from the first group. The third was revealed in the structure of Nup120, which forms a domain that fully integrates a β-propeller with an α-helical domain (Leksa et al., 2009). The α-helical segment is built around a central stalk of two long helices wrapped with 9 additional helices in an unprecedented fashion. In summary, the α-helical domains that occur in the NPC fall in different classes that provide a significant challenge for structure prediction methods. One obvious challenge is the exceedingly low sequence conservation, even between orthologs, apparent in the inconsistent nucleoporin nomenclature. Poor sequence conservation is likely due to some degree of malleability of the scaffold structure and the construction from common sequence elements (Aravind et al., 2006). Whether poor sequence conservation is further the result of adaptive evolution, linking several architectural nucleoporins to speciation, is an intriguing possibility that should be explored in more detail (Presgraves et al., 2003; Tang and Presgraves, 2009).

A

ACE1 · Nic96

Crown α5-11 · Trunk α1-3,α13-19 · Tail α21-28

B

ACE1 helix	Nic96	Nup145C	Nup85	Nup84	Sec31
1	205-228	188-196	103-117	n.d.	417-424
2	234-243	217-223	134-164	35-54	428-435
3	248-263	229-242	170-188	60-84	441-455
4	270-294	253-287	201-212	114-129	459-462
5	333-343	290-301	244-253	171-190	510-519
6	345-356	304-314	257-266	192-203	524-531
7	366-373	317-326	280-292	206-218	536-543
8	388-396	331-346	300-317	240-253	549-560
9	406-416	352-363	326-339	258-268	568-577
10	426-430	373-379	343-348	273-277	582-586
11	432-442	382-392	354-362	282-305	594-603
12	457-469	399-410	372-381	324-335	608-624
13	478-485	417-427	390-398	342-352	628-638
14	486-498	431-441	406-411	354-368	642-660
15	500-513	447-460	415-427	386-404	666-685
16	534-543	469-485	462-475	409-422	697-712
17	550-562	488-497	480-491	n.d.	717-724
18	566-583	501-514	497-509	n.d.	731-744
19	586-611	526-530	517-528	n.d.	n.d.
20	617-634	534-553	532-543	n.d.	n.d.
21	637-647	n.d.	n.d.	667- 680	n.d.
22	650-667	n.d.	n.d.	686-700	n.d.
23	685-697	n.d.	n.d.	705-723	n.d.
24	708-728	n.d.	n.d.	739-768	n.d.
25	732-742	n.d.	n.d.	783- 820	n.d.
26	765-790	n.d.	n.d.	841-867	n.d.
27	800-820	n.d.	n.d.	871-882	n.d.
28	826-835	n.d.	n.d.	895-913	n.d.

Figure 1.5 - The Ancestral Coatomer Element ACE1

Four scaffold nucleoporins, Nic96, Nup85, Nup84, and Nup145C share a distinct 65 kDa domain also found in the COPII protein Sec31, manifesting common evolutionary ancestry between the two membrane coats. (Brohawn et al., 2008; Devos et al., 2004). (A) The structures of an "average" ACE1 protein (left) and Nic96 (right) are shown in cartoon form and colored from blue (N-terminus) to white (C-terminus). The regions corresponding to the crown, trunk, and tail modules are indicated. The "average" ACE1 protein was made for illustrative purposes by superimposing the known ACE1 protein structures crown, trunk, and tail modules separately and constructing each of the 28 helices and connecting loops in the most frequently observed position. The orientation of the modules with respect to one another is shown in the average ACE1 case to most clearly show the architecture and connectivity of the fold. The position of the crown and tail modules in Nic96 are rotated 60° and 20° respectively compared to the average ACE1 structure as indicated. (B) The amino acid assignment for each of the canonical 28 ACE1 helices is indicated for all structurally characterized ACE1 proteins. All numbering is from *S. cerevisiae* except for helices 21-28 from Nup84, which correspond the human homolog Nup107. Helices are colored according to the module to which they belong as are the labels in (A). n.d. (not determined) indicates helices that fall within regions of the proteins for which no structural information is available.

31

ACE1 domains

As mentioned above, the four α-helical scaffold nucleoporins Nic96, Nup145C, Nup85, and Nup84 are constructed around a common ~65 kDa domain composed of 28 helices. Notably, this domain has to date only been identified outside the NPC in Sec31, one of the main building blocks of the COPII vesicle coat. The commonality was surprising. Sequence conservation between the five members is so low that no specific structural relationship was inferred previously (Alber et al., 2007a; Hsia et al., 2007). This domain, which we termed Ancestral Coatomer Element 1 (ACE1), is a structural manifestation of the likely common origin of the NPC and the COPII vesicle coat (Devos et al., 2004). ACE1 is constructed from three modules, crown, trunk and tail, that together form an elongated molecule of ~ 140Å x 45Å x 45Å. Structural superposition of ACE1 proteins shows that individual modules are closely aligned, while differences in linkers between modules results in significant differences in their relative orientations. These differences, as well as proteolytic susceptibility data, suggest at least modestly flexible hinges connect the modules, especially the trunk and the tail. This likely explains why all crystal constructs except for Nic96 contain either the trunk and crown (Brohawn et al., 2008; Debler et al., 2008; Hsia et al., 2007) or the tail (Boehmer et al., 2008). Even with the structural information in hand, it is difficult to find additional ACE1 proteins. Beyond a few residues conserved between orthologs, ACE1 is not characterized by a distinct sequence motif. The reason for this amazing degeneracy on the sequence level likely is that for folding the ACE1 domain only some general sequence profiles need to be satisfied. For example, helices α5, α7, α15 and α17 are typically hydrophobic, because they are incased by surrounding helices and are largely buried and solvent inaccessible. Thus, a combination of sequence profile evaluation, α-helical prediction, and overall length are currently the only indicators for the ACE1 domain. The two remaining α-helical scaffold nucleoporins without any crystallographic structural information are Nup188 and Nup192. Whether or not they belong to the ACE1 class, remains to be determined, but it appears unlikely.

Assembly

Structures of large protein assemblies are typically solved by a combination of cryo-electron microscopy and x-ray crystallography (Chiu et al., 2006), exploiting the strengths of both methods. With EM-data in the 10-15 Å range, fitting of crystal structures can often be performed fairly reliably, providing in effect a composite structure at atomic resolution. In the case of the NPC, the best tomographic EM data for the full assembly does not extend beyond ~ 58 Å (Beck et al., 2007) and thus, unfortunately, does not lend itself to directly fitting the available crystal structures. A promising step toward closing the resolution gap is the recent EM reconstruction of the Y complex at ~ 35 Å resolution, which allows at least tentative fitting of crystal structures (Kampmann and Blobel, 2009). Apart from the resolution gap, however, the biggest difficulty currently is the absence of strong experimental data on how the subcomplexes, notably the Y complex, assemble to form the core scaffold structure of the NPC.

In an alternative approach to the hybrid method of combining experimental EM- and crystallographic data, Alber et al. used a plethora of volumetric and stoichiometric data, combined with distance restraints obtained from a comprehensive co-immunoprecipitation analysis of all nucleoporins, to solve the subcomplex assembly puzzle (Alber et al., 2007a; Alber et al., 2007b)). Even though each datapoint has very limited information content by itself, a useful three-dimensional model is generated by combining all the data, conceptually similar to the way an NMR structure is computed. The resulting draft model at an estimated 5 nm resolution provides a plausible arrangement of Nups in the yeast NPC, information that is not available from the current tomographic studies. According to the computed model, 8 Y complexes each self-assemble into a cytoplasmic and a nucleoplasmic ring, defining the Z-dimension of the scaffold at ~38 nm (Figure 1.6, middle row labeled "computational model"). Sandwiched between the two rings are two 8-membered rings of Nup157/170, Nup188, and Nup192. All remaining nucleoporins, except for the filamentous nuclear basket and cytoplasmic extensions, have been positioned as well and decorate the main

scaffold. Clearly, this model has severe limitations and will have to be adjusted with more detailed information becoming available. For instance, the Y complex is modeled as a compact rod in the computational model and the coat is space-filling with very small gaps, both in apparent contrast to results from crystallographic analyses. The exciting possibility of this combinatorial approach however is that it should be possible to integrate currently available and forthcoming high-resolution structures to further refine the model. To date, crystal structures accounting for 23% of the mass of the NPC scaffold are available. Integration of these data could potentially reveal an NPC structure that would come fairly close to the reality.
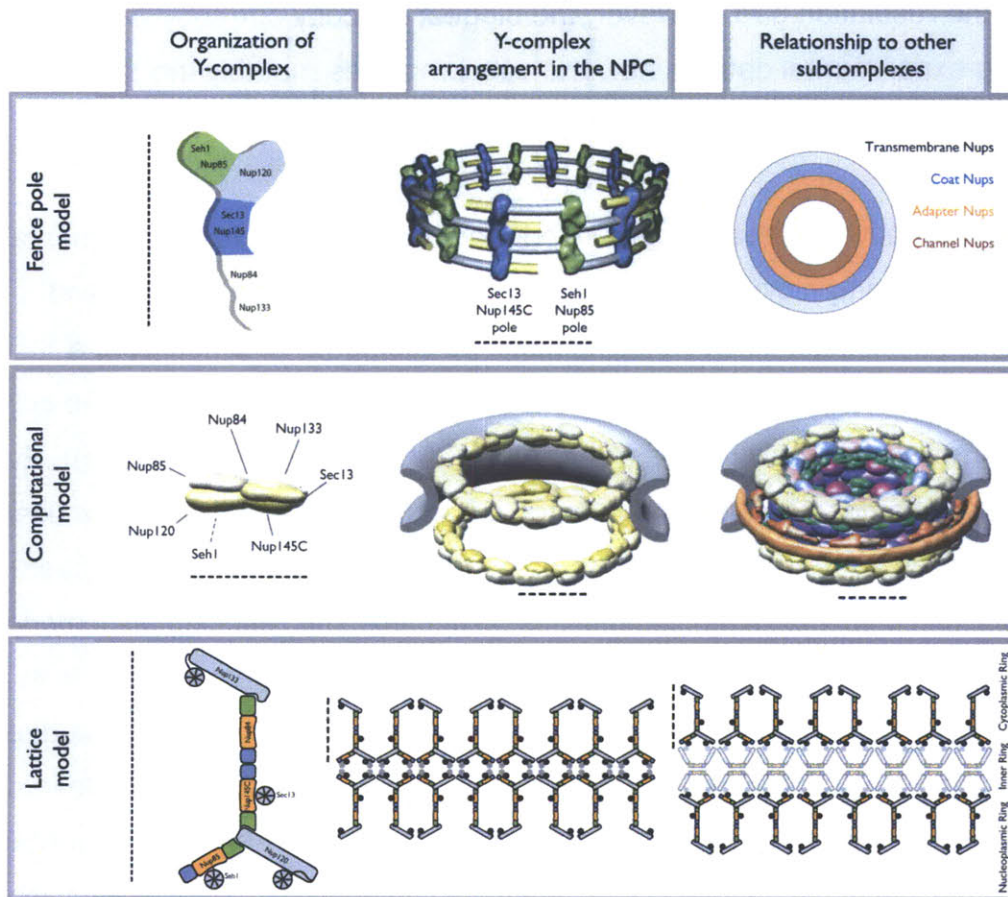
Figure 1.6 - Assembly models for the NPC
Three recently proposed models for the structural organization of the NPC are illustrated. The fence pole model (top row), computationally generated model (middle row), and lattice model (bottom row) are compared in their prediction of protein interactions within one Y complex (left column), Y complex organization

within the NPC (middle column), and placement of Y complexes relative to other NPC subcomplexes (right column). Dashed lines to the left or underneath panels represent ~40 nm (The approximate height of a Y complex) and are shown for scale. In the fence pole model, heterooctameric poles of Nup145C•Sec13 and Nup85•Seh1 observed in crystal structures organize four rings of 8 Y complexes each. These four rings form a cylinder adjacent to the transmembrane Nups on the membrane side and layers of adaptor Nups followed by channel Nups towards the transport channel (Hsia et al., 2007; Debler et al., 2008). The computationally generated model provides localization volumes for each Nup and shows 8 Y complexes arranged into two separate rings, one towards the nucleocytoplasmic and the other the cytoplasmic side of the NPC. The other Nups are arranged into membrane rings, inner rings, linkers between rings, and FG nucleoporins (Alber et al., 2007b). The lattice model is based on structural homology of ACE1 proteins in the NPC and COPII vesicle coat (Brohawn et al., 2008). ACE1 proteins are colored by module with tails green, trunks orange, and crowns blue. A model of a single Y complex incorporates the demonstrated specific interactions between domains of the 7 proteins. 8 Y complexes are assumed to form a nucleoplasmic and cytoplasmic ring, which may or may not be connected by additional lattice elements such as Nic96 and Nup157/170 in an inner ring. The illustration is meant to emphasize the predicted open, lattice-like organization of the NPC structural scaffold and is not meant to imply specific interactions between complexes. While it seems likely that the NPC lattice will be composed of ACE1-containing edge elements and vertex elements made from β-propeller interactions as observed in the COPII vesicle coat, the exact nature of the vertex elements in the NPC lattice remains to be seen (Brohawn et al., 2008).

Blobel and coworkers have taken a different experimental approach to address the subcomplex assembly problem. Their rationale is that the subcomplexes assemble only at very high protein concentrations as they are found in the assembled NPC in the living cell. Such conditions are difficult to reproduce *in vitro*, but they can potentially be mimicked in protein crystals where the protein concentration is similarly very high. Thus, crystal-packing interactions between dimeric subcomplex fragments have been used to develop an assembly model for the NPC (Debler et al., 2008; Hsia et al., 2007). Besides the overall dimensions of the scaffold, the resulting model is drastically different from the computational model. Four instead of two 8-membered rings of Y complexes are stacked on top of each other to form a continuous membrane-proximal shell. The rings are connected by alternating heterooctameric poles of Nup145C•Sec13 and Nup85•Seh1 (Figure 1.6, top row labeled "fence pole model"). The Nic96 complex is postulated to form a second, inner shell within the NPC that connects to the FG-network. Similar to the computational model, the Y complexes are

envisioned to directly contact the pore membrane via membrane-inserting ALPS-motif containing amphipathic helices predicted to be present in several Y complex components (Drin et al., 2007). However, experimental evidence supporting membrane-insertion is so far only available for an ALPS-helix found in the vertebrate Nup133 β-propeller, and the prediction of ALPS-helices is non-trivial. Indeed, several predicted ALPS-helices have been found to be well packed and to contribute to the hydrophobic core of nucleoporins (for example in Nup85 and Nup120), making their involvement in membrane-insertion unlikely.

Comparison to vesicle coats

In a groundbreaking paper, it was postulated that the NPC and vesicle coats share a common ancestor, dating back more than one billion years to the very early eukaryote (Devos et al., 2004). The principal hypothesis was that these assemblies are used to fulfill similar roles specific to eukaryotes, namely stabilizing the highly curved membranes of vesicles and the circular openings in the nuclear envelope. Judged by primary sequence analysis, however, the relationship between the components is largely undetectable. One important pillar of the 'proto-coatomer' hypothesis is that Sec13 is a bona fide component of both the COPII vesicle coat and the NPC. Structural evidence supporting the hypothesis was provided by showing that a class of four scaffold nucleoporins shares a specific 65 kDa domain, ACE1, with Sec31 of the COPII vesicle coat (Brohawn et al., 2008; Brohawn and Schwartz, 2009a). The COPII coat is organized into a membrane-proximal inner layer built from Sec23•Sec24 heterodimers and Sar1, and an outer coat assembled from Sec13•Sec31 heterodimers (Stagg et al., 2008). Sec31 in the COPII coat and Nup145C in the NPC interact very similarly with Sec13 by insertion of a 7th blade to complete the β-propeller. This interaction mode is recapitulated in Nup85 binding to the Sec13 homolog Seh1, which is also facilitated by insertion-completion of the β-propeller. Since Nup145C, Nup85, and Sec31 all share the same ACE1 domain, the structural data not only suggests a common ancestor, but it also provides clues as to how the NPC assembles.

In comparison to the NPC, the COPII vesicle coat is much simpler, has been extensively studied, and its assembly is quite well understood. Most informative with respect to the NPC assembly is the organization of the Sec13•Sec31 outer coat of COPII. Here, two Sec13•Sec31 heterodimers dock via their ACE1 crown modules to form an edge element (Fath et al., 2007). Four edge elements converge in a vertex, where the N-terminal β-propellers of Sec31 interact. COPII vesicles of different sizes can be assembled from different numbers of edge elements by adjusting the angles at the interaction sites (Stagg et al., 2008). In analogy to the Sec31•Sec31 homodimer interface in the COPII coat, it was predicted on a biochemical and structural basis that Nup145C and Nup84 form a heterodimeric edge element via a crown-to-crown interface in the NPC. Due to the predicted steric conflicts, this arrangement would be fundamentally incompatible with the hetero-octameric fence pole model discussed above (for details, see (Brohawn et al., 2008)). We predict that the entire NPC scaffold has an open, lattice-like organization with significant similarity to the COPII coat, also built from vertex and edge elements (Figure 1.6 bottom row labeled "lattice model", Figure 1.7). As much as COPII and clathrin lattices seem not to share common construction principles (Fath et al., 2007), so do similarities between NPC and clathrin lattices not extend beyond superficial characteristics.
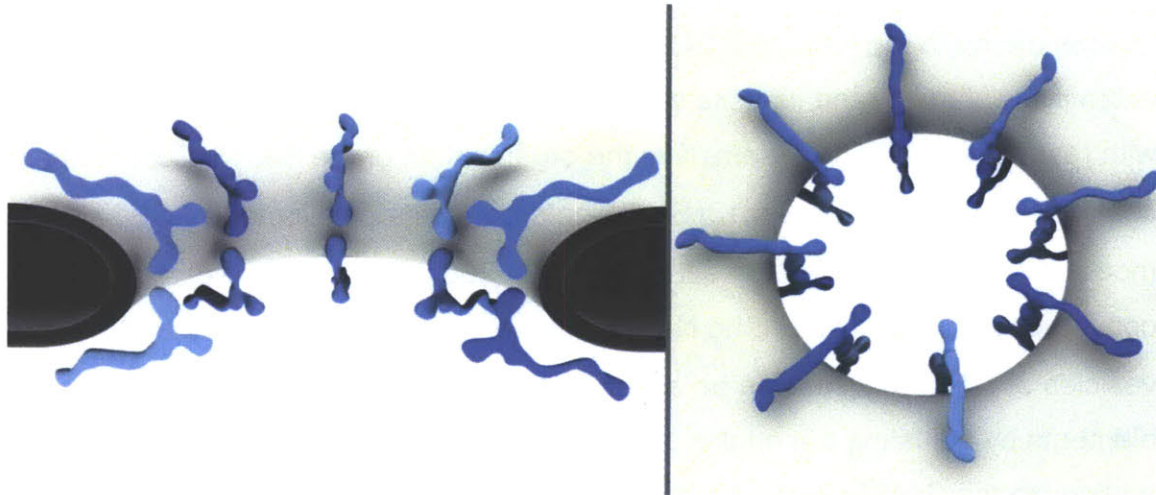
Figure 1.7 – Predicted lattice-like arrangement of Y complexes in the NPC
The predicted arrangement of Y complexes (blue) in the NPC with respect to the nuclear envelope (pale yellow) is shown in a cut-away side view on the left (with 10 of the 16 Y complexes in the NPC shown) and from a top view on the right. The Y complexes are predicted to be part of an exterior lattice in which ACE1 edge elements of Nup84•Nup145C are arranged parallel to and stabilize (but not directly contact) the positive curvature of the nuclear envelope. This is analogous to the role of ACE1 edge elements of Sec31 in the COPII outer lattice.

One important distinction between the computational and the hetero-octameric pole model on one hand and the lattice model on the other is that in the latter the Y complex and likely also the Nic96 complex components Nic96 and Nup157/170 do not directly coat the membrane, although there may be punctual contacts. By analogy to the Sec23•Sec24 inner coat of COPII, adaptor molecules instead are predicted to directly contact the pore membrane. The transmembrane Nups, especially Ndc1, are prime candidates for this function. Indeed, Ndc1 has already been shown to physically interact with the Nic96 complex via Nup53/59 (Onischenko et al., 2009). Ndc1 is found in all eukaryotes and is essential, though this could also be related to its additional role in the spindle pole body (Winey et al., 1993). Such adaptor-mediated anchoring to the pore membrane is consistent with the membrane-proximal gaps observed in cryo-tomographic reconstructions (Beck et al., 2007; Stoffler et al., 2003).

Obviously, there are some important limitations to the analogy to COPII. First, it is not yet apparent how the vertices are formed in the NPC. In COPII, Sec31

contains an additional N-terminal β-propeller, absent in the ACE1 domains of the NPC, but essential for COPII vertex assembly (Fath et al., 2007). Thus, the vertex might look different in the NPC. Further, the COPII coat is a self-enclosed lattice structure that completely encapsulates the positively-curved membrane vesicle. In the NPC, the lattice can be continuous in lateral direction around the central transport gate, but must terminate on the nucleo- and cytoplasmic sides. Further, the circular pore opening has positive curvature only in the axial direction, but negative curvature in equatorial direction. Since all vesicle coats seamlessly wrap positively curved membranes, it is not unreasonable to speculate that a specific architectural element might occur exclusively in the NPC to establish its unique geometrical requirements.

Outlook

Progress toward the structural characterization of the NPC at high resolution, one of the holy grails of structural biology (Bhattacharya, 2009), has been spectacular since our last review on the subject, a review that was intended as a roadmap toward this ambitious goal (Schwartz, 2005). The past five years have seen improvement in images of the entire NPC available, and a flurry of crystal structures of various key parts of the NPC. New computational methods were developed to integrate geometric data from various sources, and these were applied to the NPC, in part to demonstrate a general means to tackle other large assemblies (Alber et al., 2008). The path toward a high-resolution structure of the NPC will continue to be highly interdisciplinary. Single-particle reconstruction on subcomplexes has still been used only sparingly, but holds great promise to close the gap between the resolutions provided by crystallography and by tomography. It is likely that integration of structural data, particularly crystallographic data, into computationally derived models will provide further useful information. Finally, developments in super-resolution microscopy promise to allow *in vivo* measurement of distances within the NPC on the order of 10 nm, a potential source of further useful parameters – parameters so far inaccessible experimentally. These distances would complement other available structural

39

data. Judging by recent progress in the field, dramatic developments in our understanding of the NPC lie just around the corner.

Acknowledgements

Chapter two

Structural evidence of a common ancestry of the nuclear pore complex and vesicle coats

Abstract

Nuclear pore complexes (NPCs) facilitate nucleocytoplasmic transport. These massive assemblies comprise an eightfold symmetric scaffold of architectural proteins and central-channel phenylalanine-glycine-repeat proteins forming the transport barrier. We determined the nucleoporin 85 (Nup85)•Seh1 structure, a module in the heptameric Nup84- or Y complex, at 3.5 angstroms resolution. Structural, biochemical, and genetic analyses position the Y complex in two peripheral NPC rings. We establish a conserved tripartite element, the ancestral coatomer element ACE1, that reoccurs in several nucleoporins and vesicle coat proteins, providing structural evidence of coevolution from a common ancestor. We identified interactions that define the organization of the Y complex on the basis of comparison with vesicle coats and confirmed the sites by mutagenesis. We propose the NPC scaffold, like vesicle coats, is composed of polygons with vertices and edges forming a membrane-proximal lattice that provides docking sites for additional nucleoporins.

Introduction

Exchange of macromolecules across the nuclear envelope is exclusively mediated by NPCs (D'Angelo and Hetzer, 2008; Tran and Wente, 2006; Weis, 2003). Whereas much progress has been made understanding the soluble factors mediating nucleocytoplasmic transport, the structure of the ~40-60 MDa NPC itself is still largely enigmatic. Cryo-electrontomography (cryo-ET) and cryo-electronmicroscopy (cryo-EM) have established the NPC structure at low resolution (Beck et al., 2007; Drummond et al., 2006; Stoffler et al., 2003). Crystal structures of scaffold NPC components are emerging (Berke et al., 2004; Boehmer et al., 2008; Hsia et al., 2007; Jeudy and Schwartz, 2007), but the resolution gap still precludes fitting into the cryo-ET structure. Overall, the NPC has eight-fold rotational symmetry with an outer diameter of ~100 nm and a core scaffold ring ~30 nm wide. The central FG-repeat containing transport channel measures ~40 nm in diameter, defining the maximum size of substrates (Pante and Kann, 2002).

42

The modularity of the NPC assembly suggests a path toward a high-resolution structure (Schwartz, 2005). Of the ~30 *bona fide* nucleoporins (Nups) that comprise the NPC, only a core subset is stably attached (Rabut et al., 2004). In *Saccharomyces cerevisiae*, this core includes two essential complexes: the heptameric Y complex and the heteromeric Nic96-containing complex (hereafter called the Nic96 complex; unless noted all proteins are from *S. cerevisiae*). The Y complex is composed of one copy each of Nup84, Nup85, Nup120, Nup133, Nup145C, Sec13 and Seh1. It self-assembles from recombinant proteins *in vitro* and forms a branched Y-shaped structure (Lutzmann et al., 2002). Deletion or depletion of individual components of the Y complex leads to severe assembly defects in many organisms (Fabre and Hurt, 1997; Galy et al., 2003; Harel et al., 2003). The Nic96 complex is less well characterized, but appears to contain the architectural nucleoporins Nup157/170, Nup188, Nup192, Nup53, and Nup59 (Lusk et al., 2002; Marelli et al., 1998; Zabel et al., 1996). β-propellers and stacked α-helical domains form the building blocks of the constituents of the Nup84 and Nic96 complexes (Devos et al., 2006; Schwartz, 2005). Because vesicle coats (including COPI, COPII, and clathrin) share similar elements, a common ancestry has been hypothesized despite very low sequence homology and the absence of experimental structural evidence (Devos et al., 2004).

A recent computer-generated model of the NPC based on a plethora of primary data from different sources placed the Y complex at the NPC periphery, sandwiching the Nic96 complex in the center (Alber et al., 2007b). In contrast, a model solely based on the structure of the Nup145C•Sec13 heterodimer and crystal-packing interactions was proposed that is inconsistent with the computer model (Hsia et al., 2007).

Results

We solved the structure of a complex of Nup85 residues 1-564 (of 744) and intact Seh1 (referred to as Nup85•Seh1) at 3.5 Å (Table 2.1). Seh1 and Nup85 form distinct units in a tightly associated complex (Figure 2.1). Seh1 folds into an

open six-bladed β-propeller structure. The blades fan out consecutively around a central axis, typical for canonical β-propeller structures (Chaudhuri et al., 2008). Between blades 1 and 6, the N-terminus of Nup85 is inserted and forms a three-stranded blade that completes the Seh1 propeller *in trans*. Following its N-terminal insertion blade, Nup85 forms a compact cuboid structure composed of 20 helices, with two distinct modules, referred to as 'crown' and 'trunk'. Helices α1-α3 (residues 100-200) meander along one side of the trunk; the other side is formed by helices α12-α19 (residues 362-509) running in the opposite direction in an antiparallel zigzag to the C terminus. The trunk elements are separated by an intervening crown composed of helices α4 to α11 (residues 201-361) that form a distinct bundle that caps one end of the trunk. Helices α5 to α10 in the crown module are almost perpendicular to the helices in the trunk.

Table 2.1 - Data collection and refinement statistics

	$Nup85_{1\text{-}564}$•Seh1 Native	Nup85•Seh1 Selenomethionine
Data collection		
Space group	$P4_12_12$	$P4_12_12$
Wavelength (Å)	0.9792	0.9792
Cell dimensions		
$a=b, c$ (Å)	112.6, 350.5	112.5, 351.2
$\alpha=\beta=\gamma$ (°)	90	90
Resolution (Å)	40-3.4	50-3.7
R_{sym} (%)	10.7	12.7
$I/\sigma I$	13.7 (1.8)	21.9 (1.8)
Completeness (%)	96.9 (98.8)	99.1 (99.9)
Redundancy	3.1 (3.3)	6.3 (6.1)
Refinement		
Resolution (Å)	30-3.5	
No. reflections	28300	
R_{work} / R_{free}	32.6 / 36.9	
No. atoms		
Protein	9636	
Water	0	
Wilson B-factor ($Å^2$)	119	
R.m.s deviations		
Bond lengths (Å)	0.008	
Bond angles (°)	1.2	
Coordinate error (Å)	0.58	
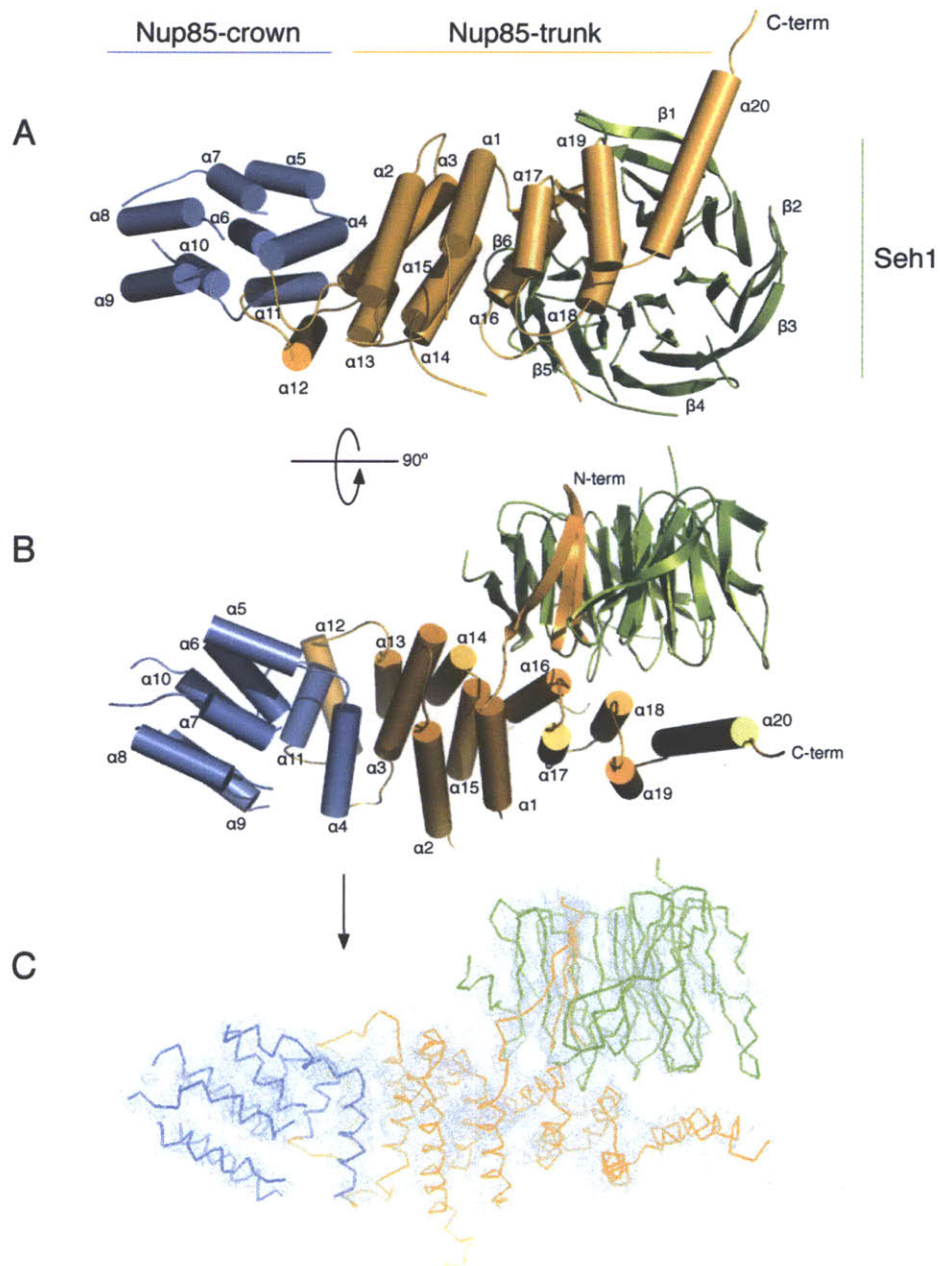
Figure 2.1 - Structure of the Nup85•Seh1 complex

The structure of the heterodimeric Nup85•Seh1 complex is shown in two views
(A, B), related by a 90° rotation around the horizontal axis. Nup85 has a trunk
(orange, helices α1-α3 and α12-α20) and a crown (blue, helices α4-α11) module.
The β-strands at the extreme N-terminus of Nup85 form an insertion blade, which
complete the Seh1 (green) β-propeller. (C) $2F_o$-F_c omit map (contoured at 1.2σ)
with a Cα-trace of the Nup85•Seh1 complex.

In the asymmetric unit of the crystal, two heterodimers are aligned along a non-crystallographic dyad, generating patches of contacts (Figure 2.2). This interaction is unlikely to be functionally meaningful because the contact residues are poorly conserved in orthologs. Moreover, analysis of Nup85•Seh1 by analytical ultracentrifugation (AUC) showed a single species of ~104 kDa with a hydrodynamic radius of 4.4 nm (Figure 2.3). This hydrodynamic radius is close to the theoretical value calculated from the atomic coordinates using HYDROPRO (Garcia De La Torre et al., 2000) and reflects the elongated shape of the 103 kDa Nup85•Seh1 complex (a spherical protein of 220 kDa would have the same radius). Gel filtration also showed that Nup85•Seh1 is a single 103 kDa heterodimer at concentrations up to 20 mg/ml (Figure 2.4). Hence, we restrict our analysis to this heterodimer.
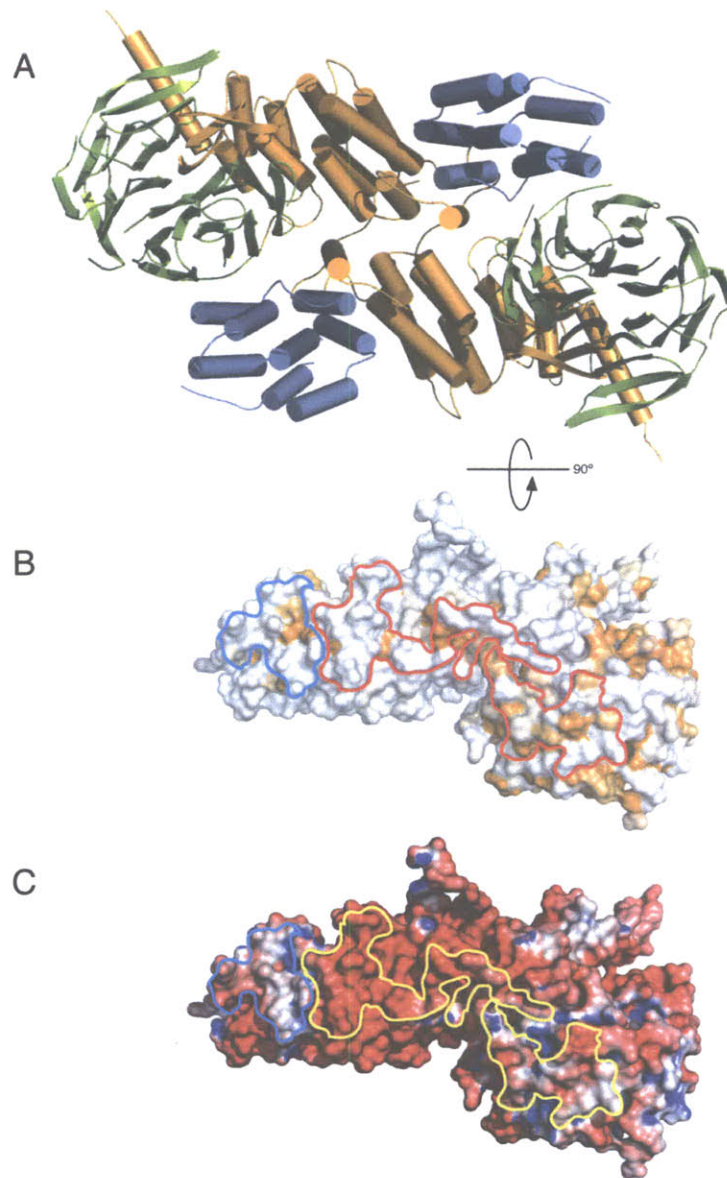
Figure 2.2 - Arrangement of two Nup85•Seh1 heterodimers in the asymmetric unit

(A) Association and orientation of the two Nup85•Seh1 heterodimers in the asymmetric unit. The heterodimers form an interface of ~1800 Å^2 and associate lengthwise along a two-fold axis. (B) Surface conservation of Nup85•Seh1 in a view 90° rotated from the lower molecule in (A) with outlines corresponding to contact regions involved in forming the interface. Red and blue outlines indicate contacts made with Nup85 and Seh1, respectively, and conservation is shaded from white (not conserved) to orange (conserved). (C) The electrostatic surface potential of the Nup85•Seh1 heterodimer (colored from red (-8 kT/e) to blue (+8 kT/e)) with outlines as in (B). Yellow and blue outlines correspond to contacts made with Nup85 and Seh1, respectively. The view is the same as in (B).
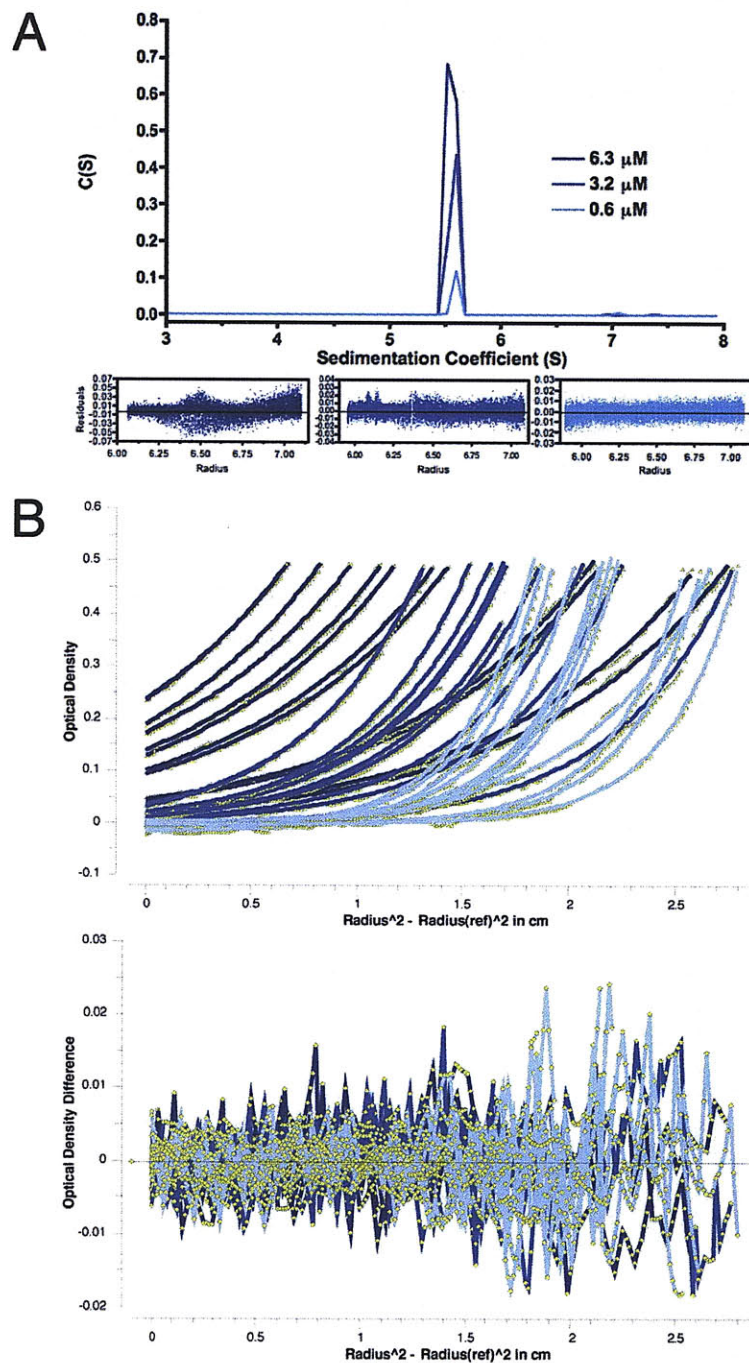
Figure 2.3 - Nup85•Seh1 is a dimer in solution as determined by analytical ultracentrifugation

(A) C(s) distribution analysis of Nup85•Seh1 sedimentation velocity data. Sedimentation data from three concentrations of Nup85•Seh1 were analyzed globally in Sedphat (Schuck, 2000) with a hybrid local continuous distribution and global discrete species model. Data was fitted from 2 to 10 s^{-13} with fixed partial specific volume. Residual plots for each concentration are shown below. The

sedimentation coefficient and corresponding rmsd values for the samples in order of decreasing concentration were 5.56 s^{-13}, 5.58 s^{-13}, and 5.60 s^{-13} and 0.0079, 0.0056, and 0.0045, respectively, which corresponds to a single species with a molecular weight of ~104 kDa and frictional ratio (f/f0) of 1.43. The calculated molecular weight for Nup85•Seh1 is 103 kDa. (B) Sedimentation equilibrium analysis of Nup85•Seh1. Sedimentation data from 6 concentrations at three speeds were analyzed globally in Ultrascan 9.0. The data was best fit by an ideal single species model. The top panel shows data points as yellow triangles with fit curves overlaid (13.5 krpm – dark blue, 17.5 krpm – medium blue, 22.8 krpm – light blue). The lower panel shows residuals of the fitted curves. The molecular weight was determined to be 99 kDa with a standard deviation of 0.4 kDa, closely matching the results obtained via sedimentation velocity.
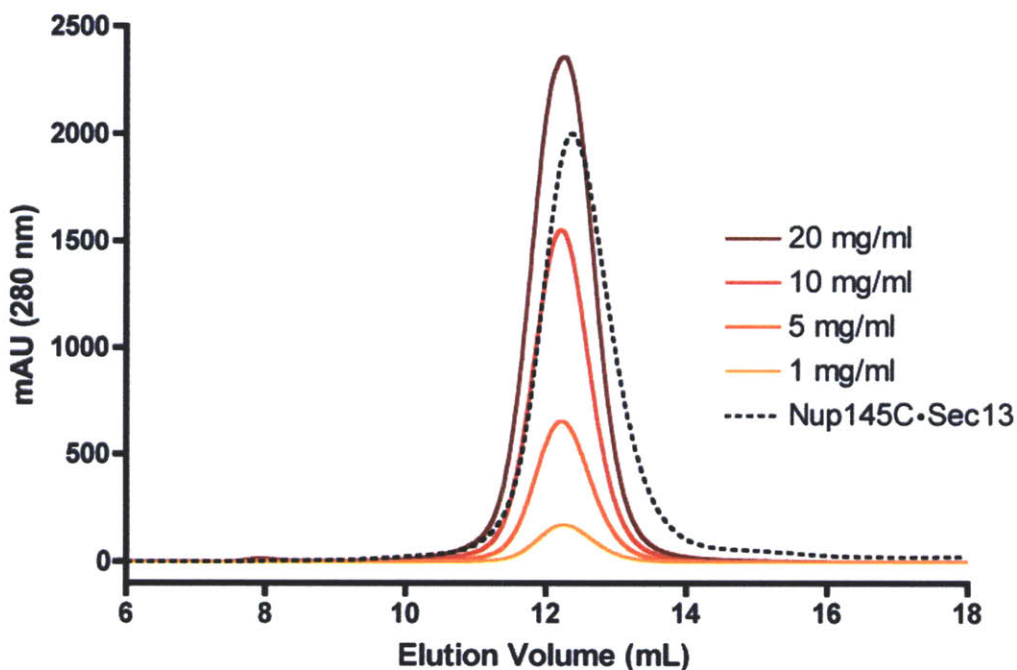
Figure 2.4 - Nup85•Seh1 is a dimer in solution as determined by size exclusion chromatography
Elution profiles of the Nup85$_{1-564}$•Seh1 complex at 1, 5, 10, and 20 mg/ml show a single peak eluting at 12.2 ml on a Superdex 200 10/300 (GE Healthcare) column indicating a hydrodynamic radius of 4.4 nm. The hydrodynamic radius was independently determined by sedimentation velocity and is consistent with the value calculated from the experimental crystal structure using HYDROPRO (Garcia De La Torre et al., 2000). The elution profile of Nup145C•Sec13 is shown for comparison (dashed black line).

The connectivity and topology of secondary structure elements and the three-dimensional folds of Nup85•Seh1 and Nup145C•Sec13 (Hsia et al., 2007) are

49

remarkably similar (Figure 2.5A), despite very low sequence identity between Nup85 and Nup145C (10%) and moderate identity between Seh1 and Sec13 (32%). Like Nup85, Nup145C has an N-terminal three-stranded β-sheet that provides a seventh blade to close the open β-propeller of Sec13. The trunk and crown modules of Nup145C are also similar to those in Nup85, although their relative orientation is modestly different in the two proteins.
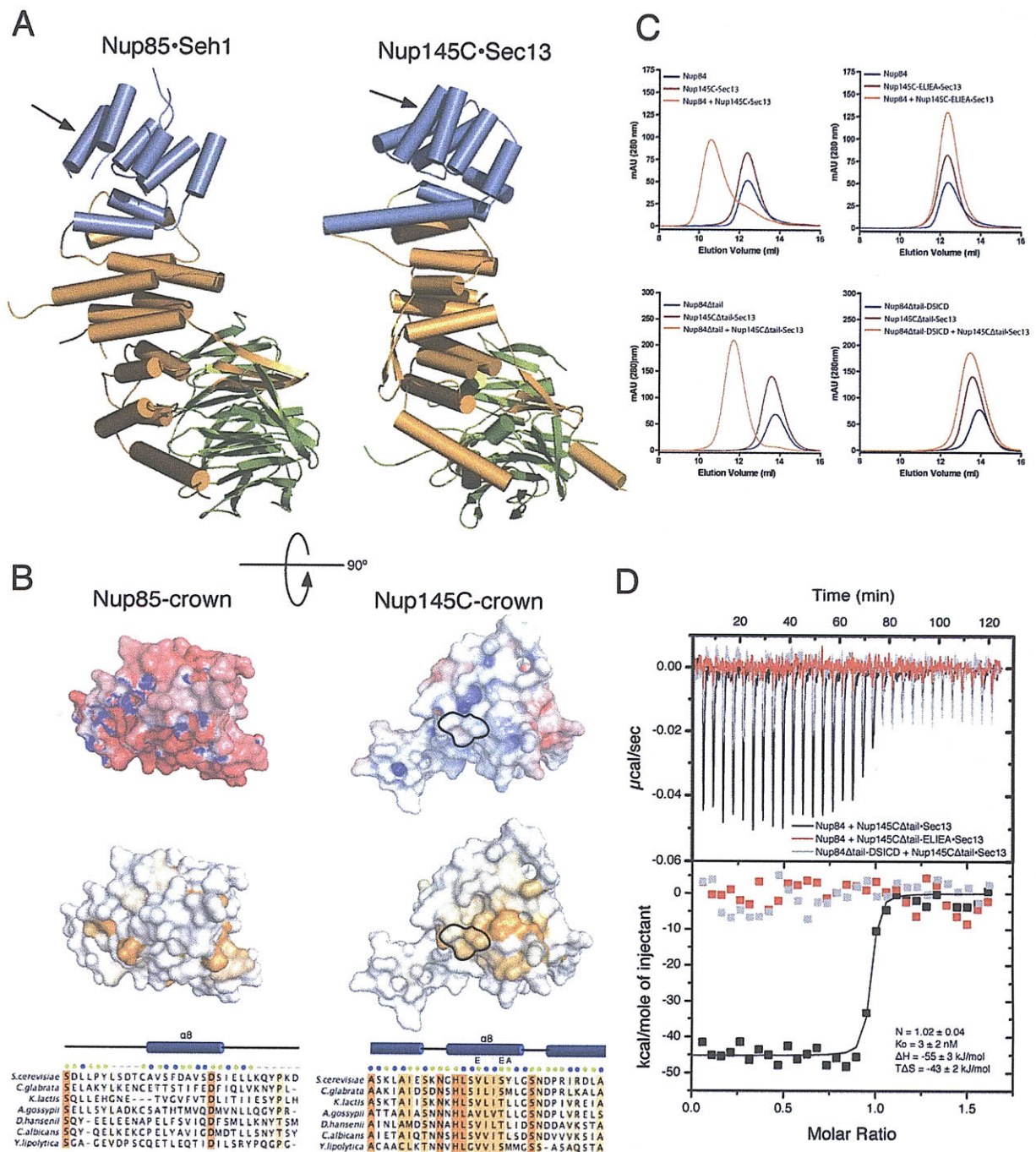
Figure 2.5 - Comparison of Nup85•Seh1 and Nup145C•Sec13 and identification of the Nup84•Nup145C crown-crown binding interface
(A) The topologies of the Nup85•Seh1 (left) and Nup145C•Sec13 (right, PDB accession code 3bg1 (Hsia et al., 2007)) complexes are shown, illustrating an overall similarity with three shared structural elements - trunk, crown, and β-propeller. Colors are assigned as in Figure 2.1. (B) Surface representations of the crowns of Nup85 and Nup145C are shown colored according to electrostatic surface potentials (top) and sequence conservation (bottom) in a view rotated 90° from (A). Sequence conservation is based on the phylogenetic tree of budding yeasts (Dujon, 2006) and is colored from white (not conserved) to orange (conserved). A partial sequence alignment of helix α8 (indicated by arrows in (A)) is also shown with surface exposed residues indicated by green dots, residues buried in the hydrophobic core by blue dots, and residues not modeled in the structure by dashes. Mutations made in this helix in Nup145C are shown above the sequence alignment and the corresponding residues are outlined in the surface representations of Nup145C. (C) In the upper panels, gel filtration data of Nup84 alone, Nup145C•Sec13 (wild type or -ELIEA mutant) alone, and Nup84 plus Nup145C•Sec13 (wild type or -ELIEA mutant) are shown. The shift in the Nup84 plus wild type Nup145C•Sec13 chromatogram indicates complex formation and is absent in the case of the -ELIEA mutant. In the lower panels gel filtration data of Nup145C•Sec13 alone, Nup84 alone (wild type or -DSICD mutant) alone, and Nup145C•Sec13 plus Nup84 (wild type or -DSICD mutant) are shown. The shift in the Nup145C•Sec13 plus wild type Nup84 chromatogram indicates complex formation and is absent in the case of the -DSICD mutant. (D) Isothermal titration calorimetry data illustrating high-affinity binding for wild type Nup145C•Sec13 and Nup84 (black). Experimental values for N, K_D, ΔH, and $T\Delta S$ are shown. In contrast, binding is lost for both crown-surface mutants Nup84-DSICD (grey) and Nup145C-ELIEA (red).

The most conserved regions of Nup85 are involved in the interaction with Seh1. The corresponding interface between Nup145C and Sec13 is also well conserved, but Nup145C has an additional highly conserved surface on the crown module around helix α8 that is not observed in Nup85 (Figure 2.5B). This region is reasonably polar and poorly conserved in Nup85 but highly conserved and distinctly hydrophobic in Nup145C, suggesting a protein-protein interaction site. Nup84 and Nup120 bind to roughly opposite sides of Nup145C•Sec13 in the Y complex (Lutzmann et al., 2002), and the C-terminal helical region of Nup145C is necessary for binding Nup120 (Figure 2.6). Thus, we hypothesized that the α8 crown surface of Nup145C is the binding site for Nup84.
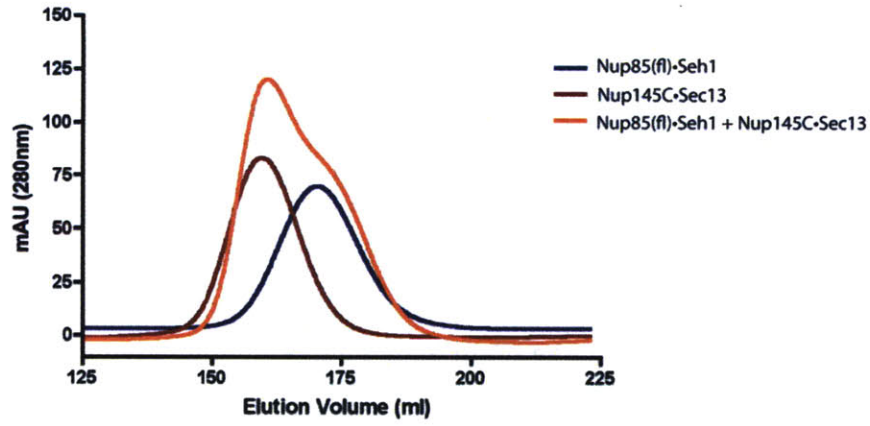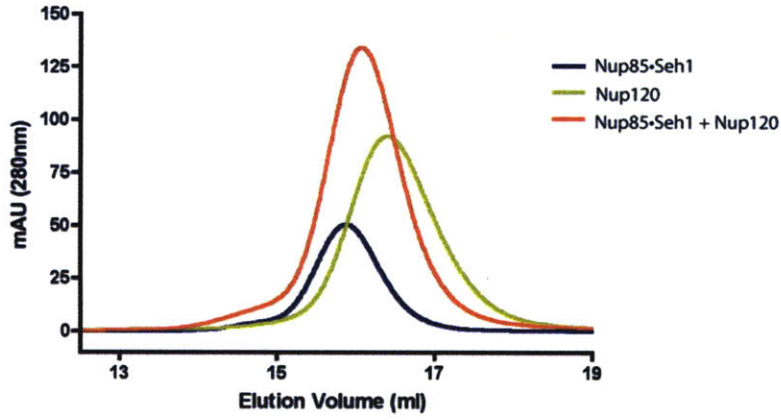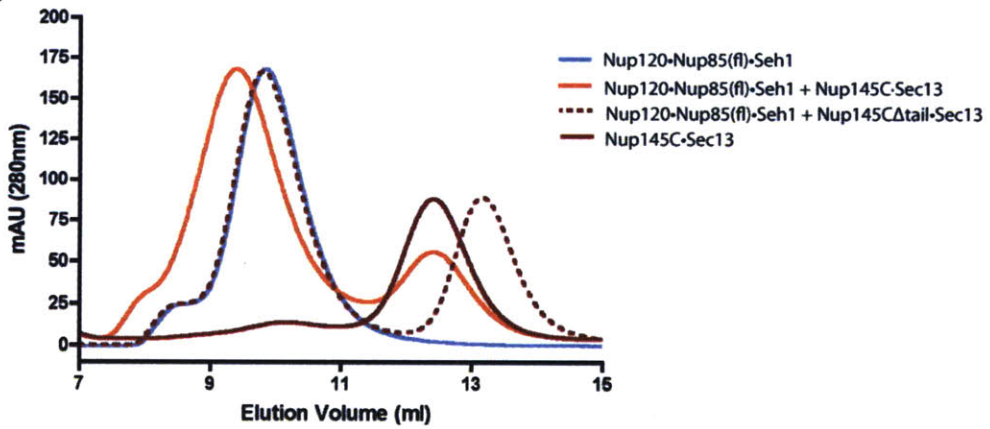
Figure 2.6 - Nup120 binds Nup145C and Nup85 via their tail modules
(A) The heterodimeric Nup85(fl)•Seh1 and Nup145C•Sec13 complexes were analyzed on a Superdex 200 HR26/60 column. A mixture of both complexes (orange) does not result in a higher molecular weight species indicating that the complexes do not directly interact. (B) The heterodimeric Nup85•Seh1 complex (without the Nup85 tail module) and Nup120 were analyzed on a Superose 6 HR 10/300 column. Again, a mixture of both samples (orange) does not result in a higher molecular weight species. (C) Nup85(fl)•Seh1 including the tail module binds Nup120 (blue), and adding Nup145C•Sec13 results in a pentameric complex (orange) (Superdex S200 HR10/300). This complex is not formed when the tail module is removed from Nup145C (dashed). Taken together, this series of experiments demonstrates that the tail modules of both Nup145C and Nup85 are responsible for Nup120 binding.

To test this hypothesis, we mutated the Nup145C sequence VLISY in $\alpha 8$ to ELIEA, introducing two negative charges and eliminating a conserved aromatic side chain on the crown surface (Figure 2.5B). The overall structure of Nup145C did not appear to be perturbed by this modification: (i) Nup145C-ELIEA•Sec13 bound to Nup120 to form a 1:1 complex indistinguishable from Nup145C•Sec13 in gel-filtration experiments (Figure 2.7); Nup145C•Sec13 and Nup145C-ELIEA•Sec13 complexes (ii) had comparable thermostability (Figure 2.8) and (iii) showed identical behavior in gel filtration (Figure 2.5C). The ELIEA mutation completely eliminated Nup84 binding. In isothermal-titration calorimetry (ITC) experiments, Nup84 bound wild-type Nup145C•Sec13 tightly ($K_D = 3 \pm 2$ nM; 1:1 stoichiometry) but not Nup145C-ELIEA•Sec13 (Figure 2.5D). Similarly, Nup84 formed a stable complex with Nup145C•Sec13 but not with Nup145C-ELIEA•Sec13 in gel filtration (Figure 2.5C). We conclude that the Nup84 binding site on Nup145C includes the exposed surface of helix $\alpha 8$.
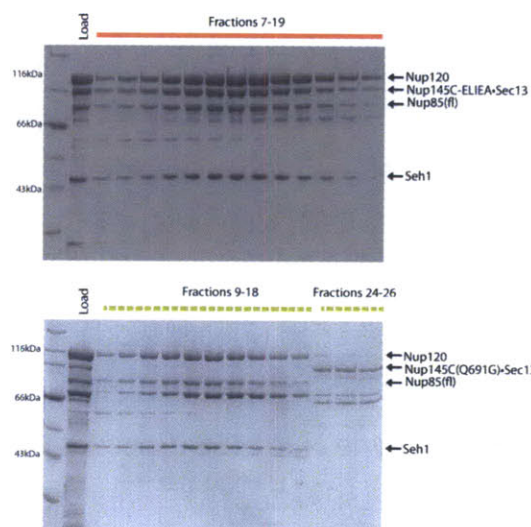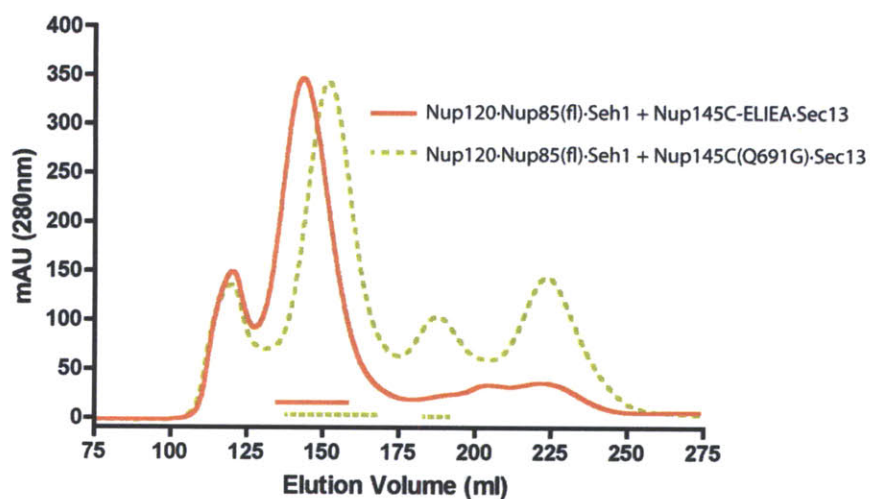
Figure 2.7 - A single point mutation in the predicted interaction helix of the Nup145C tail module disrupts Nup120 binding
Formation of a pentameric Nup120•Nup85(fl)•Seh1•Nup145C-ELIEA•Sec13 complex (orange) is disrupted by the Q691G mutation in the tail module of Nup145C (dashed) (Superdex S200 HR26/60). Formation of the Nup120•Nup85•Seh1 complex is unaffected by this mutation. Fractions were analyzed by SDS-PAGE.
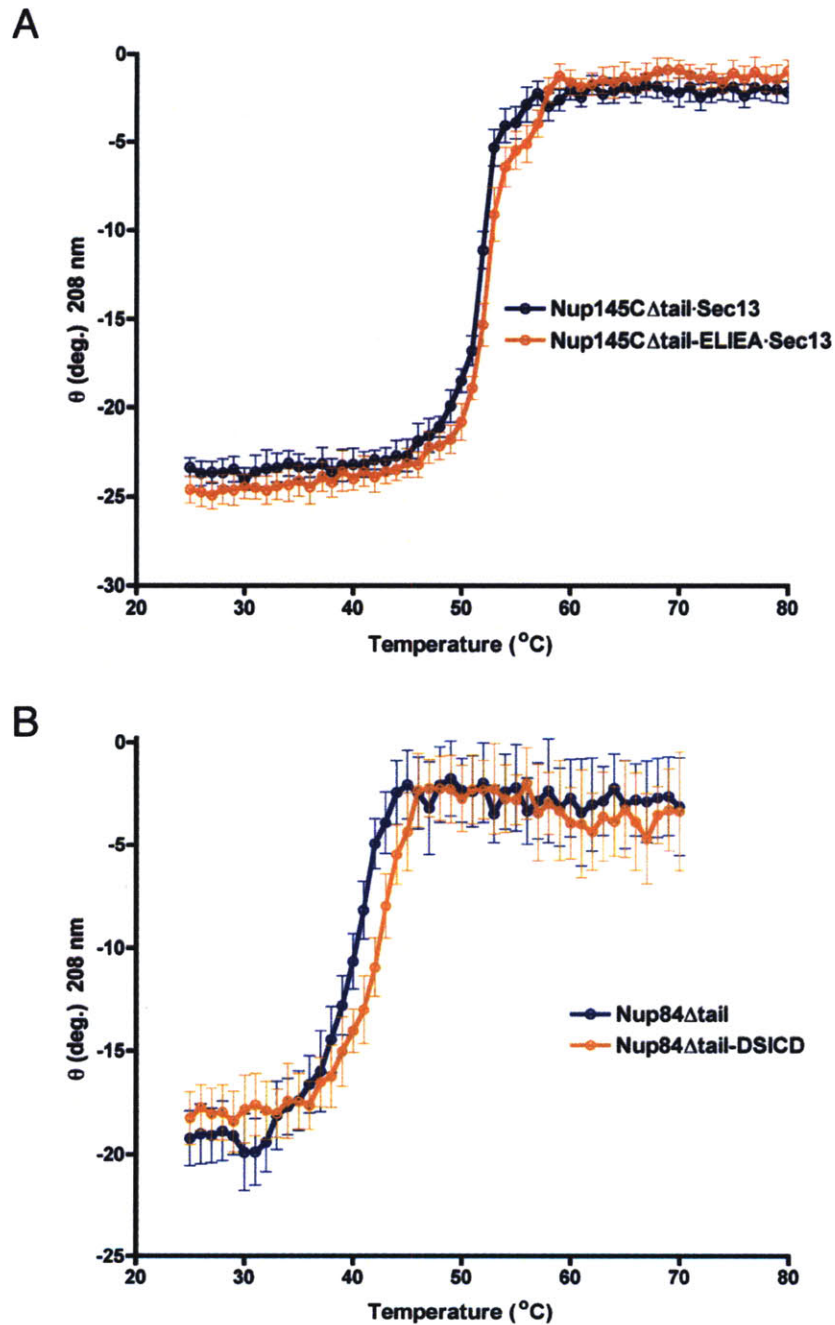
Figure 2.8 - Surface mutation of the Nup145C and Nup84 crowns does not negatively affect protein stability

(A) Heat denaturation of Nup145CΔtail•Sec13 and Nup145CΔtail-ELIEA•Sec13 monitored by circular dichroism. Ellipticity at 208nm is plotted as a function of temperature. The melting temperature, T_m, is nearly identical at ~52 °C for both protein complexes. (B) Nup84Δtail and Nup84Δtail-DISCD also have very similar thermal denaturation characteristics (T_m, ~41 °C) indicative of uncompromised protein stability.

To determine the consequences of abolishing the Nup84 binding site on Nup145C *in vivo*, we introduced the Nup145C-ELIEA mutation into the *NUP145* gene in yeast. Strains carrying *NUP145-ELIEA* in a Δ*NUP145/NUP84-GFP* or Δ*NUP145/NUP133-GFP* background displayed a marked defect in incorporating Nup84-GFP and Nup133-GFP into the NPC (Figure 2.9, 2.10). Compared with wild type, a significantly larger fraction of GFP-tagged protein was found in the cytoplasm, indicating that the Nup84 binding interface on Nup145C is crucial in recruiting both Nup133 and Nup84 to the NPC (Figure 2.11). In addition, nuclear pores were clustered into discrete foci on the nuclear envelope of the strains expressing Nup145C-ELIEA, indicative of severe NPC assembly defects and similar to Nup84 and Nup133 null strains (Doye et al., 1994; Siniossoglou et al., 1996). Cells expressing wild-type Nup145C demonstrated the expected punctate nuclear rim staining in both Nup84-GFP and Nup133-GFP strains. Thus, disruption of the Nup84 binding site on Nup145C affects NPC assembly and function and causes loss of Nup84 and Nup133 from pores. The loss of Nup133 can be rationalized because it is attached to the Y complex through a binary interaction with Nup84 (Boehmer et al., 2008; Lutzmann et al., 2002). Some Nup84 and Nup133 proteins remain associated with nuclear pores in the Nup145C-ELIEA expressing strains, arguing for the existence of additional weaker attachment sites for both proteins in the NPC. It has been shown that an ALPS membrane-binding motif is present in Nup133 (Drin et al., 2007). Because Nup133 and Nup84 are tightly associated (Boehmer et al., 2008), the ALPS motif might be weakly functional in recruiting Nup133•Nup84 to the NPC even when the Nup84•Nup145C interaction is compromised.

A

	DIC	Nup133-GFP	DNA	Overlay
NUP145				
nup145C-ELIEA				

B

	DIC	Nup84-GFP	DNA	Overlay
NUP145				
nup145C-ELIEA				
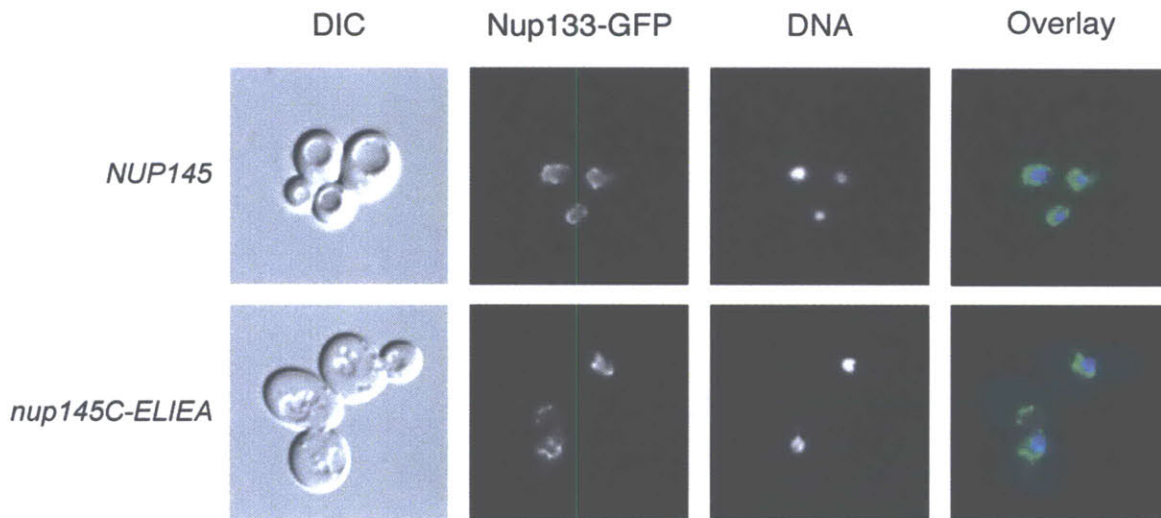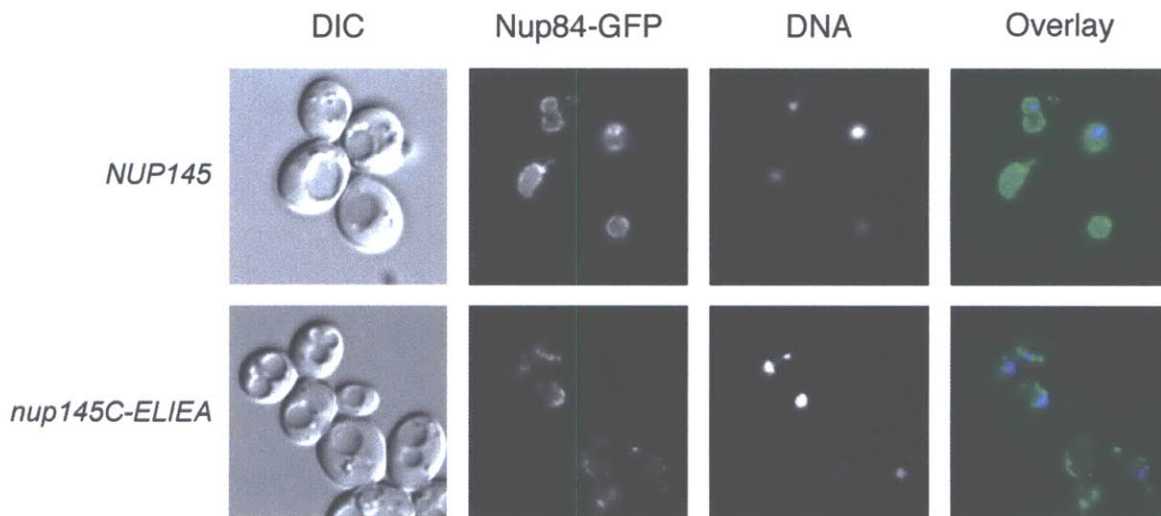
Figure 2.9 - Elimination of the Nup84 binding site on Nup145C results in nuclear pore assembly defects *in vivo*
(A) *NUP145/NUP133-GFP* and *NUP145-ELIEA/NUP133-GFP* or (B) *NUP145/NUP84-GFP* and *NUP145-ELIEA/NUP84-GFP* were grown at 24 °C and visualized by fluorescence microscopy. DIC, GFP-fluorescence, DNA (visualized with Hoechst dye), and false-colored overlay (GFP fluorescence – green, DNA – blue) images of the same field are shown in columns from left to right.
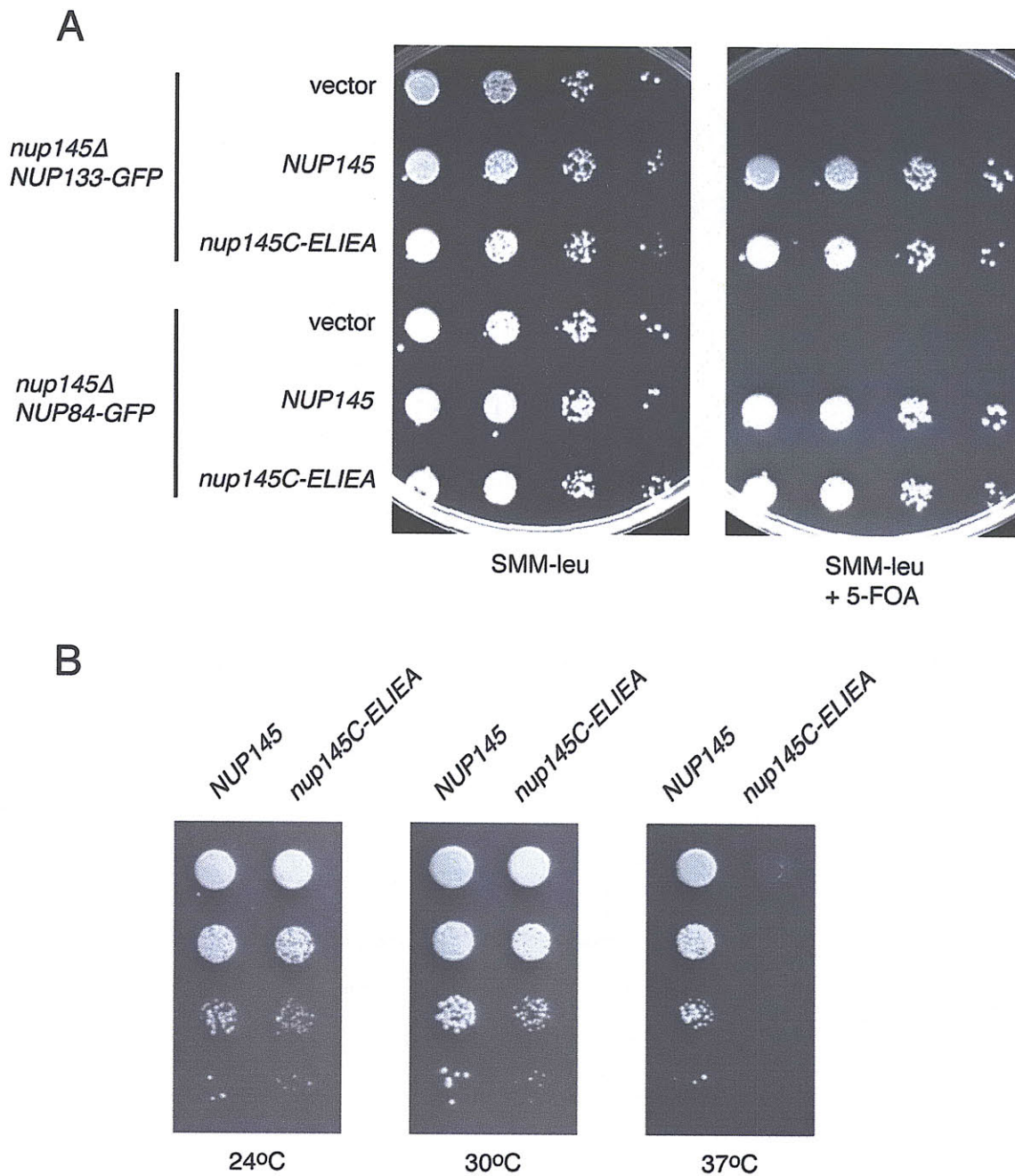
Figure 2.10 - Growth analysis of yeast strains
(A) The nup145C-ELIEA mutant supports viability. *nup145Δ/NUP133-GFP* and a *nup145Δ/NUP84-GFP* strain carrying SBYp115 (*NUP145/CEN/URA3*) and empty vector, or vector encoding *NUP145*, or vector encoding *nup145C-ELIEA* were grown in SMM-leu overnight, serially diluted and grown on SMM-leu or SMM-leu + 5-FOA plates at 30°C for two days. (B) *nup145C-ELIEA* is lethal at elevated temperatures in rich media. A *nup145Δ* strain carrying plasmid-borne *NUP145* (SBYp116) or *nup145C-ELIEA* (SBYp117) were grown in minimal medium at 24°C. Serial-diluted cells were plated onto YPD plates and grown for 2 days at 24°C, 30°C or 37°C.

A

NUP145   nup145c-ELIEA

        T   S   P   T   S   P

Nup133-GFP

Pgk1p
(cytosol)

B

NUP145   nup145c-ELIEA

        T   S   P   T   S   P

Nup84-GFP

Pgk1p
(cytosol)
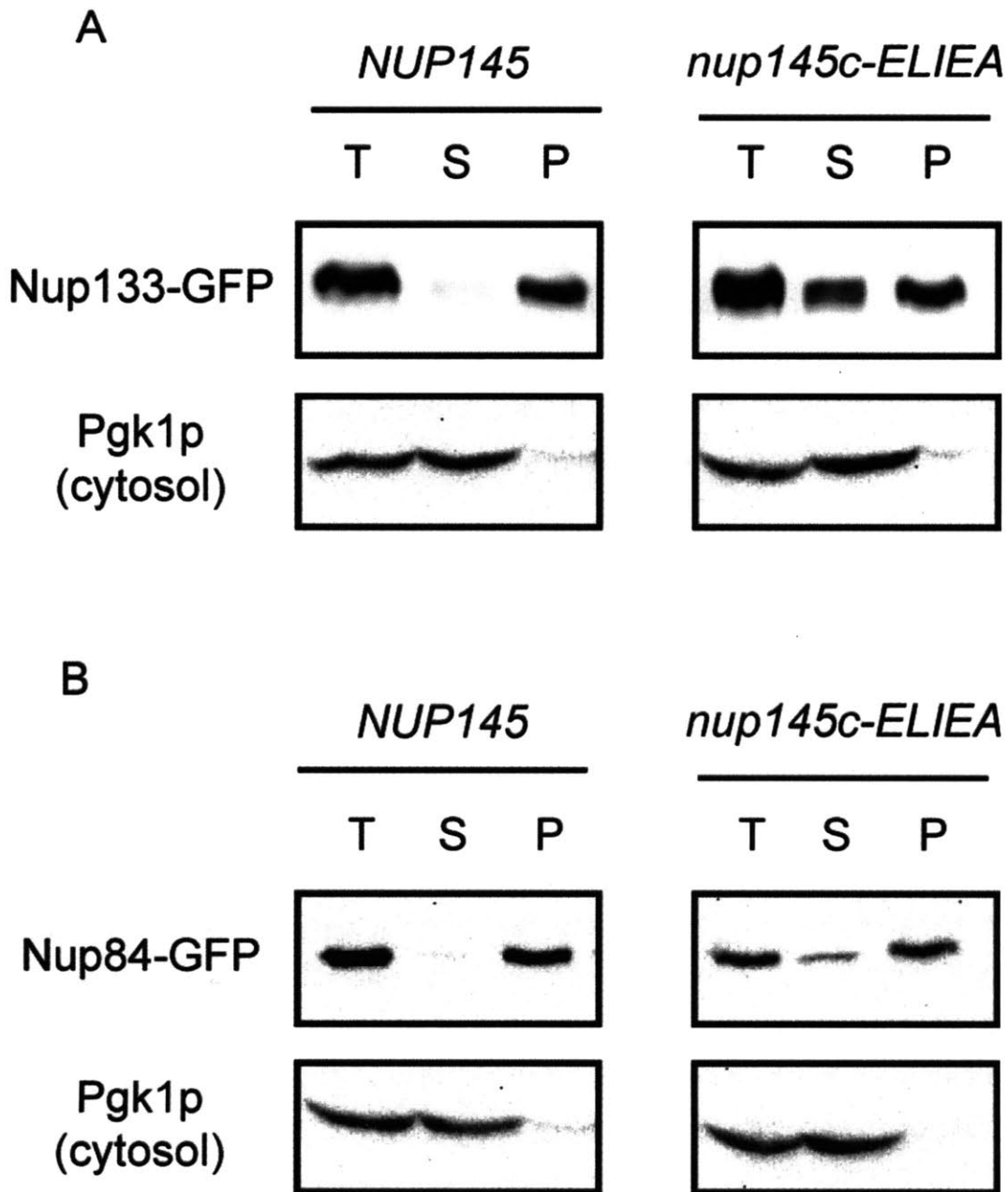
Figure 2.11 - Nup133-GFP and Nup84-GFP become more soluble in the Nup145-ELIEA strain
(A) *NUP145 NUP133-GFP* (YS221) and *nup145C-ELIEA NUP133-GFP* (YS222) or (B) *NUP145 NUP84-GFP* (YS223) and *nup145C-ELIEA NUP84-GFP* (YS224) strains were spheroplasted, and total lysates (T) were separated into 16,000xg soluble (S) and pellet (P) fractions. Equal cell equivalents from each fraction were analyzed by immunoblotting using rabbit anti-GFP or monoclonal anti-Pgk1 antibodies.

On the basis of lattice packing observed in crystals of Nup145C•Sec13, Hsia et al. (Hsia et al., 2007) proposed that Nup145C•Sec13 and Nup85•Seh1 each form heterooctameric poles that span the entire NPC in a "concentric cylinder" model of NPC structure. However, the Nup145C•Sec13 lattice contacts involved in the putative heterooctamer overlap with the crown surface of Nup145C shown here to be the Nup84-binding site. Additionally, Nup145C•Sec13 and Nup85•Seh1 behave nearly identically during gel filtration, indicative of heterodimers when their large hydrodynamic radii are taken into account (Figures 2.4 & 2.5). AUC experiments confirmed that Nup85•Seh1 is a heterodimer in solution (Figure 2.3). Thus, the heterooctameric pole model (Hsia et al., 2007) is inconsistent with our results.

The structural similarity between Nup85 and Nup145C extends to at least three other proteins (Figures 2.12 & 2.13). First, the architectural nucleoporin Nic96 (Jeudy and Schwartz, 2007) shares a common structural core (Figure 2.13) but has a distinct N terminus (Figure 2.12). The shared cores mutually superimpose with an rmsd of 3.0-3.5 Å. Nic96 has a trunk module (α1 to α3 and α12 to α19), a crown module (α4 to α11), and an N-terminal coiled-coil extension (instead of the insertion blade of Nup145C and Nup85) that tethers it to the FG-containing Nsp1 complex (Grandi et al., 1995). Apart from the N-terminal differences the three proteins differ mainly in the relative orientation of the crown and trunk modules. Although a previous comparison of Nup145C to the COPII coat component Sec31 did not reveal a strong similarity (Hsia et al., 2007), comparison with Nup85, Nup145C and Nic96 shows that Sec31 has corresponding trunk (α1 to α3 and α12 to α18) and crown (α4 to α11) modules. Sec31 homodimerizes to create an "edge element" in the COPII coat by an internal domain swap between two crown modules (Fath et al., 2007). This domain swap results in a mixed crown module that is identical in topology to the unmixed crowns in Nup85, Nup145C, and Nic96 (Fath et al., 2007). Structural prediction using Phyre (Bennett-Lovsey et al., 2008) also places Nup84 in the group containing Nic96, Nup85, Nup145C, and Sec31. Similarity extends beyond the trunk and crown modules to a 'tail'

module that has been structurally characterized in the C-terminal domain of human Nup107 (homolog of yNup84) and in Nic96 (Boehmer et al., 2008; Jeudy and Schwartz, 2007) (Figure 2.13C). The last three helices in the tail module of Nup107 form the interaction site with Nup133 (Boehmer et al., 2008). In Nic96 this region is predicted to be a protein binding site as well (Jeudy and Schwartz, 2007). Because we find this characteristic tripartite structural element of crown, trunk, and tail in architectural proteins of the NPC and the COPII coat, we term it the ancestral coatomer element 1 (ACE1).
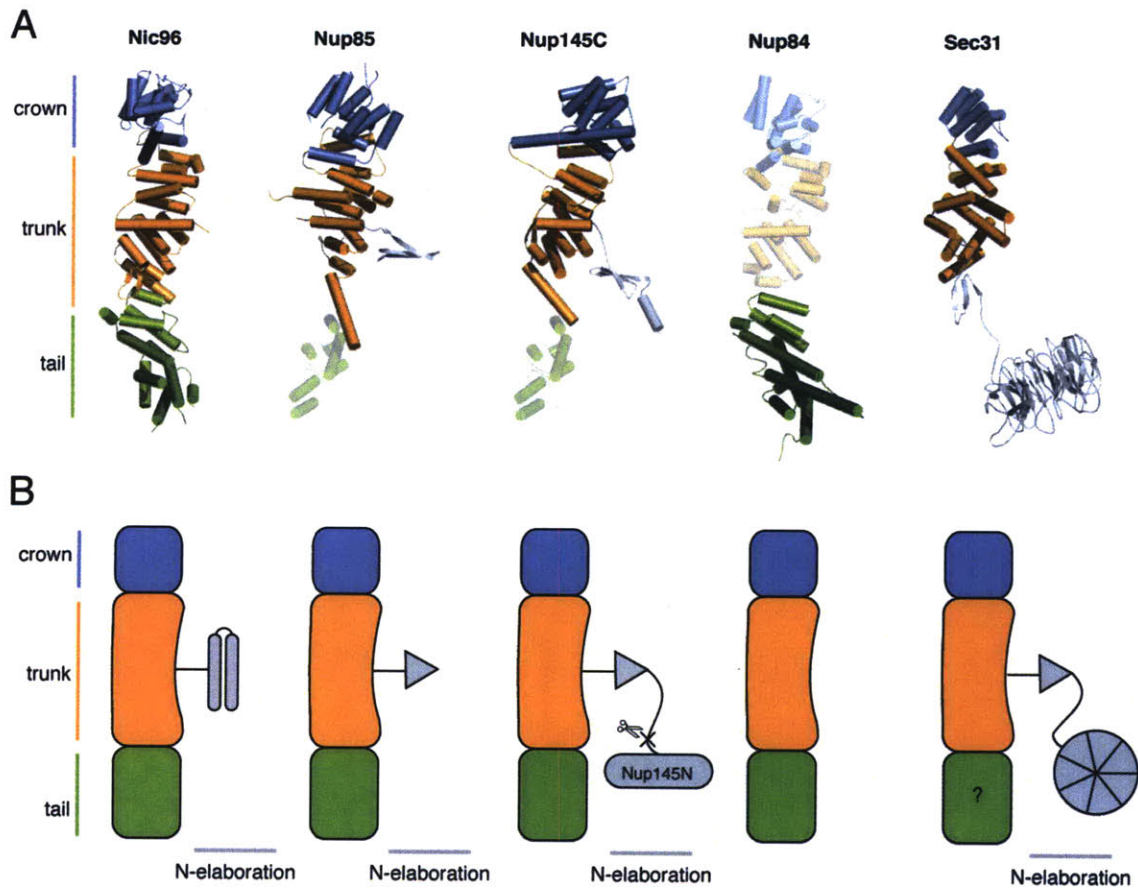
Figure 2.12 - Architecture of ACE1
(A) ACE1 containing proteins are shown as cylinders and sheets. Crowns are shown in blue, trunks in orange, tails in green, and other domains in gray. Modules with predicted structures are shown half-transparent. (PDB accession codes 3bg1, Nup145C; 2qx5, Nic96; 3cqc, Nup107; 2pm6, Sec31) (B) Cartoon illustrating the similarity and modular nature of the ACE1 element. The N-terminal elaborations are for Nic96 a coiled-coil domain that interacts with the Nsp1 complex, for Nup85 the Seh1-interacting insertion blade, for Nup145C the Sec13-interacting insertion blade preceded by an autocatalytic cleavage domain and Nup145N, and for Sec31 the Sec13-interacting insertion blade is preceded by its own N-terminal 7-bladed β-propeller. Sec31 has a unique proline rich insertion C-terminal to its trunk module followed by a conserved region predicted to be α-helical. The structure of this region was difficult to predict with high confidence. Thus, it remains to be determined whether Sec31 has a tail module similar to those in Nic96/Nup107/Nup85/Nup145C.
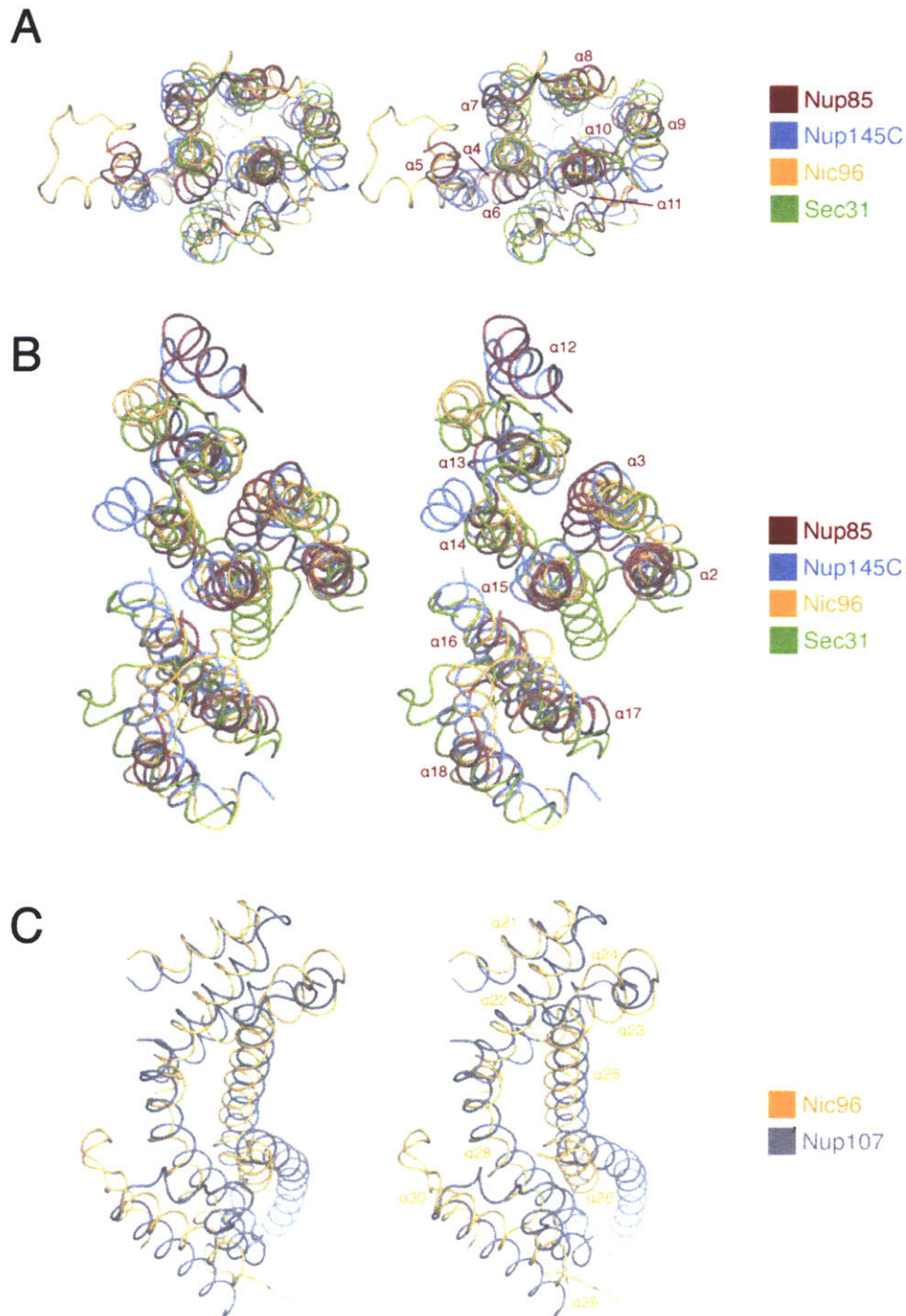
Figure 2.13 - Superposition ACE1 modules
Superposition of (A) crowns, (B) trunks, and (C) tails of ACE1 with known
structures (PDB accession codes 3BG1, Nup145C; 2QX5, Nic96; 3CQC,
Nup107; 2PM6, Sec31). Helices are labeled according to Nup85 in (A) and (B)
and Nic96 in (C). In (A) Helix α4 of Sec31 and the short β-sheet between helices
α5 and α6 in Nic96 are omitted for clarity. The rmsd between modules is 2.9-3.2Å
in (A), 3.4-3.5Å in (B) and 2.7Å in (C).

Can we predict ACE1 functional sites from established interactions? Analogy to Sec31 monomers in the COPII edge element suggests that Nup145C and Nup84 might interact crown to crown. Based on the Phyre-model and structural alignment, we constructed a surface point mutant replacing two conserved hydrophobic residues on helix α8 of the Nup84 crown with aspartate (Nup84-ISICM to Nup84-DSICD) (Figures 2.8 & 2.14). Nup84-DSICD disrupts Nup145C binding in a manner analogous to Nup145C-ELIEA, severing Nup84 binding as shown by gel filtration and ITC (Figure 2.5C,D) Thus, the Nup84•Nup145C interface is a crown-crown interaction involving α8 helices as in Sec31 homodimerization. Additionally, we found that the tail modules of Nup145C and Nup85 are necessary for interaction with Nup120 in a manner analogous to the human Nup107 interaction site for human Nup133 (Boehmer et al., 2008) (Figures 2.6 & 2.7).
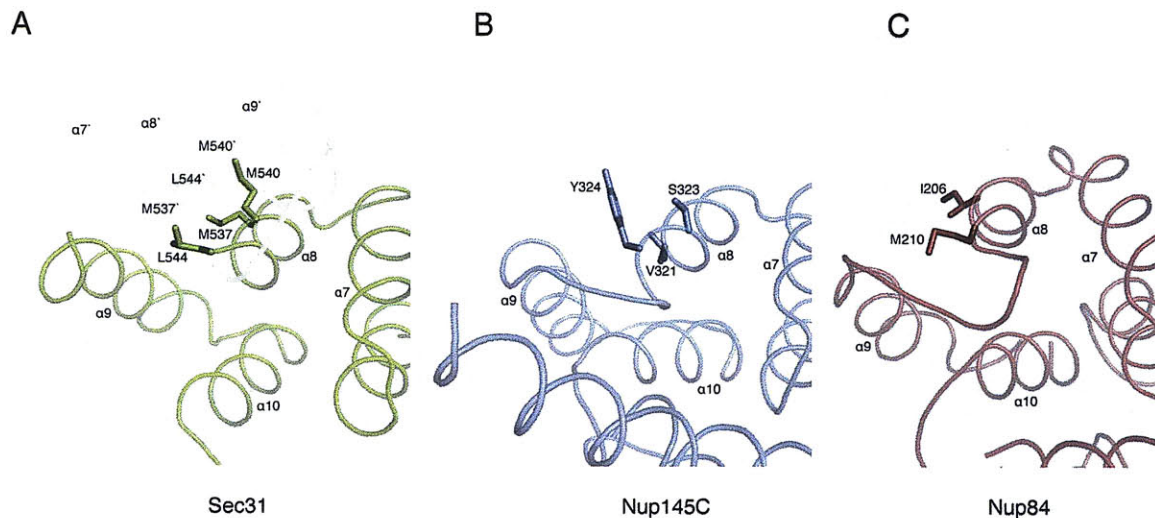
Figure 2.14 - Surface mutations of the Nup145C and Nup84 crowns
The crown helix α8 is highlighted in (A) Sec31 (PDB code 2PM6) (B) Nup145C (PDB code 3BG1) and (C) Nup84. In (A), one Sec31 monomer is shown half transparent and the three hydrophobic residues buried in the interaction between crown helix α8 in each monomer are shown as sticks. In (B), the three exposed α8 residues in Nup145C mutated to disrupt binding to Nup84 are shown (VLISY to ELIEA). In (C), the two exposed hydrophobic α8 residues in the Phyre-predicted Nup84 structure mutated to disrupt Nup145C binding are shown (ISICM to DSICD).

Here we have shown that ACE1 is abundant in the two main scaffolding subcomplexes of the NPC. To date, Nic96 is the only ACE1 protein in which all three modules (crown, trunk and tail) are structurally defined. In Nic96, the three modules form a continuous, rigid hydrophobic core (Jeudy and Schwartz, 2007; Schrader et al., 2008b). In the other four experimental structures, only a subset of the modules are present. We speculate that the three modules within ACE1 can allow hinge movements, utilized to different extents in specific family members.

Discussion

ACE1 is different from regular α-helical repeat structures, including HEAT-repeats and TPR-repeats (as discussed in (Jeudy and Schwartz, 2007)). The α-helical modules that compose ACE1 are distinctly irregular, most notably with elements that fold back onto themselves forming a U-turn within the crown module. The trunk is composed of two zigzagging helical units running in opposite directions. We propose that this architecture confers rigidity to the trunk and thus distinguishes it from regular helical repeat structures that are often inherently flexible (Conti et al., 2006). As a consequence of the specific arrangement of the helices in ACE1, several helices in the trunk and crown are encased by neighboring helices and thus have a characteristic hydrophobic character (typically helices α6 and α10, Figure 2.15). This pattern of hydrophobic helices may help to find additional ACE1 proteins. Several sequence elements, notably in the crown and at the predicted hinge regions, distinguish ACE1 from other α-helical domains (Jeudy and Schwartz, 2007). Nevertheless, these characteristics are subtle enough to remain undetected in typical primary sequence (i.e. BLAST) searches and candidate proteins need to be examined using all available tools, including phylogeny and secondary and tertiary structure analysis.

A Nup85

S.cerevisiae   MTID-----DSNRLLMDVDQFDFLDDGTAQLSNNKTDEEEQLYKRDPVSGAILVPM   51
C.glabrata   ----MLN-----DLVMDVD-LDFAD-GQSPMRTKKA----NEIIRDPISGSVVIPI   41
K.lactis   MANDEFA--DTQDLLMDVDNLDFLD-EEG-ISNESDDDINLTINVDPVSNAPMVNF   52
A.gossypii   -MSHMAN----EMLLDIDGMDFVE-DED-MGAEGE----LQFSFDPVSNAPMVSF   44
D.hansenii   MGDQEPS--TLSQKFDDIDMLEIPD-D----SASS---------------ETESFS   34
C.albicans   --MPEYT--NFESKFDDIDMLEIPE-DDD-----G----ENSIDNSISSETEGIS   41
Y.lipolytica   MFSAPPATINGADNYISPDPLTAREFGEAEMETILDGDDT----ELPEITLPGAPI   52

                                        β1 ───────▶   β2 ───────▶   β3 ──▶
S.cerevisiae   TVNDQPIEKNGDKMPL-KFKLGPLSYQNMAF----------ITAKDKYKLYPVRIPR   97
C.glabrata   DADKFPKESHN--SPL-VFKLGNGANENIMH----------VKSQKDVNLYPVPLST   85
K.lactis   -----PIEQPS-AKEL-RFKYNNVSSRSLAF----------DNSTKDNKLYPVRFLH   92
A.gossypii   PSTQQP---DL--EKL-RFKFAPVLSRSFAF----------NGGSRTGTLYGVQVPY   85
D.hansenii   DSSAS--------EPLKDKPSKQDYIRLQDWLKS---DDVIEFQFDSGQYKDSLTQ   79
C.albicans   DSSDSSKSSYDQWSPVAIFKSDKEYIDLSDWLNT---QKSLEFKFDVQKYRKGIEK   94
Y.lipolytica   PEGQVEEWQKS--RNM-GFVMDPVIARGIAWFDAKPGQDGHATKEDRTLYPTTTRV   105

                                           α1 ─────────────────────────
S.cerevisiae   L-D-TSKEFSAYVSGLFEIYRDLGDDRVF------NVPTIGVV------NSNFAK   138
C.glabrata   S-D-FSENFINYICNLFDIYQDLGSHRIF----SRPTIGVI------SSDYEL   126
K.lactis   L-D-ESKEFAQYVSKLFELYQGLGEHRKN-----DVPTIGLI------KQTSRL   133
A.gossypii   V-D-NSKEFSQYAAKLFEVYKSLGADKQF-----AVPTIGLI------NHTSTL   126
D.hansenii   N-D-VSKSYTKYINSLFKIIEGFTNDDITRIDLEDDEDPIGLISTSAKFGNSARKA   133
C.albicans   SRP-FDEKYILYVNNMFKIIQKLTEDDITRFDLEEDDSPIGLVMDS--MGSKAKQV   147
Y.lipolytica   G-DEWSDNYKGFVAEAFATTLG--PQ----------------------------   128

                                 α2 ──────────────────   α3 ────────────────
S.cerevisiae   EHNATVNLAMEAILNELEVFIGRVKDQ-----DGRVNRFYELEESLTVLNCLRTM   188
C.glabrata   EHTRVVNMALDSMIIELEILIASYDKLE----NANIARILELEQCLVILQCLRTF   177
K.lactis   EHFSIVNLAFHALVTELEFYIESIKY-----TNKLQRIGDLEECLSILNCLKTI   182
A.gossypii   EHNQTVNLALEAIVSELELFIESLKY-------SGRQQRLVDLEECLSILNCLKTV   175
D.hansenii   QRMHKIDEAFTKIVENLTQYIHDIESIG-----VDDETTEQFYYLLSILDCLHAN   183
C.albicans   RKFKKIDEALREIIDYLQDLVNSIPKEE-----QESYDSQCLQHVLYILECFEAN   197
Y.lipolytica   --------AAEDLTQRFADFIESIKHYREEIRKDVDIVEAEPLDDAYSIALCIYAV   176

                                          α4 ───────────────   α5 ──────────
S.cerevisiae   YFILDGQD-----VEENRSEFIESLLNWINRSDGEPDEEYIEQVFSVKDSTAGKKV   239
C.glabrata   QFIDS---------INDRATFFDSLLKWTNRTDGEPKVEYIQSIFGQS---DSSQV   221
K.lactis   YFLTDSPE------Y-KQEDLLESLINWVNRSDGEPSELVIQKIFDETLL-TRRKV   230
A.gossypii   QFTLDSEE------ENSRAKFIDSLINWVNRTDGEPSEAVIAKVLGDG---KQTPI   222
D.hansenii   YFCSDTRM------K-----PESIAKWINRFDYKPDKELVESVMV------NSPKP   222
C.albicans   NFYFDIQQ---------KP---ELIIKWVNTFDPKPDPELLNDVIV------NTPQP   236
Y.lipolytica   YFHNGAAGQSGFFGSQSTYEVRQPLVEWVNVSEQQPSVELGKEVMSQT---P---P   226

                         α6 ──────────   α7 ──────────   α8 ─────────────
S.cerevisiae   FETQYFWK-LLNQLVLRGLLSQAIGCIERSDLLPYLSDTCAVSFDAVSDSIELLKQ   294
C.glabrata   FFTEPFWK-LVYQLLLRGLTEQAINTLEKSELAKYLKENCETTSTIFEDFIQLVKN   276
K.lactis   YEQSDFWD-LTAQLLLRGLWSQSIQCIQNSQLLEHGNE---TVGVFVTDLITIIES   282
A.gossypii   LENPYFWR-LLCQLIIRGLFDQAVAAIEKSELLSYLADKCSATHTMVQDMVNLLQG   277
D.hansenii   YTHPQFWNTYLSQLITRGLLTQAIAAIEKSQY-EELEENAPELFSVIQDFSMLLKN   277
C.albicans   YLHPRFWNTCISQLTRGLLFTQIHEVIEHSQY-QELKEKCPELYAVIGOMDTLLSN   291
Y.lipolytica   VQHKDFWK-YIYQLVCRGMLKQASHCLHNSGA-GEVDPSCQETLEQTIDILSRYPQ   280

                         α9 ─────────────────────   α10 ──────────────
S.cerevisiae   YPKDS-SSTFREWKNLVLKLSQAFGSSATDISGE-----LRDYIEDFLLVIGGNQR   344
C.glabrata   YPLE-SEEHFREWKSFALELQQNFEDADTKIPVP-----LRKNLLDAISITAGNRD   326
K.lactis   YPLHS-ENLFREWKNSVLQLMTNWDEQE--VEPE-----IKKNLMNIFMILSGSKN   330
A.gossypii   YPRE-LEPVYREWKDLVLHLNWSQSEHKISIE-----LASSLEDCLLIMCGNKS   327
D.hansenii   YTSMSMKHQFHEWKLSCCEFRDLFLKFKANIT-DSKDLIILNQIYDLLCILTGLPK   332
C.albicans   YTSYSSKGQFTAWKLLACEFRDSLSSVRNESITETKHKLIIDQIYDLACIFTGLPK   347
Y.lipolytica   GPGN--VSFYFRQWHNSISTVQSKLLKIN---DAK-----IQKGLTTLLQVLAGDED   327

                         α11 ─────────   α12 ────────────
S.cerevisiae   KILQYSRTWYESFCGFLLYYI-PSLELSAEYLQMSLEANV-VDITND--------W   390
C.glabrata   KILQHSSFWYESLTGFLLYYI-PSEELIQEYVGLAVSKTP-IDVTNP--------W   372
K.lactis   KICEFSQYWYESYCGLMLYYI-PTLELSQEYGQLATKHNA-IDVCNN--------W   376
A.gossypii   KIIYYSKTWYECYCGLMLYYI-PSLQLSEEYLQLVLKEHP-LDVTSP--------W   373
D.hansenii   TIAIHCDKWYEIYTALSLYQVRDDDKLFKDYFNVAISEKPPALI--EDSDVLSIIS   386
C.albicans   TISSYCDTWYEVVYLALSLYQVRDSNEVYIDYFKTAVSEKPPSDIFDDENENLDGLT   403
Y.lipolytica   AIMGTNHTWYDALVTQQQYVD-PCDSRMKGYDAAVANYP-VDTFTV--------W   373

Protein multiple sequence alignment with secondary structure annotation (α13–α30).

```
                α13              α14                  α15
S.cerevisiae  EQPCVDIISGKIHSILPVMESLDSCTAAFTAMICEAKGLIENIFEGEKNSDDYSNE  445
C.glabrata    ETACLKILQNDIFPVLPILESLDNCTASFSAALCESKGLLIDRYT---ILEENSGS  425
K.lactis      ETPCYHIINDKVADVLPILESLDLATAAFVAVLCEAKGIIKTDFY-----EIFAHK  427
A.gossypii    EQACVSIITGKIYSILPVLDSLDSCTAAFTAAICEAKGLLENDLL-----DDGGLE  421
D.hansenii    EQCFVNILEEKFLKMLVSVDTIDSATAAYVARLLELKGLLNNYYS---AIN-----  434
C.albicans    EACFLNILENNFLKVLETLHELDPPTTAYLAELMELRFLLRGYYF-----D-VTVT  453
Y.lipolytica  EAGASHAIKGSLLLAIETLENCDSVIAALVSNMCYDAGLLNGY----------GAA  419

                          α16                      α17
S.cerevisiae  DNEMLEDLFSYRNGMASYMLNSFAFELCSLGDKELWPVAIGL----IALSATGTRS  498
C.glabrata    DDIEIDDIFSQRNGMASFLLNNFALELCSMKMKTLWPVAIGL----ISLSPYNTPS  477
K.lactis      SGLNSELAFENEPTIGEFLIRQVALSLCSYDDKTLWPVSVGL----ISLSSDSGDS  479
A.gossypii    VSNTGDEMFGLGNTMASYLLNQFALSIVTYEDKNLWAVAIGI----IYLTPNYSDA  476
D.hansenii    -SKDLQELIN-RKTISEYLLTTHAYQC--LNSHYLVPVGIGLLSNNDISSSSQSSV  486
C.albicans    SSASFEDLLN-RRIVSEYFLTRHAYDCL--NIHELVPVGIGLLLNNVVCTSNNAVS  506
Y.lipolytica  ----------NPAALPDYHLTNLALQCIG--YDNL--INVGI----EVLQKLRDVA  457

                α18              α19                  α20
S.cerevisiae  AKKMVIAELLPHYPFVTNDDIEWMLSICVEWRLPEIAKEIYTTLGNQMLSAHNIIE  554
C.glabrata    SKKSTIAELIPHYPFETNDDIEWALSVCAKWRLPDIAKEIYTTLGNQMLDNGNTVE  533
K.lactis      SKRAILAELLPHFPYTTNDDIEWLTVCAKWRLPDVAKEIYKLLGQDTLYQNNIVE   535
A.gossypii    AKRMMIAELLPHYPFQTNDDIEWMLTICAKWKLPQIARTIYRVLGQEALYLNNIIE  532
D.hansenii    NNRNTISEFLPHYECKTNDDLEWTLTICAKLNLIHTARKLYQVYGNKSLKDGYLYE  542
C.albicans    ANRKVIATFLPNYICNTNDDLEWALTICANLNLVSTARQLYYQAGLKSLEEGYMYE  562
Y.lipolytica  PARQILSLLLPRLPYDSAREIDWAIKKCKQNGLVEVEAEILQTLAQKAIHNNNNLM  513

                α21                      α22                  α23
S.cerevisiae  SIANFSR-----------AGKYELVKSYSW-LLFEASCMEGQKLDDP-VLNAIVS   596
C.glabrata    AIANFSK-----------AGQYELVKHYSW-LLFENAAFQGKPMEDE-VLSAIVD   575
K.lactis      AMSNFSK-----------AGEYEWVKHYSW-MIFEASILQGGPLDDI-TINSIVN   577
A.gossypii    AISNFSK-----------AGEFEWVKHYSW-MIFEASILQGAPLEDE-VVNSIVT   571
D.hansenii    SLNMFVNCYDPDSMTNESN-EGMKQVHHIIWDIVFQDSLVNNRPIKDE-LINNIIC  596
C.albicans    ALNNFVNCYDPTLLSGENHSEGMKKVHYIVWDLIFANCLVHNRPIKDE-LINNIVD  617
Y.lipolytica  GLCMFAK-------------NGNIAALRYHCW-KLFKNALVAESPVEGDKSLADYVK 556

                          α24                  α25                  α26
S.cerevisiae  KNSPAEDDVIIPQDILDCVVTNSMRQTLAPYAVLSQFYELRDRE--DWGQA---LR  647
C.glabrata    GDFNTSE--LIPDEVLNKVVTNAMRQSLSPYAVLSQFFSNVENE--RWSKA---LQ  624
K.lactis      GDTDSFN---IPKDLLNSMVTDVMRQTLSPYAVLYEFHKLAENN--ETEKA---MN  625
A.gossypii    GEYESA----IPKELVEVMVTPAMKQSLSPYAVLFNFFRATEDG--RWAEA---LK  621
D.hansenii    HKVDT----------GFEIHPVIRQCLSPYAVLVEFFKSLEQS--RSTTAMDNLS   639
C.albicans    KKVD-----------LNFKIHPVIQQCLAPYAVLKEFYETLPQRDIKLSNK---LS  659
Y.lipolytica  DPAS------------VSDVPAEVAECIAPYAVLTEYFNLRRSG--QLVKS---SE  595

                          α27                      α28
S.cerevisiae  LLLLLIEFPYLPKHYLVLLVAKFLYPIFLLDDKKLMDEDSVATVIEVIETKWDDAD  703
C.glabrata    QLMSLIEFEYLPKKYTVLLIVKYMYPLFLNDDQRMIDEETLVSLLAAIEKNWDCDD  680
K.lactis      HMIQLISFPYMPEQYCILLLAKFFYPMFLKTTETALDESQIFTILKTLEKI--KLN  679
A.gossypii    LLLSLIEFTYLPPQYLILLIGRFLYPLFLQDESKMFKEQDVLRVIKALEQFHDRPD  677
D.hansenii    KLVHLIRFNHLPKKFYPLLLCQII-P-FLIDSGMKFQLPDLIVIIELIDTFEFQAN  593
C.albicans    KLIHLLKFNYLPKKFVPLLLAQFI-P-FLSEPNTTFQLPDLVIIIELLDNYESDRN  713
Y.lipolytica  YLLALFEFPLLPTEYLGLLIHQLINYVSPISQQFNFSTSQFMTLLSAFDKWDKSQG  651

                α29                                      α30
S.cerevisiae  EKSSNLYETI-----------------I--EADKSLPSSMATLLKNLRKKLNFK   738
C.glabrata    TDSVIAYEKL-----------------SEE--IADLPKELKSMEKLIRKKLNGK   715
K.lactis      DSCVALYSQLRNDD-----------------NST--CSNLPESVDSFIQDVRKTLNYK 718
A.gossypii    EITTAMYKEL-----------------LVK--VDTLPKDPLNLINAIRTKLSLK   712
D.hansenii    DDEYTEGESLYKYSISNIENTTESYDWRNILQKSGNTLPKDLKSLIKLLRNEVVAK  749
C.albicans    EKELKEGADLYLYSINNIESDADENDWRKLVG--KDNLPKDIDSFTRLLRNLIAAK  767
Y.lipolytica  TPAFEQGEELIK---DMVDHSSDQNDWRKTFK-----IHELVMSELRLARTSVATM  699

S.cerevisiae  LCQAFM---  744
C.glabrata    LCLNLM--I  722
K.lactis      ICQSFM---  724
A.gossypii    LCHQFM---  718
D.hansenii    IGKVYVDHM  758
C.albicans    VGKVFI---  773
Y.lipolytica  SSLKGLGGR  708
```

B Nup145C

α1 · β1 · β2

```
S.cerevisiae   - D L L F S E - - C N D E I D N A K L I M K E R R F T A S Y T - - - - - - F A K F S T G S M L L T K D - - I V   166
C.glabrata     - D L L L S N - - F N D A K N K T R E I E Q E R R L N T K K Y - - - - - - N P C S F G N G H L V I G D N R S R   149
K.lactis       - D I L F P S - - F K K D L L T Y Q S V K R E R R I N S H M W - - - - - - F A K F N I N G Q L L L K D - - P S   192
A.gossypii     - D V I F N T - - F Q K D M D E F K S I R R Q R R L D S A P S - - - - - - F V R F N N D S T I T M K T - - D K   195
D.hansenii     - D I L F P E - - F N R D S L S M N H V S T P - - I N S - - - - - - - - - - - - - - - - - - - - - - K K - - F S   259
C.albicans     - D I L F S D - - F N K N V L K V S T P T K K K - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   228
Y.lipolytica   D N L L F T N T L V N V A E E E Y N K A A E K - - L R L P S Y S A A V E S S L A I F Q P S G Q L A V K L - - P H   302
```

β3 · α2 · β4 · β5

```
S.cerevisiae   G K S G V - - S I K R L P T E L Q R K F L F - D D V Y L D K E I E K V T I E A R K - S N P Y P Q I S E S S - L L   217
C.glabrata     K I L - - I T N Q V L V K D E L P - A V L T D - R T I F D N E L K T T L I - H S R S S N N Y P C V F K K E - L I   199
K.lactis       Q L S G - - C K Q I V F Q S N L P - I N K S S F D N V F A S Y L N T S I V - N S R - P N G Y P L V G K C S - L Q   242
A.gossypii     S T S G C T V T A S P L P L Q T Q - R S S I D - - S V L K K S L I D S T I - E L R - N N N Y P I V K Q F S - L T   245
D.hansenii     N L E G S - - - S L E M Y E E V Y - P R N V S - - K V I Y H M L S R S M I - S T R - S N K F P I V E S N S N F S   307
C.albicans     - - - L S V V D D I D E S Q D I Y - I D N I S - - T I F H K L L S K I V I - G H R - G N E F P K I D K T T G F E   276
Y.lipolytica   S D L - - T I G T V S K L C N F S - E N A - Q D D N F L G S V K E D M T A - A D - R N N G Y P R Y S - - - - - -   346
```

α3 · α4

```
S.cerevisiae   F K D A L D Y M E K T S S D Y N L W K L S S I L F D P V S Y P Y K T D - - N D Q - - - - V K M A L L K K E R H   266
C.glabrata     F K D V L Q C I L P T S E S Y K L W N F S S I L F D P I Q I D C P M E - - - - - - - - A K D A T L K Y K R R   245
K.lactis       F D N L A S A Y S D V P N E D K I L K L A S I L F D P L S L P Y A V N S P E - - - - - - V E K V L I K K Q R H   291
A.gossypii     F D D I A A A Y K P I P A E Y R I W K L A S I L F D P I S V S K S R S H P D D A - - - - - V K D V L V K K K Q Y   296
D.hansenii     F A D I S L N E N S T D E E E Q I L K L G S A L F D E H K L N E Y D E Y K D V N I S D S H L V K Y L E N L Q Q K   363
C.albicans     F K D I I T S H Q E R E E - K D V T I L C S A L F D N L T I N E P - - - - N P A - - - - - I S A A L V D N T R K   322
Y.lipolytica   - - - L E S V P A H A E L K G T V Y S L A S I L F Q S S T A L G L S E Y I G S A P V D G L Q A R E L E A S L R K   399
```

α5 · α6 · α7 · α8

```
S.cerevisiae   C R L T S W I V S Q I G P E I E E K I R - N S S N E I E Q I F L Y L L L N D V V R A S K L A I E S K N G H L S V   321
C.glabrata     D N I C N W I K D H V R D N V Y Q K I - K V S T T P L E K I F L H L L I N E I E T A A K I A I D S D N S H L S I   300
K.lactis       A K L C G W I V D E T R A E I D S M L S - - T A S D I Q K I L L F L S V N D I V N A S K T A I I S K N K H L S V   345
A.gossypii     E L L C D W I I N E I N S E V G A K I A - - S A G P L E K I F L Y K R D I I G A T T A A I A S N N N H L A V   350
D.hansenii     K N F T E W L K V Y N G S T I E Q L I E K N K S D M L E Y I F I K V C G G Y L K D A I N L A M D S N N A H L S V   419
C.albicans     K L L G D W L K N Y N S A T V E K L L A E Y K N D P L E T T F I Y M C S G D M I K A I E T A I Q T N N S H L S V   378
Y.lipolytica   Q R L S Q W C E K A V F E Q V Q S E L - - - G D N S V E N I V T L L T G N R V D E A C A A C L K T N N V H L G V   452
```

α9 · α10

```
S.cerevisiae   L I S Y L G S N D P R I R D L A E L Q L Q K W S T G G - - C S I D K N I S I Y K L L S G S P F E G L F S L K E   375
C.glabrata     L I S M L G S N D P R L K A L A H T Q I E Q W S G I G - - S N V E I Y V A K I Y K L L S G E L F K G P F S L I E   354
K.lactis       L I T L L G S N D P I V R E I A Q L Q L T K W K S L G - - S I V D P T V I S I Y Q L L T G N P F A S T A L - V N   398
A.gossypii     L L T L L G S N D P L V R E L S T S H L S K I K K L G - - S S L D I N I I K I Y Q L L T G S P F A E S A N - S V   403
D.hansenii     I L T L I D S N D D A V K S T A V N Q L Q Y W S D T S S L S I F I V K I H K I L S G D - - - - - - - - F S E   468
C.albicans     V I T L S D S N D V V V K S I A Q N Q L T N W K Q R Q T I S S I P S A V V K V Y Q I L A G D - F Q P - - - - - -   427
Y.lipolytica   V I S M M G S S - A S A Q S T A Q A Q L D H W Q S T N A M D L I I P D Y T Q Q V Y Q L A A G N I - - - - - D V V C   502
```

α11 · α12 · α13 · α14

```
S.cerevisiae   L E S E F S W L C L L N L T L C Y G Q I D E Y S L E S L V Q S H L D - - - K F S L P Y D - - - D P I G V I F Q L   425
C.glabrata     N A E D V N W L V L V G A A L H Y G E V D E L S L E D L I G N S M A C - L - - S D H C S - - - G V F Y L L L K L   404
K.lactis       E S N K F S R L V N L G L Q V Y Y G D I D S L T L E D L I M N A I S N P F V D T S S V S T F E N L S M N I M K L   454
A.gossypii     I S E G L S W L A T L G L Q I F Y G D I D A L S L R E L I E R G L E Y S C K D Q W P L N - - - D I S A N I L R L   456
D.hansenii     V L S G L P W N I S L A I K L F Y G D - N T L K L H E L I Q E F - Q D G I - - - - V E S - - - G P I Y D I L T L   515
C.albicans     I L E T L P W N L G L A L K L F Y G N Y N D I - - K K L I N E F - - - - - S S S I P I G - - - N P V G D V L H A   473
Y.lipolytica   K A R K L D W L R E F G L R L W F G R N R D I S A S - - I A Q - I Q S S - - - - - - S S - - - D W D L R L L R L   546
```

α15 · α16 · α17

```
S.cerevisiae   Y A - A N E T E K L Y K E V R Q R T N A L D V Q F C W Y L I Q T L R F N G T R V F S K E T S D E A T F A F A A   480
C.glabrata     F C - S T R N A D Q V L Q E I L K E S N A L G I D F W H C L Q A L E S N G I A D I N N P F C D Q I T V Q Y A Q   459
K.lactis       F A C P T I P V E Q L L D E L R S H S E I F D V R L C W F F T H M L Q R E - - - D I T E H L R D R I T L E F I D   507
A.gossypii     Y C - S D V T P D I L V G N L K V L S W F F I Q I L T R D - - - D I S P S L R D H L T L Q Y V E   508
D.hansenii     Y N - Q I H T K D K N Q A L Q L I K S S H L N I K L K W F F N K V L S - R G D A S F - E I L S T D L S L S F G N   568
C.albicans     Y V - N G I D L E - - - - - - S V T S S S L N I K L K W L F C K V L A - - - - - D F N - - - Y D T I T K E F G D   514
Y.lipolytica   Y A - G K Y N L E Q T L - - - - - T G K A F D T A V P W V L C A F F - L Q G H A K E A P D T A D R L T L Q Y A A   595
```

α18 · α19

```
S.cerevisiae   Q L E F A Q L H G H S L F V S C F L N D D K A A E D T I K R L V M R E I T L L R A S T - - - N D - H I L N R L K   532
C.glabrata     N L E L S G H I P E A L Y I C A H I K D D H L A K K L F E S I I F S N I H H L T L D - - - G K P L A I M S K L M   512
K.lactis       Q L K L D R M H K E A L F V A C F I A D D A I A K N T I D L L L S S E I L Y F T S N - - - D I K - P I L E R L Q   559
A.gossypii     Q L K L N R M F G E A L F I M C F I N D D R L A K Q Q V D H L L S S Q I T F F S Q D - - - S N Y - E L L T R L R   560
D.hansenii     F L E K I G L W K E S L F V Y S H I S D D K E N E R V I R N L V I S N I D I K S S P - D E E T - Y I T K V L K   622
C.albicans     Y L S S I D Y W K E S T V V F A H L T N D N D T G D A I T K L L I N S K I S H I K S L T I D K E Q - Y A I E V L K   569
Y.lipolytica   Q L E S M G Q V E N A L L V L G F L V S D S A Y A Q A I T K L V A R N I V A L R K L D - - - - - - - - L R S Y R   643
```

α20   α21

S.cerevisiae	IPSQLIFNAQALKDRYEGNYLSEVQNLLLGSSYDLAEMAIVTSLGPRLLLSNNPVQ	588
C.glabrata	VPNEVIYRSMAQYAKYRKCHQEELEYLLKSKDTRLAKEVFITKVAPSYILGNDD--	566
K.lactis	ISAPIIHRHLALYEKYSGDHLSEVSNLLKAGHFQEAELVTITTVGPKLIINAKW-N	614
A.gossypii	IPKSSYYAFLALLDKYNRNHLSEARNLLKAGHFQEAEKVVIVSVAPKLVLDGSA-	614
D.hansenii	VPQSLIYEAVAIQEHSLGNYWEECEALVTAKLWKKAHECIIKELGPLTVISNND--	676
C.albicans	IPRMVIYKAVAIQKSQNGDFWGECEALIEVSLWEKAHITIVNELGPKTVISNSQ--	623
Y.lipolytica	IHPQLLSESEALLARYENRPVDEVRYLLDAELWGAANTTVLNDVAPQAVIKGDA--	697

α22   α23   α24

S.cerevisiae	NNELKTLREILNEFPDSER--DKWSVSINVFEVYLKLVLDNVETQ------ETIDS	636
C.glabrata	-SKLLSLTSMLEQFDRNS----TQRNDLKVYDYYIQFK-KNSENS------NIIKS	610
K.lactis	NDYLSVLETLLQRFPSH--TIPTWEKGLGVYEKYIKLSLHNSLDP------NIIHQ	662
A.gossypii	-ANLQTLRQLLETFPAQ--QMETWTHGLGVFEKYLQIALDNNHNQ------ELLSD	661
D.hansenii	-ESKNRLQSLIAKFPESGHIIPLWSQGAGIYDNFVSLSQEEIQEKASLDVHTLLVS	731
C.albicans	-NEKSQLQNVLFKFPENGLIISDWNKGAGIYGKYL-IVLQNESDL------SAIKF	671
Y.lipolytica	-QS---LEDVIMQFPSPERHIPTWKEGGKVYLDYARFKLSKAVDV------LELAN	743

α25

S.cerevisiae	LISGMKIFYDQYKHCREVAACCNVMSQEIVSKILEKNNP--SIGDSKAKLLE-LPL	689
C.glabrata	LASELPQFYSKHQMFENVTACCNIISNNVLSAVLDLDND-SKTKEFSNLVCN-MPL	669
K.lactis	LINTLPNLCKDYSHHELVNVASSRISEAVSELFLINLNM-LSRKIPKESLLS-LPL	716
A.gossypii	LVRVLPVLATDFGSHRELSVVCCVMSKLVCHIILENYRQALETPSFKDRLLA-LPL	716
D.hansenii	LLSNLPLLQ-DYNTSK-CRIASKLMSKKVGDIALNYAD-QI--GNIKGLILS-LPL	781
C.albicans	LLDNLPLTNIDSFH---KTVALDIISKFIGNLIIENDQF---NPTDRFKILN-LPL	720
Y.lipolytica	GVSHMSKALEEKSQFE-TRVALHVMSDDIQRSLLLAPTTDD--GTKALEVLNEVTL	796

α26

S.cerevisiae	GQPEKAYLRG-EFAQDLMKCTYKI	712
C.glabrata	GQPERKYATL-HFCSR-------T	680
K.lactis	GQPESNYLKK-VLSS--------L	731
A.gossypii	GQPETIYLKR-ALAS--------T	731
D.hansenii	GEVERNYFNI-RLQLV-----N-I	798
C.albicans	GDVNKRFFEL-RLAES-------K	736
Y.lipolytica	SMKQCVYSAGTRFKAS-----F-S	814

C

Nup85   MTIDDSNRLLMDVDQFDFLDDGTAQLSNNKTDEEEQLYKRDPVSGAILVPMTVNDQ 56
Nup145C ------------------DLLFSECNDEIDNAKLIMKERRFTASYTF------- 151
α1

β1   β2   β3   α1
Nup85   PIEKNGDKMPLKFKLGPLSYQNMAFITAKDKYKLYPVRIPRLDTSKEFS--AYVSG 110
Nup145C ----------AKFSTG--SMLLTKDIV--GKSGVSIKRLPTELQRKFLFDDVYLDK 192
β1   β2   β3   α2

α2
Nup85   LFEIYRDLGDDRVFNVPTIGVVNSNFA---KEHNATVNLAMEAILNELEVFIGRLK 163
Nup145C EIEKVTIEA-----------------RKSNPYPQISESSLL--FKDAL-DYME--- 226
β4   β5   α3

α3   α4
Nup85   DQDGRVNRFYELEESLTVLNCLRTMYFILNGQDVEEN-------RSEFIESLLNWI 212
Nup145C -------KTSSDYNLWKLSSILFD---PVSYPYKTDNDQVKMALLKKERHCRLTSWI 273
α4   α5

α5   α6   α7
Nup85   NRSDGEP--DEEYIEQVFSVKDSTAGKKVFETQYFWKLVNQ-LVLRGLLSQAIGCI 265
Nup145C V-SQIGPEIEEKIRNSS----------------NEIEQIFLYLLLNDVVRASKLAI 312
α6   α7

α8   α9
Nup85   ERSDLLPYLSDTCAVSFDAVSDSIELLKQYPKDSSSAFRE-WKNLVLKLSQAFGSS 320
Nup145C ESK-----------NGHLSVLISYIG---------SNDPRIRDLAELQLQKW--- 344
α8   α9

α10   α11
Nup85   ATDISGELRDY-IEDFLLVIG-GNQR---KILQY--SRTWYESFCGFLLYYI--PS 367
Nup145C --STGGCSIDKNISKIYKLLSGSPFEGLFSLKELESEFSWLCLLNLTLCYGQIDEY 397
α10   α11   α12

α12   α13   α14
Nup85   LE--LSAEYLQMSLEANVVDITNDWEQPCVDIISG-KIHS-------ILPVMESLD 413
Nup145C SLESLVQSHLDKFS----LPYDDPIGVIFQLYAANENTEKLYKEVRQRTNALD--- 447
α13   α14   α15

α15   α16
Nup85   SCTAAFTAMICEAKGLIENIFEGEKNSDDYSNEDNEMLEDLFSYRNGMASYMLNSF 469
Nup145C VQFCWYLIQTLRFNG------------------------TRVFSKETSDEATFAF 478
α16   α17

α17   α18   α19
Nup85   AFELCSLGDKELWPVAIGLIALSATGTRSAKKMVIAELLPHY--PFVT--NDDIEW 521
Nup145C AAQLEFAQ---LHGHSLFVSCF--LNDDKAAEDTIKRLVMREITLLRASTNDHILN 529
α18   α19

α20   α21   α22
Nup85   MISICVEWRLPEIAKEIYTTLGNQMLSAHNIIESLANFSRAGKYELVKSYSWLLFE 577
Nup145C RLKIPS-QLIFNAQALKDRYEGNYL--------------SEVQNLLLGSSYDLAEMA 571
α20   α21

70

```
                    α23                              α24                                   α25
Nup85   A S C M E G Q K L D D P V L N A I V S K N S P A E D D V I I P Q D I L D C V V - T N S M R Q T L A P Y A V L S Q   632
Nup145C I V T S L G P R L - - - - - - - L L S N N P V Q N N E L K T L R E I L N E F P D S E R D K W S V S I N V F E V Y   620
                                                             α22                                        α23

                    α26                              α27
Nup85   F Y E L - R D R E D W G Q A L R L L L L L L I E F P Y L P K H Y L V L L V A K F L Y P I F L L D D K K L M D E D S   687
Nup145C L K L V L D N V E T Q E T I D S L I S G M K I F Y D Q Y K H C R E V A A C C - - - - - - - - - - N V M S Q E I V   666
                         α24                                                                α25

         α28                          α29                              α30
Nup85   V A T V I E V I E T K W D D A D E K S S N L Y E T I I E A D K S - L P S S M A T L L K N L R K K L N F K L C Q A F M   744
Nup145C S K I L E K N N P S I G D S K A K - - L L E L P L G Q P E K A Y L R G E F A Q D L M K C T Y K I - - - - - - - - - -   712
                                                α26
```

Figure 2.15 - Sequence alignments
Multiple sequence alignments of (A) Nup85 and (B) Nup145C covering the phylogenetic spectrum of budding yeasts. The alignment for Nup145C begins after the N-terminal unstructured region at residue 123. Sequence alignments were performed with 3D-Coffee (O'Sullivan et al., 2004) (using the known structures) and illustrated with Jalview (Clamp et al., 2004). Sequence conservation is colored from white (not conserved) to orange (highly conserved). The secondary structure is shown above the sequences and was assigned using information from the known structures and predictions from PredictProtein (Rost et al., 2004). Dashed lines denote the C-terminal regions absent from the known structures. (C) Pairwise alignment of Nup85 and Nup145C from *S. cerevisiae*. The alignment was made by combining DALI (Holm and Sander, 1995) results using the structures with an alignment from T-Coffee (Notredame et al., 2000) corresponding to the tail modules. The secondary structure for Nup85 and Nup145C is shown above and below the sequences, respectively, and was assigned from the structures and predictions from PredictProtein.

Based on distance constraints and stoichiometric considerations, the heptameric Y complex has been placed in two concentric eight-membered rings on the nucleoplasmic and cytoplasmic faces of the NPC (Alber et al., 2007b). But how is it oriented, and how is it connected to the inner ring of the scaffold? Nup133 is anchored to the structural scaffold by its interaction with Nup84, positioning it at the periphery of the pore (Boehmer et al., 2008). Nup84 is the link between Nup133 and Nup145C. Thus, we position the extended arm of the Y composed of Nup145C•Sec13, Nup84, and Nup133 facing outward (Figure 2.16). Excluding the Nup133•Nup84 pair, the remaining pentamer forms a roughly symmetrical triskelion that conceptually resembles the vertex elements that form polygonal

cages in vesicle coats. EM analysis showed that the triskelion measures approximately 20 nm between the tips (Lutzmann et al., 2002). An eight-membered ring of the Y complex around the central transport channel has a ~50 nm diameter if the edges were to touch at the tips. Alternatively, the Y complexes might connect through a yet unidentified adaptor protein.
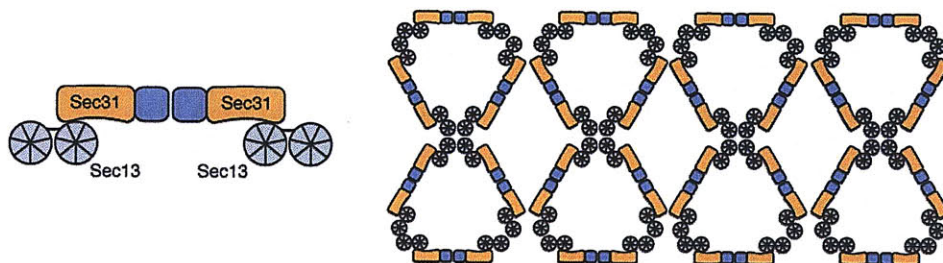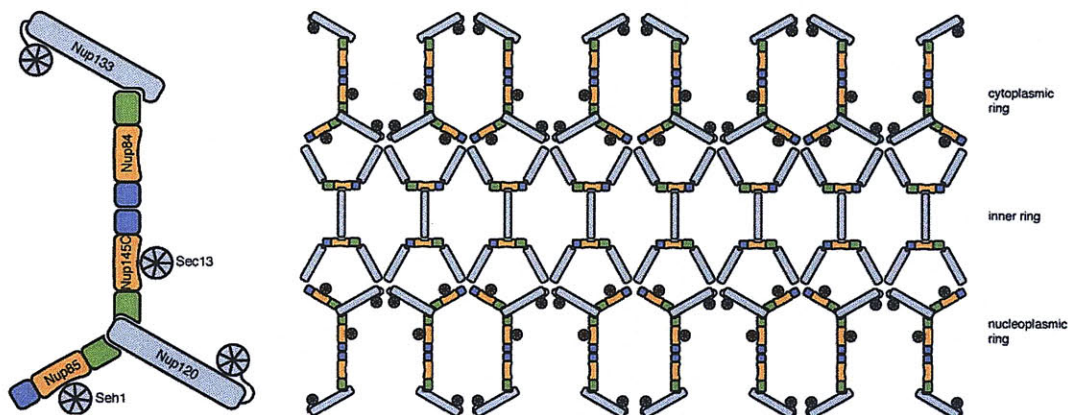
Figure 2.16 - Current model for the Y complex and the structural scaffold of the nuclear pore complex
The ACE1 proteins Nup85, Nup145C, Nup84, Sec31, and Nic96 are colored according to Figure 2.12. (A) Schematic diagram of COPII outer coat organization. The Sec31•Sec13 cuboctahedron composed of 24 edge elements (Sec31•Sec13 heterotetramers) is shown unwrapped and laid flat in 2 dimensions. The Sec31•Sec31 crown-crown interactions make edge elements while propeller-propeller interactions are vertex elements (Fath et al., 2007). (B) Schematic diagram of the predicted lattice-like organization of the structural scaffold of the NPC. The entire scaffold (8 scaffold segments) is illustrated unwrapped and laid flat in two dimensions. The Y complex comprises the nuclear and cytoplasmic rings, while the Nic96-containing complex makes up the inner ring. The relative position and interactions between the seven proteins in the Y complex are shown with Sec13, Seh1, Nup133, and Nup120 colored in gray. The remainder of the complex (Nup157/170, Nup188, and Nup192) is illustrated in gray. The illustration is not meant to predict relative positions of proteins or structure of the inner ring *per se*, rather it is meant to show the lattice-like organization of the structural scaffold similar to vesicle coating complexes.

Two of the three interface types observed in the outer COPII coat are also found in the NPC coat (Figure 2.16). Nup145C and Nup84 heterodimerize via their crown modules similar to Sec31 homodimerization, and the insertion of a seventh blade into an incomplete propeller domain is a recurring theme in Sec13•Sec31, Sec13•Nup145C, and Seh1•Nup85. Because Nic96 shares an ACE1 element, we predict that the inner scaffold ring is branched and lattice-like, as are the peripheral rings. We postulate that the Y and Nic96 complexes are both vertex elements in the assembly of the NPC structural scaffold. This would generate a lattice-like NPC coat similar to clathrin and COP coats (Fath et al., 2007; Fotin et al., 2004) (Figure 2.16B). This model explains how the relatively small mass of Nup subcomplexes fills the large volume observed for the scaffold structure of the NPC (Alber et al., 2007b) and is generally consistent with low-resolution images of NPCs (Beck et al., 2007; Stoffler et al., 2003). Notably, COP and clathrin cages do not directly contact membranes, but use adaptor protein (AP) complexes to span the ~10 nm gap between the surfaces (Owen et al., 2004). Consistent with a related architecture, a similar sized gap has been observed between the scaffold ring and membrane surface in intact NPCs (Beck et al., 2007).

The modular nature of COP and clathrin coats enables the construction of assemblies varying in composition and size (Cheng et al., 2007; Stagg et al., 2008) because the polygonal elements can be arranged in different ways. If the same principle applies to the NPC, it could explain the existence of a subset of NPCs that do not obey eightfold rotational symmetry (Hinshaw and Milligan, 2003) or further allow for the assembly of NPCs of distinct composition, possibly tailored to more specific tasks. These possibilities are now testable and will bring us closer to fully understanding the many functions of the NPC (See Appendix for additional discussion published separately).

Methods

Protein production and purification

Full-length Seh1 and a proteolytically-mapped Nup85 fragment (residues 1-564) from *S. cerevisiae* were cloned into the bi-cistronic pCOLA-Duet bacterial expression plasmid (EMD Biosciences) using the BamHI/NotI and NdeI/XhoI restriction sites, respectively, resulting in N-terminally His-tagged Seh1 and untagged Nup85 and transformed into *E. coli* BL21(DE3)-RIL. Nup85$_{1-564}$•Seh1 is referred to as Nup85•Seh1 in the text for simplicity. Cells were grown at 30 °C in Luria-Bertani broth supplemented with 0.4% glucose to OD$_{600}$= 0.8 and induced with 0.2 mM IPTG at 18°C for 18 hours. Cells were harvested by centrifugation, resuspended in 40 mM potassium phosphate pH 8.0, 500 mM NaCl, 20 mM imidazole, and 3 mM β-mercaptoethanol, and lysed using a French press. The crude lysate was centrifuged at 15,000g for 15 minutes. The soluble fraction was then incubated with 1 ml Ni-NTA per 1000 ODs for 30 minutes at 4°C and loaded onto a disposable column (Pierce). The column was washed with four bed volumes of 50 mM potassium phosphate pH 8.0, 400 mM NaCl, 30 mM imidazole, and 3 mM β-mercaptoethanol and the Nup85•Seh1 complex eluted in 4 bed volumes of 50 mM potassium phosphate pH 8.0, 250 mM NaCl, 250 mM imidazole, and 3 mM β-mercaptoethanol. Eluted protein was dialyzed against 10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.1 mM EDTA, and 1 mM DTT and the 6xHis-tag cleaved with PreScission protease. The protein was purified by anion exchange chromatography on a HiTrapQ column (GE Healthcare) via a linear NaCl gradient and twice by size exclusion chromatography using a Superdex S200 26/60 column (GE Healthcare) run in 10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.1 mM EDTA, and 1 mM DTT.

Selenomethionine derivatized Nup85•Seh1 was prepared by growing cells in M9 Medium (Sigma) supplemented with 1 mM MgSO$_4$, 6.6 μM CaCl$_2$, 1 ml FeSO$_4$ 4.2 g/L, 0.4% glucose, and 100 μl 0.5% (w/v) thiamine at 37°C (modified from (Doublie, 1997)). At OD$_{600}$= 0.5 solid amino acid supplements were added (100

74

mg/ml L-lysine, L-phenylalanine, and L-threonine; 50 mg/ml L-isoleucine, L-leucine, L-valine, and L-selenomethionine). After 30 minutes, cultures were transferred to 22°C for 20 minutes and then induced with 0.2 mM IPTG for 18 hours. The selenomethionine-derivatized Nup85•Seh1 complex was purified as described for the native version. Incorporation of selenomethionine was confirmed by mass spectrometry (data not shown). Both the native protein and selenomethionine derivative were concentrated to 30 mg/ml for crystallization.

Full length Nup85 in complex with Seh1 (Nup85$_{1-744}$•Seh1, referred to as Nup85(fl)•Seh1) was cloned and purified identically to Nup85$_{1-564}$•Seh1. Full-length Sec13 and Nup145C$_{109-712}$ from *S. cerevisiae* were cloned into the bi-cistronic pET-Duet bacterial expression plasmid using the BamHI/NotI and NdeI/XhoI restriction sites, respectively, resulting in N-terminally His-tagged Sec13 and untagged Nup145C. In order to stabilize the complex, the C-terminus of Sec13 and N-terminus of Nup145C were linked with a short flexible linker to generate a single chain. This linked version behaved identically to the complex made from separate chains with the advantage of increased stability and is here referred to as Nup145C•Sec13 for simplicity. The C-terminal tail module of Nup145C was removed from the construct to generate Nup145C$_{109-555}$•Sec13 by PCR and is referred to as Nup145CΔtail in the text. Nup145C-ELIEA•Sec13 and Nup145C(Q691G)•Sec13 were generated by PCR mutagenesis. Nup145C•Sec13 was produced identically to Nup85•Seh1 except that gel filtration was performed in 10mM Hepes/NaOH pH 7.4, 150mM NaCl, 1mM DTT, 0.1mM EDTA.

Full-length Nup84 from *S. cerevisiae* was cloned into a pET-Duet-derived plasmid using the BamHI/NotI restriction sites. The C-terminal tail module of Nup84 was removed from the construct to generate Nup84$_{1-482}$ by PCR and is referred to as Nup84Δtail in the text. Nup84-DSICD was generated by PCR mutagenesis. The protein was purified via metal-affinity and size exclusion chromatography as described for Nup145C•Sec13.

Full-length Nup120 from *S. cerevisiae* was cloned into a pET-Duet-derived plasmid using the BamHI/NotI restriction sites. A trimeric complex between full length Nup120, Nup85(fl), and Seh1 was produced essentially as in (Lutzmann et al., 2002) and purified as described for Nup145C•Sec13.

Crystallization

The Nup85•Seh1 complex was crystallized in 18 % (v/v) PEG 3350, 0.2 M sodium citrate, and 0.1 M bis-tris-propane pH 8.5 by the hanging drop vapor diffusion method at 18°C. Crystals grew within 4-10 days forming rods with dimensions of 150 μm x 150 μm x 400 μm. The selenomethionine derivative crystallized in the same conditions. Native crystals were cryoprotected in reservoir solution with 15% (v/v) PEG 400. Selenomethionine crystals were cryoprotected in the reservoir solution supplemented with 4% (v/v) additional PEG 3350 before flash freezing in liquid nitrogen. Both the native and selenomethionine Nup85•Seh1 complex crystallized in space group $P4_12_12$ with two molecules per asymmetric unit. Crystal screening was performed at beamline X6A at National Synchrotron Light Source (NSLS) and final X-ray data was collected at the NE-CAT beamline 24ID-C at Argonne National Laboratory.

Structure determination

Data reduction was carried out using HKL2000 (Otwinowski and Minor, 1997). The structure was solved with the single anomalous dispersion (SAD) technique using the selenomethionine derivative. The positions of 2*16 selenium sites (out of 2*20 possible) were found with the program SHELXD (Adams et al., 2002; Sheldrick, 2008) and were used for phasing. The NCS-averaged, solvent-flattened 3.7 Å experimental electron density map was of sufficient quality to trace the backbone of most of the model. The selenium positions served as markers to unambiguously assign the sequence for Nup85. Assigning the sequence of Seh1 was assisted by superimposing the structure of the homologous Sec13. The final model was refined against native data extending to

3.5 Å. Model building was carried out with Coot (Emsley and Cowtan, 2004), for refinement the PHENIX suite was used (Adams et al., 2002). Only few packing interactions exist in the crystal, resulting in a relatively high Wilson B-factor of 118 Å^2. Thus, B-sharpened maps were generated with CNS (Brunger et al., 1998; DeLaBarre and Brunger, 2006) and were used to assist side chain placement in the early stages of model building. NCS-restraints were applied throughout the refinement process. The final model has good stereochemistry and no Ramachandran outliers (84.1% of residues in preferred regions, 14.4% in additional allowed regions) according to Molprobity (Davis et al., 2007). All secondary structure elements of the 100 kDa heterodimer have been traced, however several loops connecting either helices in the trunk and crown of Nup85 or strands in Seh1 are omitted due to poor electron density in those regions.

Analytical gel filtration

For Nup145C•Sec13 and Nup84 binding experiments, equimolar amounts of Nup145C•Sec13 (or Nup145C-ELIEA•Sec13) and Nup84 were mixed and incubated at 4 °C for 5 minutes. Similarly, equimolar amounts of Nup145CΔtail•Sec13 and Nup84Δtail (or Nup84Δtail-DSICD) were mixed and incubated at 4 °C for 5 minutes. Reactions (500 μl) were injected onto a Superdex S200 10/300 column (GE Healthcare) and run in 10 mM Hepes/NaOH pH 7.4, 150 mM NaCl, 1mM DTT, 0.1 mM EDTA at a flow rate of 0.8 ml/min. Nup85•Seh1 complex at various concentrations (1, 5, 10, and 20 mg/ml) was loaded onto a Superdex S200 10/300 column and run in 10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.05 mM EDTA, and 0.5 mM TCEP at a flow rate of 0.8 ml/min. In Figure 2.6A, equimolar amounts of Nup145C•Sec13 and Nup85(fl)•Seh1 were mixed and incubated at 4 °C for 5 minutes before injection on a Superdex200 HR 26/60 column. In Figure 2.6B, equimolar amounts of Nup120 and Nup85•Seh1 were mixed and incubated at 4 °C for 5 minutes before injection on a Superose 6 HR 10/300 column (GE Healthcare). In Figure 2.6C, Nup120•Nup85(fl)•Seh1 was mixed with a 2-fold molar excess of Nup145C•Sec13 or Nup145CΔtail•Sec13 and incubated at 4 °C for 5 minutes before injection on a Superdex S200

HR10/300 column. In Figure 2.7, Ni-NTA elutions of Nup120•Nup85(fl)•Seh1 co-expressed with Nup145C-ELIEA•Sec13 or Nup145C(Q691G)•Sec13 (where both Nup120 and the Nup145C•Sec13 variant are 6xHis-tagged) were run on a Superdex 200 HR26/60 column. Gel filtration was carried out in 10 mM Hepes/NaOH pH 7.4, 150 mM NaCl, 1mM DTT, 0.1 mM EDTA at the columns recommended flow rates.

Analytical ultracentrifugation

Purified Nup85•Seh1 complex was gel-filtered in 10mM Tris/HCl pH 8.0, 150mM NaCl, 0.5mM TCEP, and 0.05 mM EDTA immediately prior to the experiments. Analytical ultracentrifugation experiments were carried out with an Optima XL-A centrifuge using An50Ti (6 hole, equilibrium runs) or An60Ti (4 hole, velocity runs) rotors.

Samples for sedimentation velocity (440 µl sample or 450 µl buffer) were loaded into Epon-charcoal filled 2 channel centerpieces, fit with sapphire windows, and spun at 42,000 rpm. Concentrations of 6.3, 3.2, and 0.63 µM Nup85•Seh1 were used. Sedimentation velocity data was analyzed globally to generate a c(s) distribution using the hybrid local continuous distribution and global discrete species model in SEDPHAT (Schuck, 2000). The data was fit from 2 to 10 s[-13] with Sedanal calculated vbar = 0.7319 cm^3/g, η = 1.0182 cP, and ρ = 1.00472 g/cm^3.

Samples for sedimentation equilibrium (110 µL sample or 120 µl buffer) were loaded into Epon-charcoal filled 6 channel centerpieces, fit with quartz windows, and spun at 13,500, 17,500, and 22,800 rpm, respectively. Two replicates of 6 concentrations (7.6, 6.3, 5.0, 3.8, 2.5, and 1.3 µM) were analyzed. Approach to equilibrium was monitored with Winmatch. Absorbance data was collected at 280 nm at the smallest possible step size with 5 replicates per step. Sedimentation equilibrium data (36 datasets total) was fit globally with Ultrascan 9.0 (http://ultrascan.uthscsa.edu) with a single ideal species model.

Isothermal titration calorimetry

Purified Nup145C•Seh1, Nup145C-ELIEA•Seh1, Nup145CΔtail•Seh1, Nup84, Nup84Δtail, and Nup84Δtail-DSICD were gel-filtered into 10mM Hepes/NaOH pH 7.4, 150mM NaCl, 1mM DTT, and 0.1mM EDTA immediately prior to the experiment. Protein concentrations were determined spectrophotometrically at 280 nm. ITC was performed using a VP-ITC microcalorimeter (MicroCal, Northampton, MA). Titrations were performed at 25 °C by injecting 6 µl aliquots of Nup145C•Seh1, Nup145C-ELIEA•Seh1, or Nup145CΔtail•Seh1 at 4.8 µM into the ITC cell containing 1.43 ml of Nup84, Nup84Δtail, or Nup84Δtail-DSICD at 0.4 µM. Binding stoichiometry, enthalpy, and entropy as well as the equilibrium dissociation constant was determined by using the "single set of independent sites" model of molecular association (MicroCal Origin 2.9; MicroCal).

CD spectroscopy

Nup145CΔtail•Sec13, Nup145CΔtail-ELIEA•Sec13, Nup84Δtail, and Nup84Δtail-DSICD were purified as described above and gel filtered into 5 mM Hepes/NaOH pH 7.4 and 150 mM NaCl immediately prior to the experiment. An Aviv Model 202 CD spectrometer was used for all experiments. CD signal at 208nm of 1.3 µM protein in a 1 mm pathlength cell was recorded at every degree during a 25-80°C temperature ramp with two minutes of equilibration time at each step.

In vivo localization experiments

Strains were grown in YPD (1% yeast extract, 2% yeast peptone, 2% glucose) at 24°C to an OD_{600} of 0.5-0.8. Cells were added to an equal volume of phosphate buffered saline (PBS, pH 7.3) with 20 mg/ml Hoechst dye (Invitrogen) for 30 minutes at room temperature to stain DNA. Cells were harvested by centrifugation, washed once in phosphate buffered saline (PBS) and viewed using a Nikon E800 fluorescent microscope (Melville, NY) mounted with a Hamamatsu digital camera (Bridgewater, NJ). Images were captured using Improvision OpenLabs 2.0 software (Lexington, MA) with identical exposure

times for each sample. Final images were constructed using Adobe Photoshop CS3 and Adobe Illustrator CS3 (Adobe Systems).

Yeast plasmid construction

The CEN/ARS plasmid SBYp115 containing the entire *NUP145* gene with the *NUP145* promoter and 3' UTR was constructed by gap repair in yeast. Sequences upstream and downstream of the *NUP145* ORF (−800 to −100 and +1 to +400) were amplified by PCR using Phusion DNA polymerase (New England BioLabs) with primer combinations oES143 (5'-aaaggatccGCAACACTTTCAATTGCATTTCTTCAA-3') with oES144 (5'-tttgaattcCAAACGAGTTAATTCTTTCTAATTTTT-3') and oES145 (5'-tatgaattcGACTGAAGCTAACGCTTTTGGAGTAAT-3') with oES146 (5'-aaagtcgacGAAAGAGATAGATTTCTGTAAGAAGGC-3'), respectively. The PCR products were cloned into the *Bam*HI/*Sal*I sites of pRS316 to make pES323. Gap repair of *Eco*RI-digested pES323 resulted in full-length *NUP145* in pRS316 (SBYp115). Plasmid SBYp116 (*NUP145, LEU2, CEN*) was constructed by ligating a *Bam*HI-*Sal*I fragment from SBYp115 into pRS315 (Sikorski and Hieter, 1989). Plasmid SBYp117 (*nup145C-ELIEA, LEU2, CEN*) was constructed by ligating a 1.5 kb *Aat*II-*Avr*II fragment from pSB210 into the same site of SBYp116.

Strain construction

Yeast strains used in this study are listed in Table 2.2. Genomic tagging of *NUP133* and *NUP84* in a *NUP145/nup145::KANMX4* diploid (BY4743 background, *Saccharomyces cerevisiae* deletion consortium) was done by homologous recombination (Janke et al., 2004) using pYM28 (*eGFP:HIS3MX6*) as template. Strains were selected on SMM-histidine plates and verified by western blotting with anti-GFP antibody and nuclear rim localization of the Nup-GFP chimeric proteins. The resulting diploids were transformed with SBYp115 (*NUP145, URA3, CEN*) to allow viable colonies following sporulation and tetrad dissection. Haploid strains were then transformed with SBYp116 (*NUP145,*

LEU2, CEN) or SBYp117 (*nup145C-ELIEA, LEU2, CEN*), grown in medium lacking leucine, and plated on 5-fluoroorotic acid plates (Figure 2.10).

Table 2.2 - Yeast strains used in this study

Strain	Genotype
YS221	MATa*nup145::KANMX4 NUP133-GFP:HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0* [SBYp116;(*NUP145, LEU2, CEN*)]
YS222	MATα *nup145::KANMX4 NUP133-GFP:HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0* [SBYp117; (*nup145C-ELIEA, LEU2, CEN*)]
YS223	MATα*nup145::KANMX4 NUP84-GFP:HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0* [SBYp116; (*NUP145, LEU2, CEN*)]
YS224	MATα *nup145::KANMX4 NUP84-GFP:HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0* [SBYp117; (*nup145-ELIEA, LEU2, CEN*)]
YS225	MATα*nup145::KANMX4 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0* [SBYp115;(*NUP145, URA3, CEN*)]

Cell fractionation

Strains expressing Nup133-GFP or Nup84-GFP were grown to log phase (OD$_{600}$ = 0.5-0.7) in YPD at 30 °C. Twenty-five OD$_{600}$ units were harvested by filtration, washed with cold water, and collected by centrifugation. Cells were pre-treated with 100mM Tris/HCl, pH 9.4, 0.5% 2-mercaptoethanol for 10 minutes at 30°C and spheroplasted in S buffer (40mM HEPES pH 7.5, 5 mM MgCl$_2$, 1.2 M sorbitol) containing 0.2mg/ml Zymolyase (100T) for 1 hour at 30 °C. Spheroplasts were washed 3 times with S buffer, resuspended in 0.5mL lysis buffer containing 20 mM HEPES, pH 7.5, 5 mM MgCl$_2$, 1 mM PMSF and protease inhibitor cocktail (Roche) and lysed on ice by Dounce homogenization. A portion of the total lysate was removed (T), and the remaining lysate was centrifuged at 16,000 xg for 30 minutes at 4 °C resulting in a soluble (S) and pellet (P) fraction. Equal cell equivalents were resolved by SDS-PAGE, transferred to nitrocellulose, and proteins were detected using antibody against GFP (1:20,000) or the cytosolic protein Pgk1p (1:3,000).

Acknowledgements

Chapter three

The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

This chapter was previously published as Leksa, N.C.*, Brohawn, S.G.* & Schwartz, T.U. The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture. Structure 17, 1082-1091 (2009). *These authors contributed equally to this work.

N.C.L. designed, conducted and analyzed crystallographic and *S. cerevisiae* experiments and wrote the manuscript; S.G.B. designed, conducted and analyzed crystallographic and biochemical experiments and wrote the manuscript; T.U.S. advised on all aspects of the project and wrote the manuscript.

Abstract

Nucleocytoplasmic transport is mediated by nuclear pore complexes (NPCs), enormous protein assemblies residing in circular openings in the nuclear envelope. The NPC is modular, with transient and stable components. The stable core is essentially built from two multiprotein complexes, the heptameric Y complex and the Nic96 complex, arranged around an eightfold axis. We present the crystal structure of Nup120$_{1-757}$, one of the two short arms of the Y complex. The protein adopts a compact oval shape built around a novel bipartite α-helical domain intimately integrated with a β-propeller domain. The domain arrangement is substantially different from the Nup85•Seh1 complex, which forms the other short arm of the Y. With the data presented here, we establish that all three branches of the Y complex are tightly connected by helical interactions and that the β-propellers likely form interaction site(s) to neighboring complexes.

Introduction

The main feature that distinguishes eukaryotes from prokaryotes is the confinement of the genetic material into a membrane-enveloped nucleus. Because gene transcription and mRNA processing occur inside the nucleus and protein translation is restricted to the cytoplasm, transport across the double-layered nuclear envelope (NE) is essential for cellular homeostasis. The exchange of all molecules, including ions, proteins, and RNAs, is facilitated exclusively by nuclear pore complexes (NPCs) (D'Angelo and Hetzer, 2008; Lim et al., 2008; Tran and Wente, 2006; Weis, 2003). NPCs are large protein assemblies of 40-60 MDa that are embedded in the nuclear envelope and exhibit an 8-fold rotational symmetry around a central axis in addition to an imperfect two-fold symmetry across the plane of the NE (Beck et al., 2007; Stoffler et al., 2003)(Beck et al., 2007; Stoffler et al., 2003). Composed of multiple copies of ~ 30 proteins, termed nucleoporins (nups), the NPC has an outer diameter of ~100 nm whereas the central channel measures ~40 nm in width. Transmembrane nups directly connect the NPC to the NE, while the phenylalanine-glycine (FG)

repeat-containing Nups line the interior of the pore. These FG-filaments mediate nucleocytoplasmic transport of cargo molecules across the NE. FG-filament bearing nups are anchored to the NPC scaffold built from architectural nucleoporins arranged in two large multiprotein complexes that form a membrane-proximal layer. The scaffold structure is very stable and undergoes virtually no turnover in the quiescent cell (D'Angelo et al., 2009), whereas many other nucleoporins have variable dwell times at the NPC (Rabut et al., 2004). In consequence, the NPC is a highly modular structure (Schwartz, 2005). Understanding the structure of the NPC therefore depends upon elucidating its basic scaffold.

The two essential architectural building blocks of the NPC are the Nup84 or Y complex and the Nic96 subcomplex. The components of the Nic96 subcomplex likely include Nic96, Nup53/59, Nup157/170, Nup188 and Nup192 (yeast nomenclature), as inferred from co-immunoprecipitations (co-IPs) (Alber et al., 2007b; Hawryluk-Gara et al., 2008; Marelli et al., 1998; Onischenko et al., 2009) and yeast-two-hybrid screens (Wang et al., 2009; Yu et al., 2008). Judged by immunolabeling, the Nic96 subcomplex might form a central ring within the NPC sandwiched between peripheral rings formed by Y complexes (Alber et al., 2007b). In comparison to the Nic96 subcomplex, the Y complex is substantially better understood. It has 7 universally conserved members (yeastNup84/humanNup107, yNup85/hNup75, yNup120/hNup160, Nup133, yNup145C/hNup96, Sec13, and Seh1) and three additional members (Nup37, Nup43, and ELYS/Mel-28) to date found mainly in metazoa (Cronshaw et al., 2002; Gillespie et al., 2007; Loiodice et al., 2004; Rasala et al., 2006). In the fungus *Aspergillus nidulans*, distant Nup37 and ELYS orthologs have been described recently (Liu et al., 2009). The heptameric core Y complex assembles tightly as shown by co-IPs and *in vitro* assembly (Harel et al., 2003; Lutzmann et al., 2002; Siniossoglou et al., 2000; Walther et al., 2003). Negatively-stained electronmicrographs of the assembled Y complex reveal a branched Y-shaped

structure, with two short arms and a kinked stalk connected at a central hub (Lutzmann et al., 2002).

Crystallographic analysis of the Y complex has progressed quickly. The kinked stalk ends with a flexibly attached β-propeller domain (Berke et al., 2004) at the N-terminus of Nup133 followed by an irregular C-terminal helical stack domain that connects end-to-end to Nup84 (Boehmer et al., 2008; Whittle and Schwartz, 2009). The Nup84•Nup133 interface defines at least one kink in the stalk. The opposite end of Nup84 links to Nup145C (Brohawn et al., 2008). Nup145C•Sec13 (Hsia et al., 2007) resides proximal to the hub (Lutzmann et al., 2002). Seh1•Nup85 forms one of the two short arms of the Y complex (Brohawn et al., 2008; Debler et al., 2008). Nup84, Nup85, and Nup145C are structurally related (Brohawn et al., 2008), despite very low sequence conservation, as are the β-propeller proteins Seh1 and Sec13.

Nup120 is the last remaining Y complex component without structural information. Here we report the 3.0 Å crystal structure of Nup120 (residues 1-757 of 1037), which reveals a compact and rigid structure composed of an N-terminal β-propeller domain tightly integrated into a novel bipartite α-helical domain. Our structure largely defines the second short arm of the Y complex. Comparison with other members of the Y complex, phylogenetic analysis, *in vitro* binding experiments, and *in vivo* localization data suggest a role for Nup120 consistent with our lattice-like model of the NPC.

Results

Structure determination

After systematic C terminal truncation, a stable fragment comprising most of Nup120 (residues 1-757 of 1037 total) from *S. cerevisiae* was recombinantly expressed in *E. coli* and purified. The protein is a monomer in solution (data not shown). Native protein readily crystallized and selenomethionine derivatized crystals were obtained after microseeding with native crystals. Though both crystal forms were optically identical, the selenomethionine crystals diffracted better and were used exclusively in structural analysis. The structure of Nup120 was solved with one molecule per asymmetric unit by single-wavelength anomalous dispersion (SAD) on very strong Se-Peak data (all 9 Se sites are well ordered). The model is complete except for 26 residues at the C-terminus and three flexible loops and was refined to Rwork / Rfree of 24.4 % / 29.9 % (Table 3.1).

Table 3.1 - Data collection and refinement statistics

	$Nup120_{1\text{-}757}$ Selenomethionine
Data collection	
Space group	$P2_12_12$
Cell dimensions	
a, b, c (Å)	114.6, 153.7, 53.0
α, β, γ (°)	90, 90, 90
Resolution (Å)	50-3.0 (3.1-3.0) *
R_{sym} (%)	5.1
$I/\sigma I$	20.4 (1.8)
Completeness (%)	99.8 (99.6)
Redundancy	3.0 (3.0)
Refinement	
Resolution (Å)	50.0 - 3.0
No. reflections	35841
R_{work} / R_{free}	24.4 / 29.9
No. atoms	
Protein	5305
Water	0
B-factors ($Å^2$)	
B-Propeller	114
α-helical insertion	89
α-helical domain	83
Ramachandran plot (%)	
Favored/allowed/outliers	93.6 / 5.3 / 1.1
RMSD bond lengths (Å)	0.017
RMSD bond angles (°)	1.915

*Values in parentheses are for highest-resolution shell.

Crystal structure of Nup120

Nup120 folds into a continuous, prolate disk with overall dimensions of 90 Å x 55 Å x 35 Å. One half of the structure is formed by an N-terminal β-propeller domain that is intimately connected to a compact central domain built from two closely packed α-helical segments (Figure 3.1). Overall, the structure is better resolved in the α-helical segment than the β-propeller, likely a result of a paucity of packing contacts involving the latter. The β-propeller of Nup120 contains 7 consecutive blades that fan out from a central axis. The blades are formed by a β-sheet of 4 consecutive antiparallel strands, labeled A-D. Blade 7 is built from the very N-terminus of the polypeptide chain forming strand 7D and joining strands 7A-C to close the propeller in a Velcro-like closure commonly observed in β-propeller domains (Chaudhuri et al., 2008). Blade 1 is 5-stranded, with

strand 7D extending to form the additional strand 1E before connecting to strand 1A (Figure 3.1E). Blade 3 is somewhat unusual in that the outermost strand 3D is only loosely connected to strand 3C with a hydrogen-bonding network hardly visible in our structure.
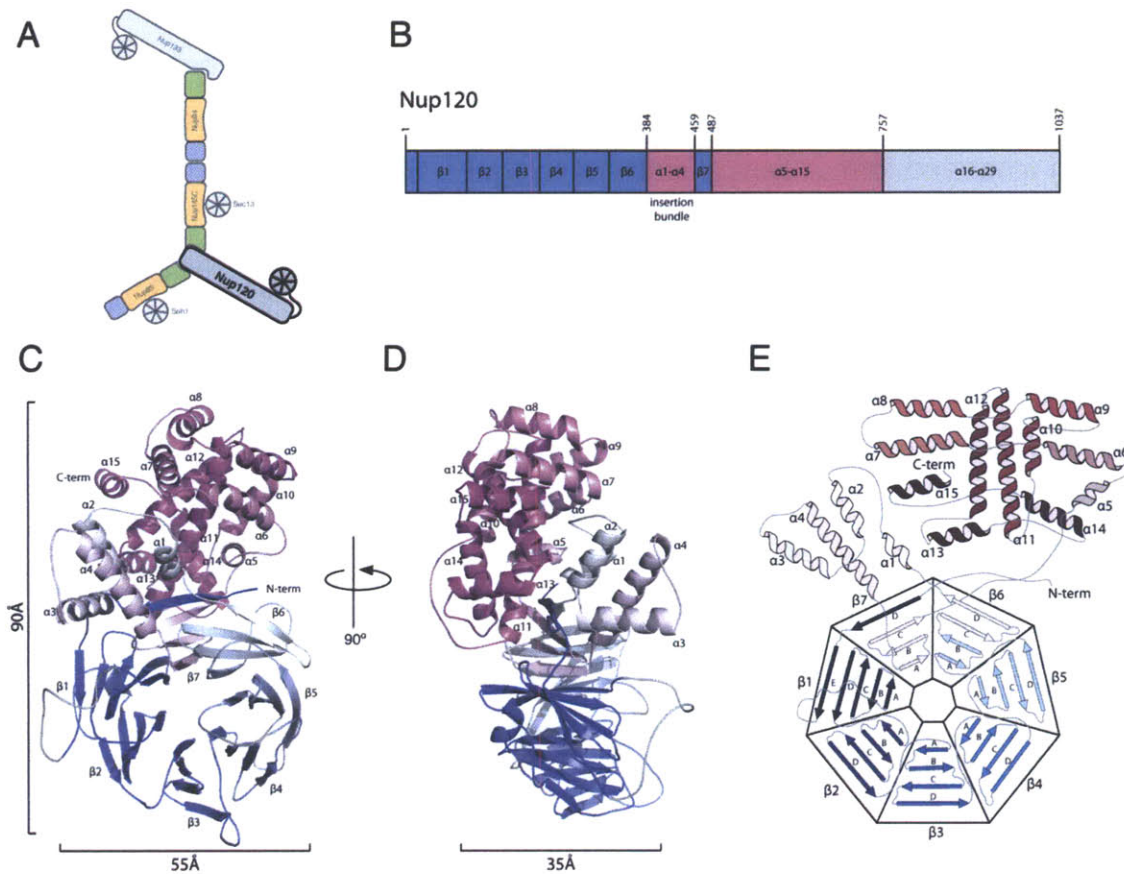
Figure 3.1 - Overall topology of Nup120
(A) Current model of the Y complex. The relative position of Nup120 is highlighted. (B) Schematic of full-length Nup120 from *S. cerevisiae*. Residues that form the β-propeller are colored blue, those that form the α-helical domain are purple, and those not present in the crystallized construct are in gray. (C,D) The overall topology of Nup120 (residues 1-757 of 1037) is shown in two views rotated by 90°. The structure is gradient-colored from blue to white to magenta from N- to C-terminus. At its N-terminus, Nup120 forms a 7-bladed β-propeller. A 4-helix bundle (α1-α4) between blades 6 and 7 packs against the remainder of the helical domain (α5-α15), composed of helices wrapping around a central hydrophobic stalk of the two long helices α11 and α12. Unstructured loops absent from the final model are shown in gray. (E) A topological diagram of the Nup120 structure is shown, illustrating the 4-helix insertion between blades 6 and 7 of the propeller as well as the two central helices of the helical domain.

89

The α-helical domain that forms the second half of the molecule is constructed in a unique discontinuous manner. In total the domain contains 15 helices, labeled α1-α15. The first 4 helices form a compact bundle and are inserted between blades 6 and 7 of the β-propeller. The remaining 11 helices are C-terminal to the β-propeller and pack tightly against the 4-helix bundle to form one compact entity. The arrangement of the helices is highly irregular. The most prominent feature of the domain are two long helices, α11 and α12, which pack against each other and form a central stalk, defining the long axis of the domain. Helices α5-α9 wrap up and around this element, with helices α6/α7 and α8/α9 arranged in two stacked braces oriented perpendicular to the stalk. Helices α1, and α13-α15 meander back down and around the other side to bury most of the hydrophobic stalk. The remaining surface area of the two central helices is closed by the 4-helix insertion bundle. Taken as a whole, the structure of $Nup120_{1-757}$ consists of a bipartite helical domain that is interrupted by a β-propeller.

The main crystal contact is formed by a domain swap

Other than a collection of spurious small contacts, crystal packing is mainly achieved by a domain swap of the terminal helices α15 and α15' exchanging between two neighboring molecules (Figure 3.2A). The interface measures 1355 $Å^2$, is entirely hydrophobic and highly complementary (Figure 3.2B). Domain swaps are regularly found in crystals (Liu and Eisenberg, 2002) and, as stated above, we do not observe dimerization of Nup120 in solution. We cannot rule out the possibility that the interface is physiologically relevant; sterically the domain swap is conceivable in the context of the entire molecule including the C-terminal 280 residues omitted in our construct. It is however more likely that the exposed hydrophobic patch is artificially generated by the truncation of the domain, since we also do not observe particularly high sequence conservation within helix α15. We speculate that *in vivo* the patch likely accommodates one of the additional helices from the C-terminal domain, or alternatively, is involved in

interaction with a neighboring molecule. Whether the C-terminal domain is rigidly or flexibly tethered to Nup120$_{1-757}$ is an open question.
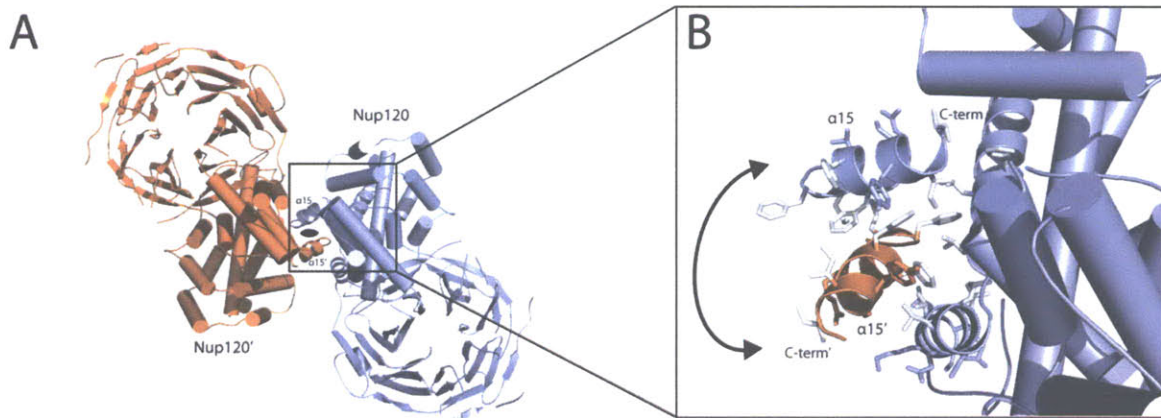
Figure 3.2 - Crystal contacts between two symmetry-related molecules
(A) One molecule of Nup120 in blue, and its symmetry mate, related by a 2-fold rotation, in orange. The β-propellers are at opposite ends while the helical domains engage in a putative domain swap between helices α15 of both molecules. (B) Close-up of the domain-swapped region, illustrating the hydrophobic nature of helix α15 and the surrounding pocket (hydrophobic residues are shown in white). In monomeric Nup120, helix α15 likely folds under (arrow) and occupies the position taken by helix α15' (orange) of the symmetry-related molecule in the crystal.

Conservation of Nup120 and comparison to the human ortholog Nup160

Overall, sequence conservation between Nup120 orthologs is weak as is typically observed in scaffold nucleoporins (Brohawn et al., 2008; Jeudy and Schwartz, 2007). Most of the better-conserved residues are buried in the hydrophobic core of the protein and are involved in maintaining the structural integrity of the protein. On the protein surface we find few conserved patches (Figure 3.3A). Most distinct is an area on the edge of the β-propeller, corresponding to the outer strands of blade 3 and the loop leading into blade 4. The conserved sequence begins in the 3BC loop and continues into strand 3C itself. Although generally buried in canonical β-propellers, here strand 3C is quite exposed. This is probably the result of weaker interactions with strand 3D, which is flanked by two large loops and peels away from the core of the propeller. Additional conserved residues are spotted around this area, creating a relatively

large conserved patch. The potential significance of this observation is discussed below.

A

Bottom View   Top View   β-Propeller View

B

Bottom View   Top View   β-Propeller View
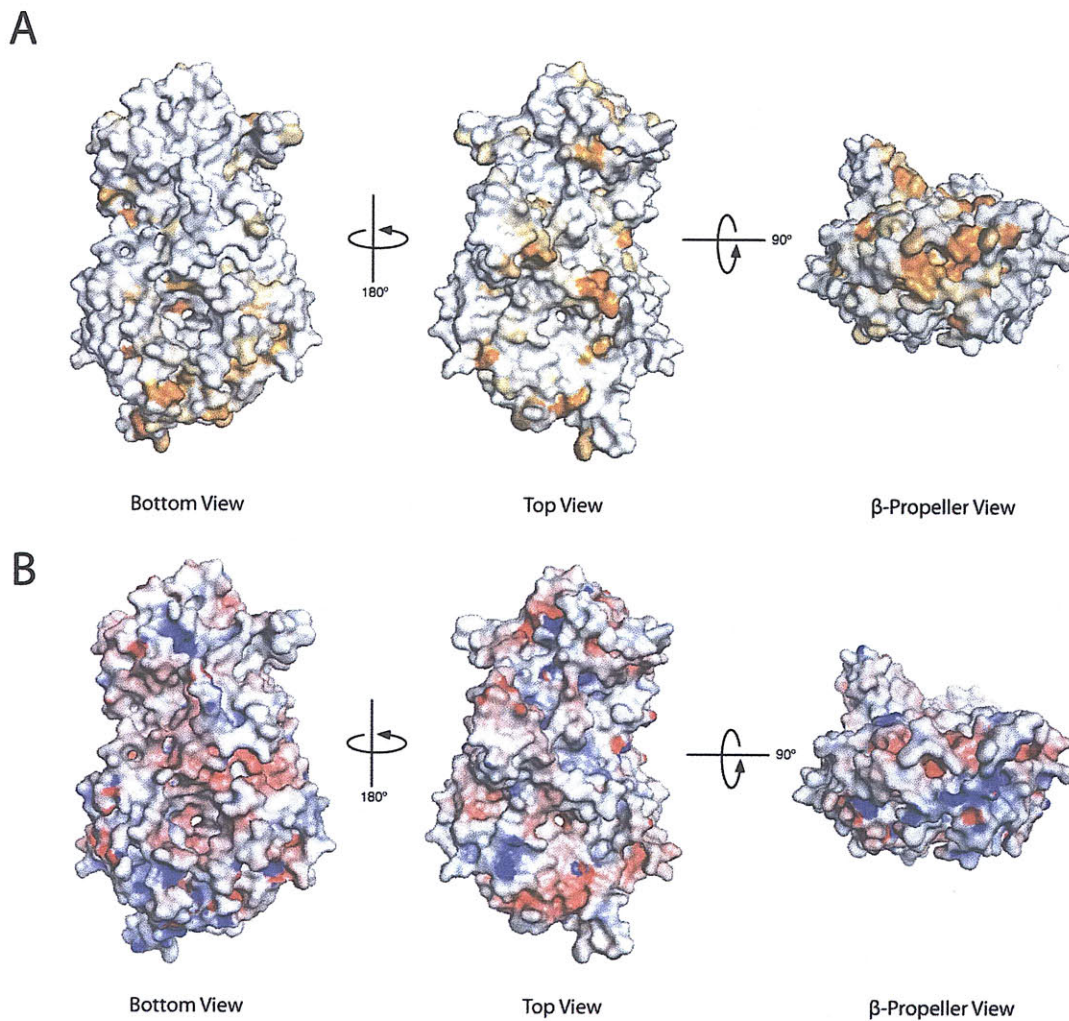
Figure 3.3 - Surface conservation and electrostatics of Nup120
(A) Surface conservation of Nup120 is shown from three different views. To illustrate the conservation of residues on the surface of Nup120, a multiple sequence alignment sampling the phylogenetic tree of budding yeasts was generated and mapped onto the surface, colored from white (not conserved) to orange (highly conserved). The view in the middle panel corresponds to the view shown in Figure 3.1C. A patch of highly conserved residues is apparent on the outer face of the propeller domain of Nup120. (B) The electrostatic surface potential of Nup120 is shown in the same views as in (A) and is colored from red (-10 kT/e) to blue (+10 kT/e).

We analyzed the charge distribution on the surface of Nup120 (Figure 3.3B). Since Nup120 is part of the scaffold structure of the NPC, we asked whether it may be possible that it directly juxtaposes the pore membrane. This would also be consistent with a membrane-curvature sensing ALPS motif, predicted in helix α5- α6 of Nup120 (Drin et al., 2007). The surface charge of $Nup120_{1-757}$, however, is fairly mixed without conserved positive patches that might suggest direct membrane interaction. The ALPS motif is deeply embedded in the structure and it is rather unlikely that it would swing out and insert in the membrane. Thus we suggest that $Nup120_{1-757}$ does not directly touch the nuclear membrane.

Structure-guided sequence comparison of Nup120 and its human ortholog Nup160 strongly suggests that both proteins adopt the same unique fold despite a low sequence identity of ~10 %. Both non-canonical characteristics of Nup120 (the helical insertion between blades 6 and 7 of the N-terminal β-propeller and the long central stalk helices forming the hydrophobic core of the central domain) are clearly conserved in Nup160. The 279 additional residues of Nup160 are dispersed over several regions and mostly correspond to different loop lengths connecting α-helices and β-strands. Of note, the C-terminal domain of Nup160, which is not present in the Nup120 crystal structure described here, has 5 additional predicted helices, possibly indicating a vertebrate-specific extension. Despite these differences, the Nup120 crystal structure is likely generally representative of all Nup120/Nup160 orthologs.

The C terminus of Nup120 directly binds Nup145C and Nup85

We sought to map the interaction of Nup120 with its direct binding partners in the Y complex, Nup145C and Nup85. In a gel filtration assay, we tested for the formation of a pentameric Sec13•Nup145C•Nup120•Nup85•Seh1 complex (Figure 3.4). Incubating $Nup120_{766-1037}$ or $Nup120_{1-757}$ with both Nup145C• Sec13 and Nup85•Seh1 resulted in complex formation only for the C-terminal Nup120 domain, but not for the crystal construct. In combination with previous interaction

mapping experiments (Brohawn et al., 2008), we conclude that the helical tails of the ACE1 domains of both Nup145C and Nup85 each interact directly with the helical Nup120$_{766-1037}$. This positions the C-terminus of Nup120 at the center of the hub of the Y complex.
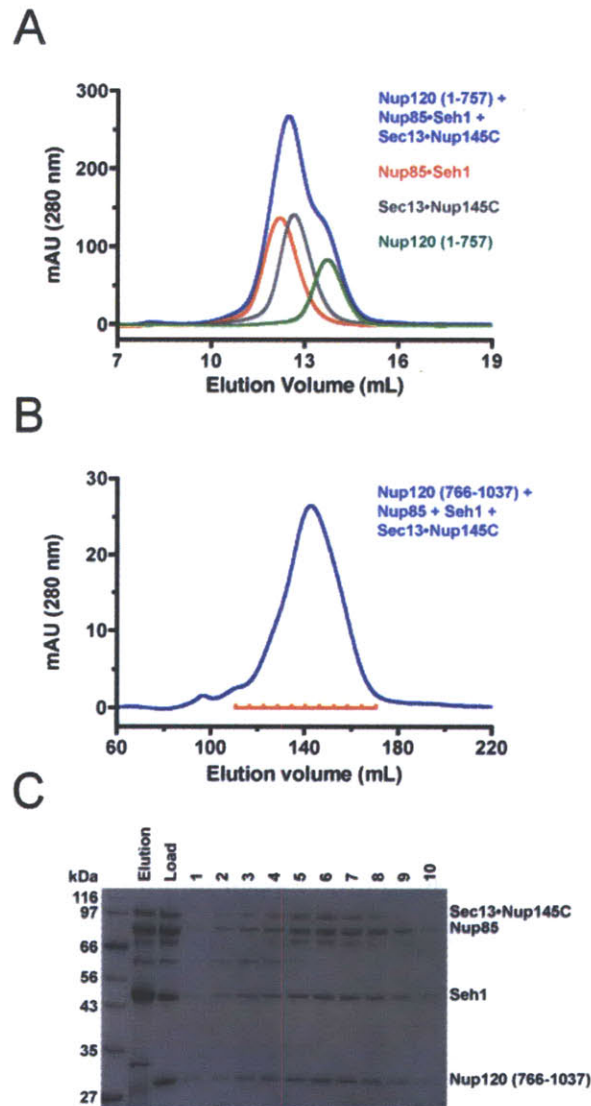
Figure 3.4 - The C-terminus of Nup120 is necessary for binding Nup85•Seh1 and Nup145C•Sec13

(A) Nup85•Seh1 (red), Nup145C•Sec13 (gray), and Nup120$_{1-757}$ (green) were run individually and in combination (blue) on a Superdex S200 10/300 gel filtration column. (B) Nup85•Seh1, Nup145C•Sec13, and Nup120$_{766-1037}$ were incubated together and run on Superdex S200 26/60 and eluted in a single peak. (C) Fractions from the gel filtration experiment in B were analyzed by SDS-Page.

Co-migration of Nup85•Seh1, Nup145C•Sec13, and Nup120$_{766-1037}$ indicates that the C-terminus of Nup120 is necessary for the formation of the pentameric complex that comprises the hub of the Y complex.

Without its C-terminal domain Nup120 does not properly localize to the NPC

Having established that Nup120$_{766-1037}$ is sufficient to bind both Nup145C•Sec13 and Nup85•Seh1 *in vitro*, we sought to examine the integration determinants of Nup120 into the NPC *in vivo*. NUP120 is not essential in yeast but nup120Δ cells exhibit a pore clustering phenotype (Aitchison et al., 1995; Heath et al., 1995) that is reminiscent of, but less severe than, the pore clustering observed for other scaffold nucleoporins including Nup84 and Nup133 (Li et al., 1995; Pemberton et al., 1995; Siniossoglou et al., 1996). We genomically GFP-tagged full length Nup120 and replaced the C-terminal 280 residues of genomic Nup120 with an in frame GFP-tag to create strains expressing Nup120-GFP or Nup120$_{1-757}$-GFP in a BY4741 background and examined the localization of the proteins via immunofluorescence (Figure 3.5). Nup120-GFP properly localizes to the NPC and shows typical nuclear rim staining, superimposing well with mAb414-staining of FG-Nups (Aris and Blobel, 1989). Nup120$_{1-757}$-GFP, on the other hand, does not properly localize to the nuclear envelope and shows staining throughout the cell. This result is consistent with our *in vitro* data and suggests that the integration into the Y complex is important for proper localization of Nup120.
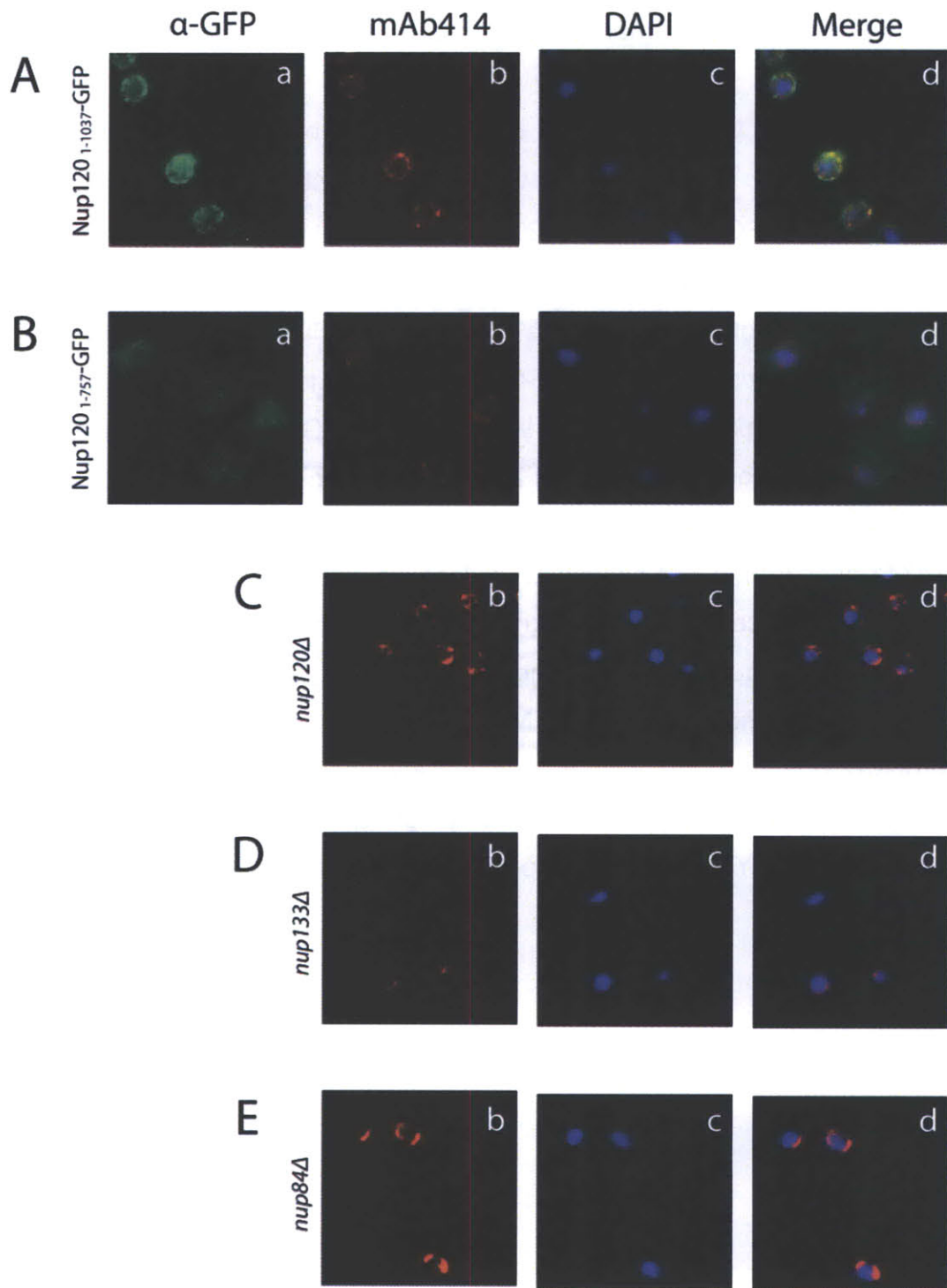
Figure 3.5 - Nup120$_{1-757}$ does not localize to the nuclear envelope
(Aa-Ad) Nup120-GFP is targeted to the nuclear envelope, as confirmed by co-localization with mAb414 (staining FG-Nups), while Nup120$_{1-757}$-GFP (Ba-Bd) is distributed throughout the cell. Mislocalization indicates that the C-terminus of Nup120 is necessary for proper recruitment to the NPC. *Nup120Δ, nup133Δ,* and *nup84Δ* cells (in the same BY4741 strain background) are shown for comparison (Cb-Cd,Db-Dd,Eb-Ed). Nuclear rim was visualized using mAb414, GFP-tagged Nup120 using goat α-GFP, and DNA using DAPI. Merged images are shown on the right.

Nup120 is topologically different from other scaffold nucleoporins

A recent surge in the X-ray crystallographic analysis of components of the NPC has greatly increased the repertoire of available structures of nucleoporins constituting the structural scaffold of the NPC. These structures (including those of Nic96 (Jeudy and Schwartz, 2007; Schrader et al., 2008b), Nup133-NTD (Berke et al., 2004), Nup133•Nup107 interaction complex (Boehmer et al., 2008), Nup145C•Sec13 (Hsia et al., 2007), and Nup85•Seh1 (Brohawn et al., 2008; Debler et al., 2008) as well as associated biochemical experiments, have led to a deeper and broader understanding of how the scaffold of the NPC is assembled from its constituent parts.

The structural subunits of the NPC were initially predicted to be composed of simple combinations of regular α-helical solenoids and β-propellers (Devos et al., 2006). Experimental data now allows us to specify these broad classifications, which should help to more specifically address the ancestry of the NPC. Both Sec13 and Seh1 form open, 6-bladed propellers that are completed in trans by the N-terminal insertion blades of their binding partners Nup145C and Nup85, respectively. Furthermore, helical nucleoporins Nic96, Nup145C, Nup85, and Nup84 are built around a common and distinct ancestral coatomer element (ACE1) shared with Sec31 of the outer coat of COPII vesicles (Brohawn et al., 2008). In ACE1 proteins, a specific N-terminal elaboration is followed by a tripartite helical domain composed of a trunk, a crown, and a tail module. The ~ 30 helices within ACE1 follow a J-like pattern, zigzagging up on one side of the trunk, making a U-turn within the crown domain, and then following down on the

opposite side of the trunk (Figure 3.6A, right panel). The tail module is often attached with modest flexibility to the trunk and is missing in most crystal constructs. In the case of Nup145C and Nup85 the N-terminal elaborations are the aforementioned insertion blades that bind to Sec13 and Seh1. Nup145C•Sec13 and Nup85•Seh1 heterodimers form the two proximal segments of the Y complex and are tethered together by Nup120 (Brohawn et al., 2008).

Based on structure predictions and its overall size, it was reasonable to suggest that Nup120 may take on a structure similar to Nup145C•Sec13 and Nup85•Seh1, with the only major difference being that the β-propeller and the α-helical domains are fused into one polypeptide chain. However, comparison between the structure of Nup120 and the Nup85•Seh1 heterodimer reveals a marked difference in topology (Figure 3.6A). Whereas the ACE1 architecture of Nup85 forms an elongated α-helical domain, the central α-helical domain of Nup120 is nearly as wide as it is long, forming an almost globular structure. The ACE1 trunk module covers the bottom face of the Seh1 β-propeller, while in Nup120 the helical domain is attached to and integrated into an edge of the β-propeller. Further, the ACE1•β-propeller interaction is accomplished by the addition of an insertion blade N-terminal to ACE1, while in Nup120 the β-propeller domain inserts a 4-helix bundle into the central α-helical domain. This helical insertion fits snugly into a pocket formed by helices α5-α7 and α11-α13 and buries a surface of nearly 600 Å^2 (Figure 3.6B).

The extensive interaction between the β-propeller and α-helical domain of Nup120 buries a total surface area of 2175 Å^2 and creates a rigid interface. In contrast, the largest buried surface area between ACE1 and its β-propeller partner is at the insertion blade/β-propeller interface. Additional contact areas in ACE1•β-propeller complexes are smaller in comparison to the corresponding interfaces in Nup120 and, importantly, far less hydrophobic. Thus, for the

ACE1•β-propeller assembly one has to consider substantial flexibility about the interaction joint, while the Nup120 structure presented here is very likely inflexible. Not only does the structure of Nup120 significantly differ from the ACE1•β-propeller heterodimers, but additional emerging evidence suggests that it also lacks similarity to Nup170 and Nup133, the two other scaffolding nucleoporins of similar size and domain composition with an N-terminal β-propeller followed by an α-helical domain (Whittle and Schwartz, 2009).
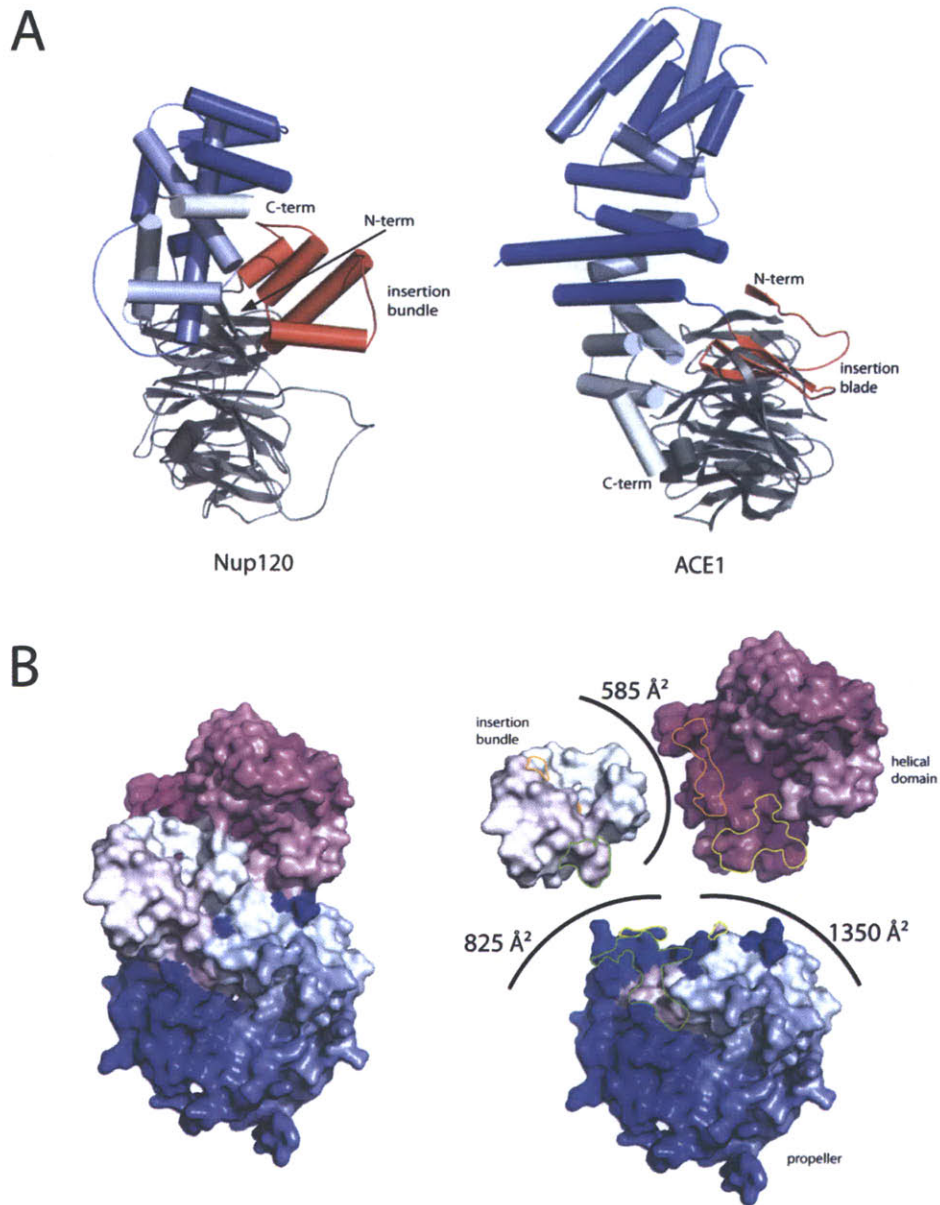
Figure 3.6 - Nup120 is composed of a combined β-propeller - α-helical domain distinct from ACE1•β-propeller

(A) The overall architectures of Nup120 and the ACE1 motif of Nup85•Seh1 are distinctly different. Nup120 is characterized by a bipartite helical domain (blue to white from N- to C-terminus) that is interrupted by a β-propeller domain (gray). The Nup85 ACE1 motif is characterized by an elongated helical stack (colored blue to white from N- to C-terminus) that makes a U-turn in the crown domain of the molecule. At its N-terminus, Nup85 inserts a blade (in red) into the open, 6-bladed Seh1 β-propeller. In contrast, the β-propeller of Nup120 contributes a helical insertion bundle (red) to the helical domain. The view of Nup120 is the same as that in Figure 3.3B. (B) Surface representations of intact Nup120 are shown on the

left, while on the right the three modules of Nup120 – the propeller, the helical insertion, and the helical domain – are shown pulled apart to illustrate the buried surface areas in between. Interacting surfaces between the propeller and the insertion bundle are outlined in green, between the propeller and the helical domain in yellow, and between the insertion bundle and the helical domain in orange. The molecule is N-to-C gradient-colored from blue-to-white-to-magenta.

Discussion

Here we report the crystal structure of Nup120, a large, universally conserved architectural nucleoporin. This structure adds substantially to the growing inventory of crystallographically characterized nucleoporins. As a result of these studies, we learn that the NPC is constructed from nucleoporins with a limited set of domain architectures. While other α-helical and β-propeller domains of scaffold nucleoporins fall into distinct classes, likely pointing to gene duplication in the early evolution of the NPC, the Nup120 architecture appears to be quite distinct. A search for structurally related proteins fails in detecting similarity beyond the isolated β-propeller scaffold or the arrangement of more than 6 α-helices. Within the list of crystallographically uncharacterized nucleoporins, none is likely to match the Nup120 structure closely.

Nup120 in the context of the NPC scaffold

Nup120 forms one of the two short arms of the universally conserved, 0.6 MDa Y complex, the essential building block of the NPC scaffold. The assembly of the Y complex from its 7 members is fairly well understood and has been studied using many different techniques. All of these studies profit from generally very high affinities observed between the interacting proteins within the Y, which generated largely consistent co-immunoprecipitation and yeast two-hybrid results and facilitated the crystallization of several complex crystal structures. However, models for the assembly of the NPC structural scaffold and the integration of the Y complex are still controversial.

Based on a combination of computational, structural, biochemical, and *in vivo* experiments, a model was proposed where the Y complex is positioned in two 8-membered rings located at the periphery of the NPC sandwiching two equally wide rings composed of Nup157/170,Nup188 and Nup192 in between (Alber et al., 2007b).

Blobel and coworkers proposed a concentric cylinder model based on crystal-packing interactions where four 8-membered rings of the Y complex are stacked and placed directly adjacent and in contact to the curved membrane (Debler et al., 2008; Hsia et al., 2007). Further, Seh1•Nup85 and Sec13•Nup145C are both supposed to form heterooctameric fence poles spanning the NPC vertically, thereby connecting the 4 stacked rings. Nup157/170, Nup188, Nup192 and Nic96 are suggested to form a second inner layer bridging to a third layer composed of FG-nups. With a scaffold twice the mass of the computer-generated model, the concentric cylinder model generates a densely packed NPC coat.

In contrast, we proposed a lattice-like model for the NPC, extrapolated from the assembly of COPII vesicle coats and substantiated by the structure and assembly principles of core components of the NPC scaffold (Brohawn et al., 2008; Brohawn and Schwartz, 2009a). We propose the Y complex does not directly coat the pore membrane (in analogy to the COPII outer coat), but is anchored by another set of proteins, likely involving the essential transmembrane nucleoporin Ndc1 and/or its direct binding partners (Onischenko et al., 2009). It is of interest to discuss this issue in respect to the membrane-inserting ALPS motif that was experimentally characterized within a loop structure in hNup133-NTD and that was predicted to occur as well in yNup85 and yNup120 (Drin et al., 2007). Based on the structural data now available on both yNup85 and yNup120, it appears unlikely that the predicted ALPS motif in both proteins is functional in membrane-binding since neither is in an exposed region of the protein, or is likely to become exposed. This is in contrast to the ALPS motif in hNup133-NTD, where it is well exposed in the crystal structure, and also highly conserved in

metazoa (Berke et al., 2004). Taking all the available data together, it appears more reasonable to suggest a specific function for the ALPS motif in metazoan Nup133 rather than a general function in anchoring of the NPC to the pore membrane. Since Nup133-ALPS is only poorly conserved in yeast, it is tempting to speculate that it may have a specific role in NPC assembly in open mitosis (Guttinger et al., 2009).

We predict that the lattice scaffold of the NPC is built from edge and vertex elements, following similar assembly principles as established for COPII. However, in the absence of definitive inter-subcomplex interaction data, any detailed NPC assembly model is still premature and has to be interpreted cautiously.

The fact that inter-subunit interactions are still obscure suggests that these interactions are rather weak and hard to establish. Each short arm of the Y complex contains one β-propeller domain, while the stalk contains two (Figure 3.1). For the assembly of the extensions of the Y, direct interactions between the α-helical domains is essential, however this does not exclude the participation of the β-propellers. It is reasonable to suggest that the β-propellers are prime candidates for the elusive inter-subcomplex contacts. The vertices of the outer coat of COPII vesicles are assembled exclusively via β-propeller interactions, which have still only been inferred by fitting crystal structures into EM maps (Fath et al., 2007; Stagg et al., 2008). β-propellers make excellent protein-protein interfaces due to their inherent ability to pair with a binding partner in multiple modes. Binding to peptides via the face of the β-propeller is well known (Jawad and Paoli, 2002). Additionally, each blade exposes on its edge (typically on strand D) a stretch of ~ 6-8 residues available for intermolecular β-sheet formation, which can be likened to one half of a zipper. In Nup120, 5 of the 7 blades are exposed this way, two are buried in the hydrophobic core shared with the attached α-helical domain. In addition to these interactions being relatively weak, another inherent difficulty in identifying them is that they are likely very

poorly conserved at the sequence level because the contacts are mediated via the backbone rather than side chains. Based on the available data, it is conceivable that the Nup120 β-propeller is involved in inter-Y complex contacts. It is also possible that it is used to bridge to the Nic96 complex, but we can also not exclude that it may be an anchor for dynamic nucleoporins or other accessory proteins. The relatively mild nup120Δ phenotype (Figure 3.5) compared to nup133Δ or nup84Δ and the behavior of Nup120$_{1-757}$-GFP suggests that if the Nup120 β-propeller has an integral role in the NPC scaffold, it is either redundant or can be functionally replaced by another nucleoporin.

In summary, we show that Nup120 adopts a unique architecture to build one of the two arms of the multimeric Y complex, the linchpin of the NPC scaffold. The atomic structure of the universally conserved heptameric core of the Y complex is now nearing completion. With reliable data on inter-subcomplex contacts the construction of a basic NPC architecture is within reach in the close future.

Methods

Protein expression and purification

Nup120 from *S. cerevisiae* (residues 1-757 of 1037) was expressed at 18°C in E. coli strain BL21(DE3)-RIL as a 6xHis N-terminal fusion protein from a pET-Duet-derived plasmid. Cells were pelleted and resuspended in lysis buffer (50mM potassium phosphate pH 8.0, 500mM NaCl, 40mM imidazole, 5mM β-mercaptoethanol). Cells were lysed using a French press and the cleared lysate incubated in batch with Ni-affinity resin. After washing the resin in batch with lysis buffer, the protein was eluted with lysis buffer containing 250mM imidazole. After cleavage of the purification tag, Nup120 was subjected to size exclusion chromatography on Superdex S200 equilibrated in 10mM Tris/HCl pH 8.0, 150mM NaCl, 0.1mM EDTA, and 1mM DTT. Nup120 eluted as a monomer of 88 kDa. Selenomethionine-derivatized protein was prepared as previously

described (Brohawn et al., 2008) and Nup120-SeMet was purified identically to the native version.

Full length Nup85 in complex with Seh1 and a single chain version of full-length Nup145C in complex with Sec13 from *S. cerevisiae* were cloned as described (Brohawn et al., 2008), purified as for Nup120 (residues 1-757), and are referred to in the text as Nup85•Seh1 and Nup145C•Sec13. The C-terminal helical domain of Nup120 (residues 766-1037) was generated from a full length Nup120 construct by PCR. A 5-protein complex of Nup120 (residues 766-1037), Nup85•Seh1, and Nup145C•Sec13 was prepared by co-expression of a trimeric complex of Nup120 (residues 766-1037)•Nup85•Seh1 (Brohawn et al., 2008) and the single chain version of Nup145C•Sec13 in BL21(DE3)-(RIL) cells and was purified as for Nup120 (residues 1-757). The Ni-NTA elution was pooled, digested with human rhinovirus 3C to remove fusion tags, and subjected to size exclusion chromatography using a Superdex S200 26/60 column equilibrated in 10mM Tris/HCl pH 8.0, 250mM NaCl, 1mM DTT, and 0.1mM EDTA.

Protein crystallization

Nup120 concentrated to 20 mg/ml was crystallized in 15% (w/v) PEG 3350, and 0.1M Tris/HCl pH 7.5, 0.2M KSCN by the hanging drop vapor diffusion method at 18°C in 2μl drops. Crystals grew within 3-6 days forming rhomboid prisms with dimensions of 60μm x 60μm x 20μm. The selenomethionine derivative crystallized in the same condition, while the highest quality crystals were obtained by microseeding with native crystals. Both native and derivative crystals were cryo-protected by serial transfer of the crystals into reservoir solutions supplemented with increasing amounts of PEG200 (10%-25% (v/v), 5% steps) before flash freezing in liquid nitrogen. Both native and derivative protein crystallized in space group $P2_12_12$ with one molecule per asymmetric unit. Data was collected at beamline 24ID-C at Argonne National Laboratory.

Structure determination

Although the native crystals were larger and optically superior, the selenomethionine-derivatized crystals diffracted significantly better and were exclusively used for data analysis. A complete dataset was collected at the Se-Peak wavelength and data reduction was carried out using the HKL2000 package (Otwinowski and Minor, 1997). All 9 selenium sites were found using SHELXD (Sheldrick, 2008). After refinement of the Se positions and density modification with SHARP, an excellent experimental electron density map was obtained, allowing for the assignment and building of the majority of the structure. Sequence assignment was aided by using the selenium positions as markers. Model building was done with Coot (Emsley and Cowtan, 2004) and refinement was carried out using the PHENIX suite (Adams et al., 2002). The model is complete except for residues 31-52, 188-200, 303-313, and 731-757 for which only spurious electron density was observed. Blades 3 and 4 of the β-propeller have the highest temperature factors and are not as well packed as the remainder of the molecule.

Analytical size exclusion chromatography

For Nup120 (residues 1-757), Nup145C•Sec13, and Nup85•Seh1 binding experiments, equimolar amounts of each component were incubated alone or in combination for 30 minutes at 4°C in binding buffer (10mM Tris/HCl pH 8.0, 250mM NaCl, 1mM DTT, 0.1mM EDTA). Reactions were injected onto a Superdex S200 HR 10/300 column (GE Healthcare) equilibrated in binding buffer, and run at a flow rate of 0.8 ml/min (Figure 3.4A).

Yeast strain construction

Deletion strains were taken from the Yeast Deletion Consortium (Winzeler et al., 1999), C-terminal GFP-tagging was done by homologous recombination in a BY4741 background, using pFA6a-GFP(S65T)-kanMX6 as template for C-terminal modifications (Longtine et al., 1998). Strains were selected on G418 plates (200 µg/ml) and verified by PCR.

Fluorescence microscopy

Strains were grown overnight in YPD (1% yeast extract, 2% yeast peptone, 2% glucose) at 30°C, diluted 20-fold into fresh YPD, and grown for 4-5 hours at 30°C to OD600 ~0.5. Cells were harvested by centrifugation, fixed for 3 minutes in 3.7% formaldehyde/0.1M potassium phosphate pH 6.5, and prepared for immunofluorescence as previously described (Kilmartin and Adams, 1984). Samples were incubated with mAb414 (abcam, 1:1000) alone or in combination with goat anti-GFP (1:500) for 90 minutes at room temperature. Bound antibodies were detected by incubation with Cy5-conjugated anti-mouse (1:500) alone or in combination with Cy2-conjugated donkey anti-goat (1:200) for 45 minutes at room temperature. DNA was stained with 0.05 µg/ml 4',6-diamidino-2-phenylindole (Sigma-Aldrich) and samples were mounted for imaging in 1 mg/ml p-phenylenediamine and 90% glycerol. Fluorescence microscopy was performed on a Zeiss AxioImager.Z1 microscope and images were taken with a Zeiss AxioCam HRm camera.

Accession number

The atomic coordinates of the Nup120 structure described here will be deposited with the Protein Data Bank (PDB) under the accession code 3HXR.

Chapter four

Molecular architecture of the Nup84•Nup145C•Sec13 edge element in the nuclear pore complex lattice

This chapter was previously published as Brohawn, S.G. & Schwartz, T.U. Molecular architecture of the Nup84•Nup145C•Sec13 edge element in the nuclear pore complex lattice. Nature Structural & Molecular Biology 16, 1173-1177 (2009).

S.G.B. designed, conducted, and analyzed biochemical, biophysical, and crystallographic experiments and wrote the manuscript; T.U.S. advised on all aspects of the project and wrote the manuscript.

Abstract

Nuclear pore complexes (NPCs) facilitate all nucleocytoplasmic transport. These massive protein assemblies are modular, with a stable structural scaffold supporting more dynamically attached components. The scaffold is made from multiple copies of the heptameric Y complex and the heteromeric Nic96 complex. We previously showed that members of these core subcomplexes specifically share an ACE1 fold with Sec31 of the COPII vesicle coat, and we proposed a lattice model for the NPC based on this commonality. Here we present the crystal structure of the heterotrimeric 134-kDa complex of Nup84•Nup145C•Sec13 of the Y complex. The heterotypic ACE1 interaction of Nup84 and Nup145C is analogous to the homotypic ACE1 interaction of Sec31 that forms COPII lattice edge elements and is inconsistent with the alternative 'fence-like' NPC model. We construct a molecular model of the Y complex and compare the architectural principles of COPII and NPC lattices.

Introduction

In eukaryotic cells, physical separation of the nucleus and cytoplasm by the nuclear envelope necessitates a conduit for nucleocytoplasmic molecular traffic. This gateway is solely provided by NPCs, proteinaceous channels that stud the nuclear envelope (Brohawn et al., 2009; D'Angelo and Hetzer, 2008; Tran and Wente, 2006; Weis, 2003). NPCs are among the largest assemblies in the cell, at ~50 MDa, and have crucial roles in cellular homeostasis. The overall structure of the NPC is generally conserved across species (Elad et al., 2009). NPCs have a central scaffold ~30–50 nm in height, with approximately eight-fold rotational symmetry about the transport axis, an outer diameter of ~100 nm, a central transport channel ~40 nm in diameter, and attached cytoplasmic and nucleoplasmic filaments (Alber et al., 2007b; Beck et al., 2007; Stoffler et al., 2003). The NPC is a modular structure composed of multiple copies of ~30 proteins (nucleoporins, or Nups) arranged into distinct subcomplexes (Brohawn et al., 2009; Rout et al., 2000; Schwartz, 2005). It is also dynamic, with components possessing widely ranging resident times (Dultz et al., 2008; Rabut

et al., 2004). The most stable nucleoporins form the structural scaffold and are largely organized in the heteromeric Nic96 subcomplex and the heptameric Y complex. The scaffold connects to the nuclear envelope through interaction with transmembrane nucleoporins; it anchors phenylalanine-glycine repeat-containing nucleoporins that form an extended meshwork projecting into the central pore channel that constitutes the main transport barrier (Frey and Gorlich, 2007; Rout et al., 2003).

Considering its central role in transport and other cellular processes, a high-resolution structure of the NPC is much sought after. Significant effort has focused on elucidating the central scaffold architecture. Its fundamental importance is evident, as assembly defects and transport deficiencies result from knockout or knockdown of scaffold components (Boehmer et al., 2003; Fabre and Hurt, 1997; Galy et al., 2003; Harel et al., 2003; Makio et al., 2009; Walther et al., 2003). Still, the arrangement of the Nic96 subcomplex (including Nic96, Nup53, Nup59, Nup157, Nup170, Nup188 and Nup192; nomenclature is from *Saccharomyces cerevisiae* unless noted) remains largely enigmatic. However, the structures of major portions of Nic96, Nup170 and Nup53/59 are now available and combined with functional data will help narrow down the problem (Handa et al., 2006; Jeudy and Schwartz, 2007; Onischenko et al., 2009; Schrader et al., 2008b; Whittle and Schwartz, 2009). Organization of the Y complex (composed of Nup133, Nup84, Nup145C, Sec13, Nup120, Nup85 and Seh1) is better understood. The complex is tightly associated and forms a Y, as observed by EM, with two short arms and a long kinked arm connected at a central hub (Harel et al., 2003; Kampmann and Blobel, 2009; Lutzmann et al., 2002; Siniossoglou et al., 2000; Walther et al., 2003). High-affinity connections within the Y complex involve binary interactions between α-helical domains of its constituents (Boehmer et al., 2008; Brohawn et al., 2008; Lutzmann et al., 2002). The long arm terminates with a flexibly attached N-terminal β-propeller of Nup133 followed by an irregular α-helical stack domain that interacts end-to-end with Nup84 (Berke et al., 2004; Boehmer et al., 2008; Whittle and Schwartz,

2009). Nup84 in turn interacts with Nup145C. Nup120 and Nup85 form the short arms of the Y complex, with the C-terminal region of Nup120 forming the central hub that interacts with the C-terminal tail modules of Nup145C and Nup85 (Leksa et al., 2009). Nup145C and Nup85 bind the related β-propeller proteins Sec13 and Seh1, respectively, by addition of an N-terminal insertion blade to complete the open six-bladed β-propellers *in trans* (Brohawn et al., 2008; Debler et al., 2008; Hsia et al., 2007). Although crystallographic data on single proteins and some binary complexes are available, we are still lacking a detailed structural description of the entire Y complex, notably including all domain-domain and protein-protein interfaces.

A common evolutionary origin of the NPC and vesicle coats has been proposed based on their shared role in stabilizing curved membranes and predicted similarities in fold composition of constituent proteins (Devos et al., 2004). Structural evidence of a common ancestor has been obtained, as the nucleoporins Nic96, Nup85, Nup145C and Nup84 are homologous to the COPII vesicle coatomer Sec31 and together constitute the unique fold class ACE1 (ancestral coatomer element 1; (Brohawn et al., 2008)). In the COPII vesicle coat lattice, two molecules of the ACE1 protein Sec31 interact to form edge elements, whereas the β-propellers of Sec31 and Sec13 interact to form vertex elements (Fath et al., 2007; Stagg et al., 2008). We proposed that the NPC structural scaffold forms a similar lattice-like coat for the nuclear envelope (Brohawn et al., 2008; Brohawn and Schwartz, 2009a). However, the absence of structural knowledge of ACE1 organization in the NPC precluded any direct comparison of the NPC and COPII lattices.

We set out to determine the molecular architecture of an edge element in the NPC lattice and here present the 4.0-Å crystal structure of the heterotrimeric Nup84•Nup145C•Sec13 unit of the Y complex of the NPC from *S. cerevisiae*. The ACE1 interaction between Nup84 and Nup145C is architecturally related to the Sec31 edge element in the COPII lattice. As in the COPII coat, the edge

element in the NPC lattice is arranged in a manner consistent with its role in stabilizing membrane curvature at the nuclear envelope. We further present a composite atomic model of the Y complex, propose how it is arranged in the NPC, and compare the NPC lattice architecture to that of vesicle coats.

Results

Structure of Nup84•Nup145C•Sec13 trimeric complex

We solved the crystal structure of a trimeric complex between Nup84, Nup145C, and Sec13 from *S. cerevisiae* by multiple isomorphous replacement with anomalous scattering using selenomethionine and tantalum bromide derivatives (Figure 4.1 and Table 4.1). The crystallized construct includes $Nup84_{1-424}$, $Nup145C_{109-555}$ and full-length Sec13. Despite modest resolution and relatively high B factors, initial phase estimates were excellent and resulted in high-quality electron density maps (Figure 4.2). Still, confident model building at this resolution is aided by the availability of high-resolution models of fragments of the structure (Brunger et al., 2009). To this end, we crystallized a complex of the minimal interaction domain of $Nup145C_{109-179}$ in complex with full-length Sec13, solved the structure by molecular replacement at 2.6-Å resolution, and refined it to an R_{work}/R_{free} of 21.5%/25.3%. We placed this partial model and $Nup145C_{180-555}$ from the *S. cerevisiae-H. sapiens* hybrid Nup145C•Sec13 structure (Hsia et al., 2007) into the map using real-space methods. Phase improvement using the experimental data with the partial model resulted in clearly interpretable maps into which we built the final Nup84•Nup145C•Sec13 model and refined it to an R_{work}/R_{free} of 28.2%/32.9%.

Table 4.1 – Data collection and refinement statistics

	Nup145C$_{109-179}$•Sec13 Native	Nup84•Nup145C•Sec13 Selenomethionine	Nup84•Nup145C•Sec13 [Ta$_6$Br$_{12}$]$^{2+}$
Data collection			
Space group	P2$_1$2$_1$2	P6$_2$22	P6$_2$22
Cell dimensions			
a, b, c (Å)	68.3, 93.9, 55.0	170, 170, 271	170, 170, 270
α, β, γ (°)	90, 90, 90	90, 90, 120	90, 90, 120
Resolution (Å)	40-2.6 (2.66-2.6) *	50-4.0 (4.14-4.0)	50-4.4 (4.56-4.4)
R_{sym} (%)	12.6 (64.9)	17.2 (97.4)	17.0 (79.9)
$I/\sigma I$	12.4 (2.0)	10.6 (1.5)	8.6 (1.3)
Completeness (%)	97.9 (96.9)	99.9 (99.9)	92.2 (76.6)
Redundancy	3.5 (3.3)	5.9 (5.0)	3.7 (2.2)
Refinement			
Resolution (Å)	35-2.6	50-4.0	
No. reflections	11148	37016	
R_{work} / R_{free}	21.7 / 25.4	28.2 / 32.9	
No. atoms			
Protein	2621	8671	
Water	67	0	
B-factors (Å^2)			
Protein	71	189	
Water	70	n/a	
R.m.s deviations			
Bond lengths (Å)	0.003	0.004	
Bond angles (°)	0.679	0.762	

*Values in parentheses are for highest-resolution shell.

A.

B.

Nup84 **Nup145C** **Sec13**

C.

90°
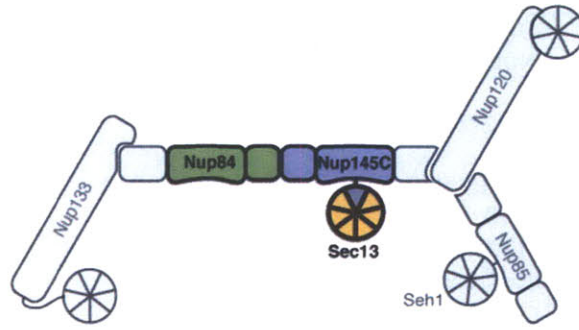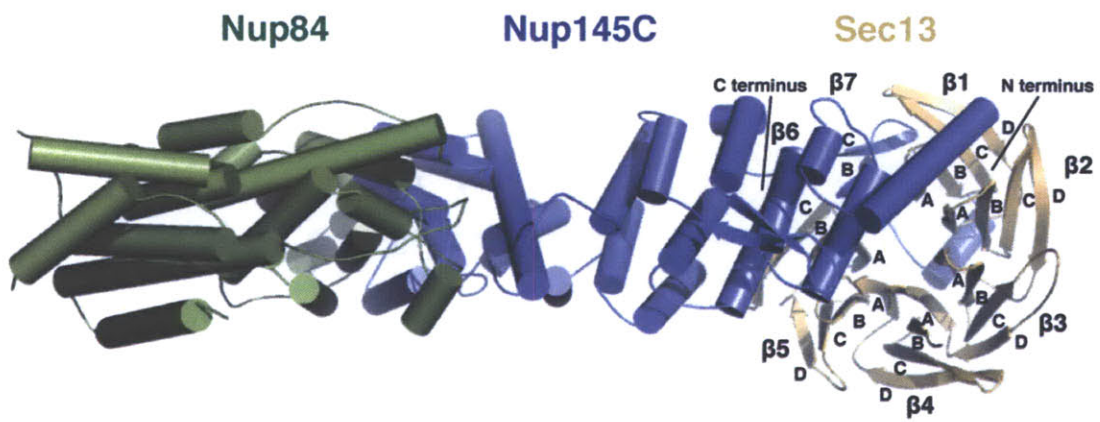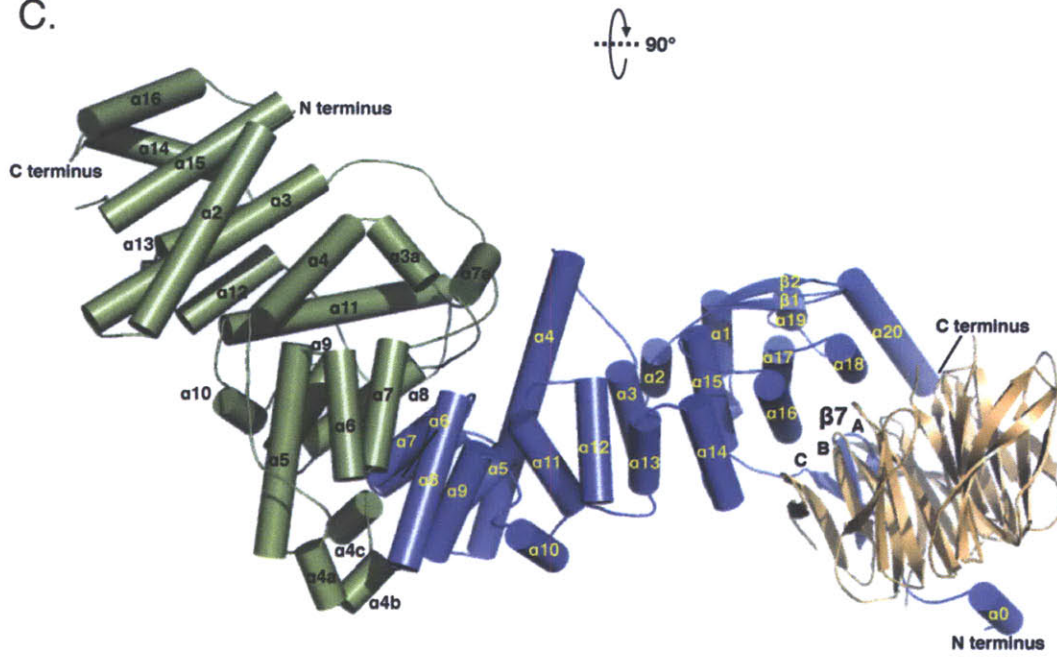
Figure 4.1 - Structure of Nup84•Nup145C•Sec13

(A) Y complex of the NPC. The crystallized trimeric segment is colored. (B) Overall structure of the heterotrimeric Nup84•Nup145C•Sec13 complex, with Nup84 in green, Nup145C in blue, and Sec13 in light orange. β-propeller composed of blades 1–6 from Sec13 and blade 7 from Nup145C is labeled. (C) Structure is rotated by 90°, with secondary structure elements of Nup84 and Nup145C labeled.

Nup84   Nup145C   Sec13
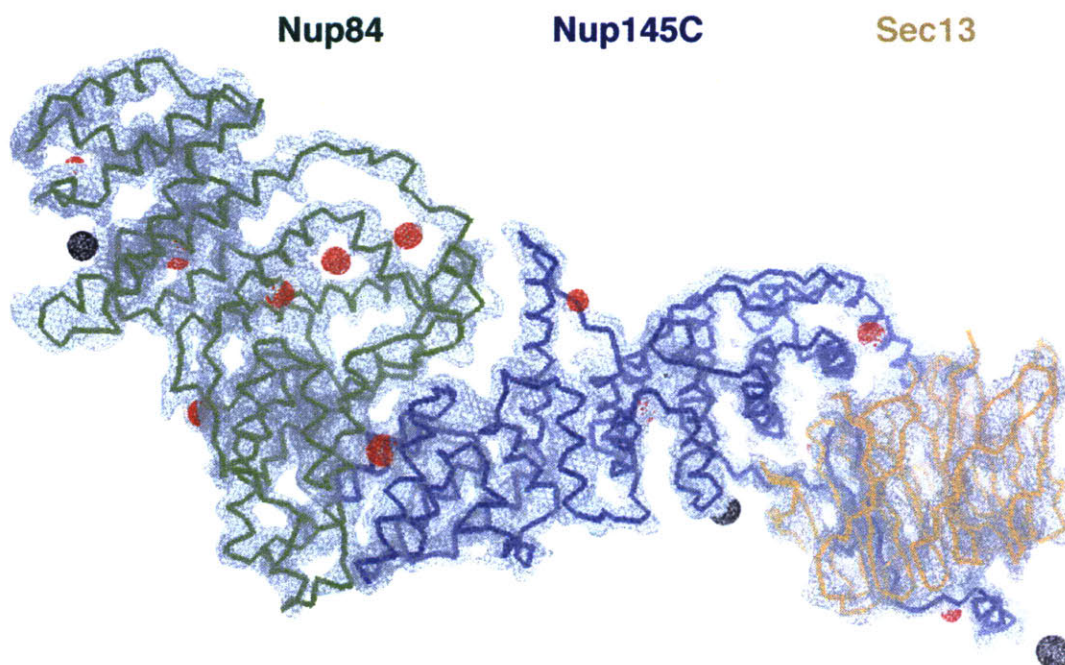
Figure 4.2 - Electron density for the Nup84•Nup145C•Sec13 model

Phase combined $2F_o$-F_c 4.0 Å electron density map contoured at 1.5 σ, shown in gray. Heavy atom electron density displayed for the 12 selenium sites (red, 8 σ) and 4 tantalum sites (black, 6 σ). The structure is shown as a Cα trace.

The trimeric complex has the approximate shape of a kinked rod, with dimensions ~150 × 30 × 30 Å (Figure 4.1). Nup84 and Nup145C both form α-helical blocks with dimensions ~65 × 30 × 30 Å that interact at the kink in the rod, creating a 2,040-Å^2 interface. The N terminus of Nup145C forms an insertion blade that completes the open six-bladed β-propeller of Sec13 *in trans*. The higher-resolution fragments both superimpose well with the same regions in the trimeric structure, with mostly minor deviations observed (see Methods). In comparison to the hybrid human Sec13•yeast Nup145C structure (Hsia et al., 2007), there is a ~10° rotation of the propeller unit about its central axis. This may be indicative of flexibility of the β-propeller unit relative to the ACE1 domain, which could be important in the assembly of the NPC lattice (Stagg et al., 2008) (see below).

ACE1 nucleoporins Nup84 and Nup145C interact crown to crown

As predicted by structural modeling, Nup84 adopts an ACE1 fold despite very low sequence homology to other ACE1 members (Brohawn et al., 2008) (Figure 4.3). ACE1 is a tripartite, J-shaped helical fold composed of three modules: crown, trunk, and tail.

The Nup84•Nup145C•Sec13 structure contains the trunk and crown modules of Nup145C and Nup84 (Figure 4.3). The two proteins form an extensive interface between their crown modules, with α6, α7 and α8 of Nup84 packing antiparallel to α6, α7 and α8 of Nup145C, completely burying α7 from each protein in the interface (Figure 4.4). The surfaces of helices α6, α7 and a8 in each protein are distinctly hydrophobic and highly conserved (Figure 4.5). Homodimerization of Sec31 is similarly accomplished by the antiparallel interaction of helices α6, α7 and α8 ((Fath et al., 2007); Figure 4.4). Notably, a domain swap between crown helices α5, α6 and α7 from each Sec31 monomer is observed in the crystal structure. Whether this is the physiologically relevant manner of interaction is

116

unclear, but a long loop that allows for the domain swap is conserved in length in Sec31. Regardless, the interaction likewise juxtaposes and buries α7 from each Sec31 molecule.
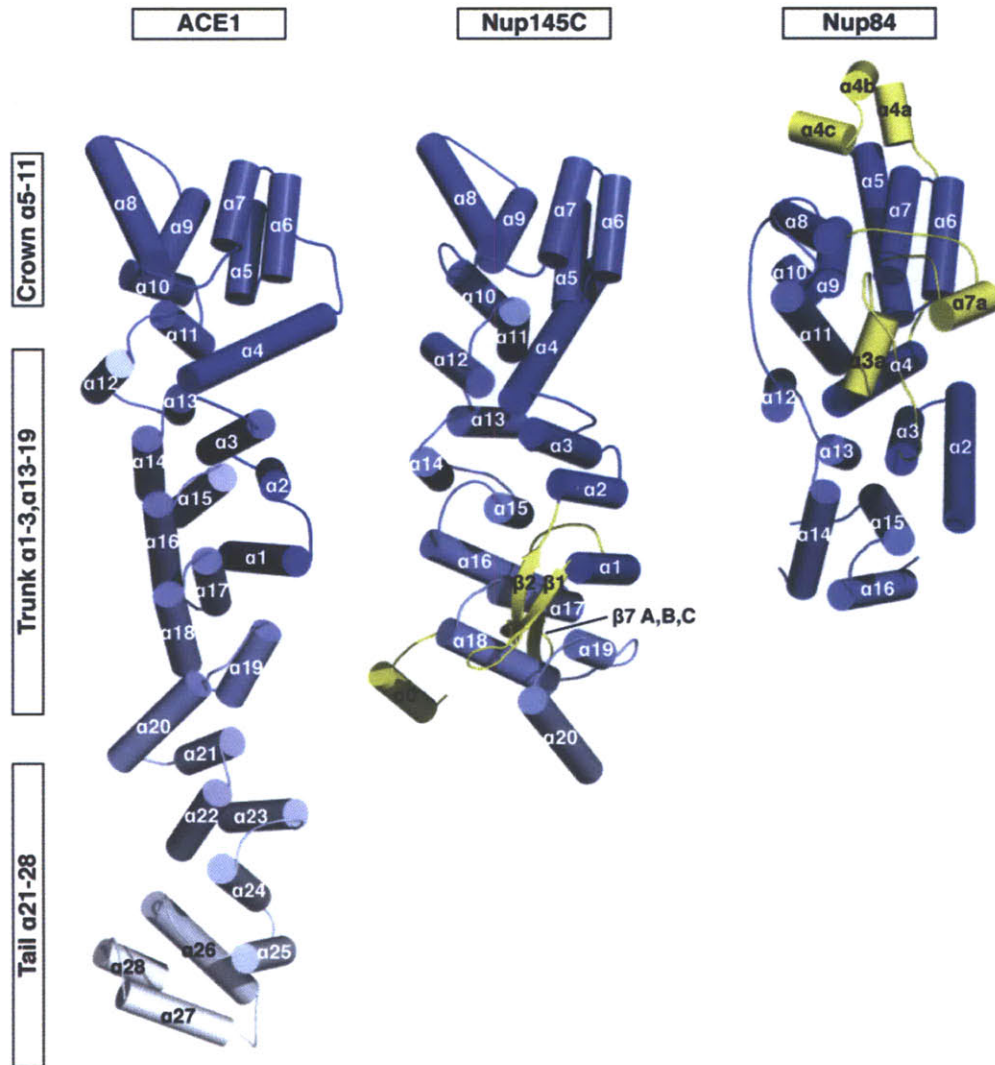
Figure 4.3 - Nup84 is an ACE1-containing protein
Structural alignment of the canonical ACE1 fold, Nup145C, and Nup84. The 28 conserved ACE1 helices are colored in a blue-white gradient from N to C terminus, protein-specific elaborations in Nup145C and Nup84 are in yellow. Nup145C contains an additional N-terminal helix, the Sec13-interacting insertion blade, and a β-hairpin following α1, Nup84 contains a few short 1-3 turn α-helices following α3, α4, and α7. The canonical ACE1 structure was generated for illustrative purposes by superimposing the crown, trunk, and tail modules of ACE1 structures individually and constructing the conserved helices and intervening loops in the most frequently observed position. The Nup84•Nup145C•Sec13 structure includes the crown and trunk modules of Nup145C and Nup84. The characteristic α-helical architecture of the ACE1 fold is clearly observed in Nup145C and Nup84.
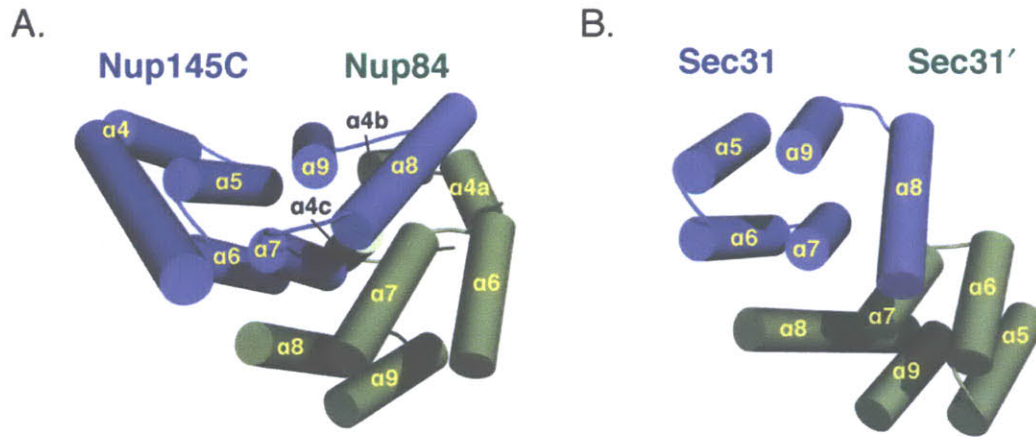
Figure 4.4 - The crown-crown interaction of Nup84•Nup145C is analogous to that of Sec31•Sec31

(A) ACE1 crown-crown interaction between Nup145C and Nup84. (B) The same interaction between two Sec31 molecules (PDB 2PM6; (Fath et al., 2007)). The Sec31 interaction is shown without domain swapping, and the rest of the molecules are removed for clarity (see Results). Analogous juxtaposition of crown helices α6, α7 and α8 is observed in both (A) and (B).

Figure 4.5 - The hydrophobic and highly conserved Nup84•Nup145C interface

The Nup84•Nup145C interface is shown in an "open-book" view relative to Figure 5.1C. In (A), the surface buried upon interaction is yellow and contributing helices and residues from each protein are labeled. Residues mutated on the surface of helix α7 in each protein that eliminate binding *in vitro* and *in vivo* (Brohawn et al., 2008) are labeled in red. The interface is highly complementary and largely hydrophobic. In (B), the buried surface of each protein is outlined in black and the protein surface is colored by conservation from white (not conserved) to orange

(highly conserved). Surface corresponding to the residues mutated to abolish binding is indicated by an inner black outline. A sequence alignment from budding yeast for each protein from α6-α8 is shown in (C) and colored as the surface in (B). A plot of buried surface area vs. residue is shown as a bar graph above the alignment. The mutations made are indicated above the appropriate residues (I206D/M210D in Nup84 and V321E/S324E/Y325A in Nup145C). The mutated residues are all among the largest contributors to buried surface area and most conserved in the interface.

Unique features of each ACE1 unit in the Nup84•Nup145C•Sec13 structure form additional interaction sites that frame the primary surface of α6, α7 and α8 (Figure 4.5). Two extended and conserved loops in Nup84 (between α3 and α4, and between α7 and α8) pack against the long and kinked helix α4 of Nup145C. On the opposite side of the α6, α7 and α8 interface, Nup84 has an insertion of three short helices between ACE1 helices α4 and α5 (α4a, α4b and α4c) that together form an interface with two crown loops (between α6 and α7, and between α8 and α9) of Nup145C.

Structural evidence for lattice model of NPC

By analogy to the Sec31 interaction in COPII coats, Nup84 and Nup145C have been predicted to interact through their crown modules (Brohawn et al., 2008). Point mutation resulting in surface residue alterations in Nup145C α7 (V321E, S324E and Y325A) and corresponding alterations in the then-predicted Nup84 helix α7 (I206D and M210D) abrogate high-affinity binding. The structure presented here definitively shows that the interaction between Nup84 and Nup145C occurs through ACE1 crown modules and allows the mutant data to be explained from a structural perspective. The altered sites on each protein are intimately involved in the interaction surface: I206 and M210 account for 11% of the total area of Nup84 buried (223 of 2,024 Å^2), whereas V321, S324 and Y325 form 12% of the total area of Nup145C buried (257 of 2,059 Å^2). Introduction of charged residues into, or loss of large side chains from, the hydrophobic and complementary interaction surface is highly destabilizing, resulting in specific disruption of the interaction. That these point mutations eliminate binding indicates that the α6, α7 and α8 surface is the primary binding determinant, and

121

the secondary framing interactions are insufficient to independently maintain interaction.

The Nup84•Nup145C•Sec13 structure presented here fully supports our lattice model for the NPC and provides conclusive evidence against the alternative 'fence-like' model, based primarily on crystal contacts observed in the yeast Nup145C•human Sec13 hybrid structure (Hsia et al., 2007). In that crystal, Nup145C•Sec13 units stack by homotypic crown-crown interaction of Nup145C. Superposition of the Nup145C interaction observed in that crystal with the Nup84•Nup145C interface reported here shows that formation of the two interfaces is mutually exclusive (Figure 4.6).

A.

Nup84

Nup145C

Sec13

**Nup84•Nup145C•Sec13
NPC edge element**

B.

Sec13'

Nup145C'

Nup145C

Sec13

**Nup145C•Sec13 / Nup145C'•Sec13'
Crystal packing interaction**
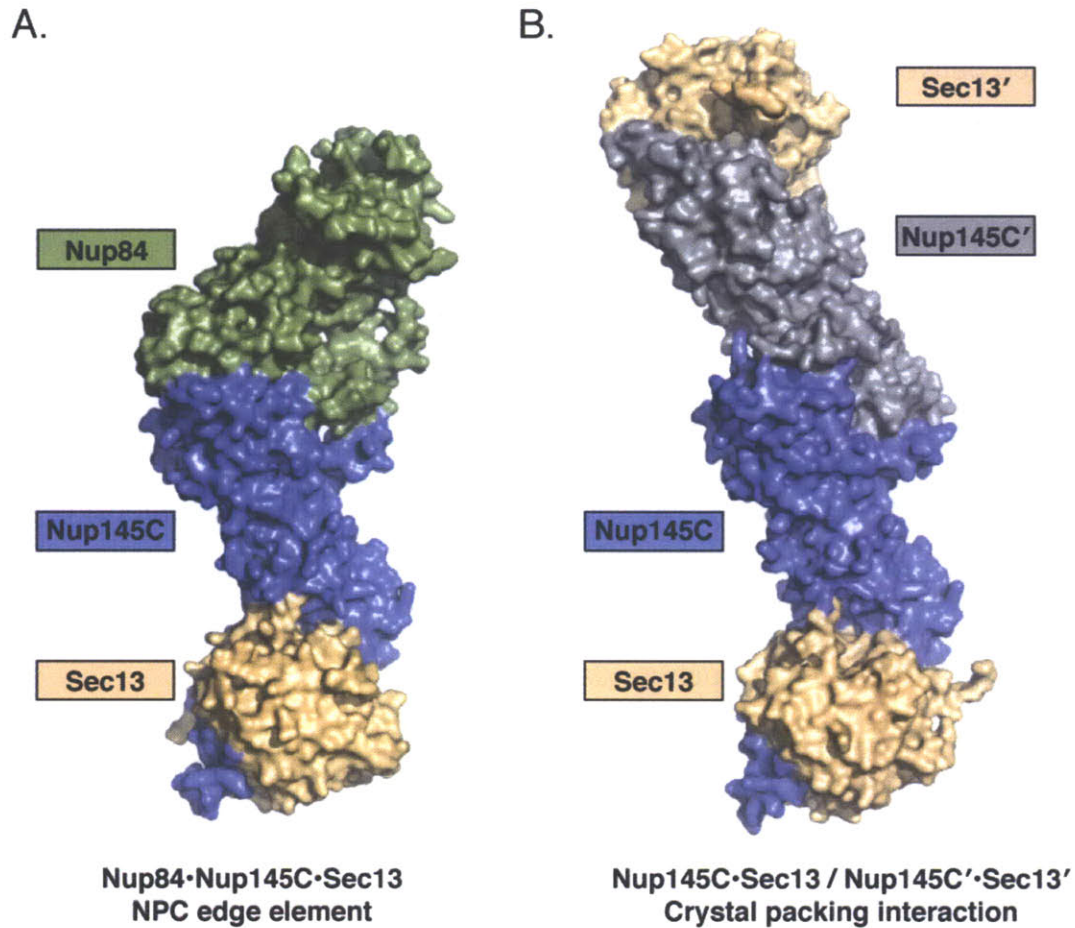
Figure 4.6 - The Nup84•Nup145C•Sec13 edge element sterically clashes with a proposed "fence-like" pole in the NPC
(A) Surface representation of the entire Nup84•Nup145C•Sec13 structure. (B) Two heterodimers of the yeast Nup145C•human Sec13 structure (Hsia et al., 2007) involved in a crystal contact are shown with the lower heterodimer aligned to (A). The view in (A,B) is rotated 90° clockwise from Figure 5.1C. The position of Nup84 in the physiologically relevant trimeric complex (A) sterically clashes with the position of the second Nup145C•Sec13 dimer related by crystal packing (B) thus providing conclusive evidence against this interaction as part of "pole" in the "fence-like" model of the NPC (Debler et al., 2008; Hsia et al., 2007).

Comparison of edge elements in NPC and COPII coat

The similarity between the Nup84•Nup145C•Sec13 structure and the

Sec13•Sec31 edge element in the COPII cage is immediately apparent (Figure

4.7), so we refer to the Nup84•Nup145C unit as an edge element in the NPC

lattice. The shared binding mode between crown modules in the two structures

123

results in analogous relative orientations of the interacting ACE1 units. The interface between Nup145C and Nup84 creates an angle of ~120° between ACE1 units. The interface between Sec31 molecules is ~165° in the crystal structure, though it was modeled to be ~135° by normal mode analysis for fitting into both COPII coat EM reconstructions (Fath et al., 2007; Stagg et al., 2008). A hinge movement about the crown-crown interface was thus postulated to be one mechanism that allows the coat to adapt its size to vesicles of different diameters (Fath et al., 2007). EM reconstructions of the Y complex have similarly shown plasticity in the angles of the long arm (Kampmann and Blobel, 2009). A similar hinge at the Nup84•Nup145C interface could be used in rearrangements of the NPC lattice in assembly and transport. Consistently, we observed hinge movement at the crown-crown interface in normal mode analysis of the Nup84•Nup145C•Sec13 structure (data not shown).

Nup84-Nup145C **Sec31-Sec31**
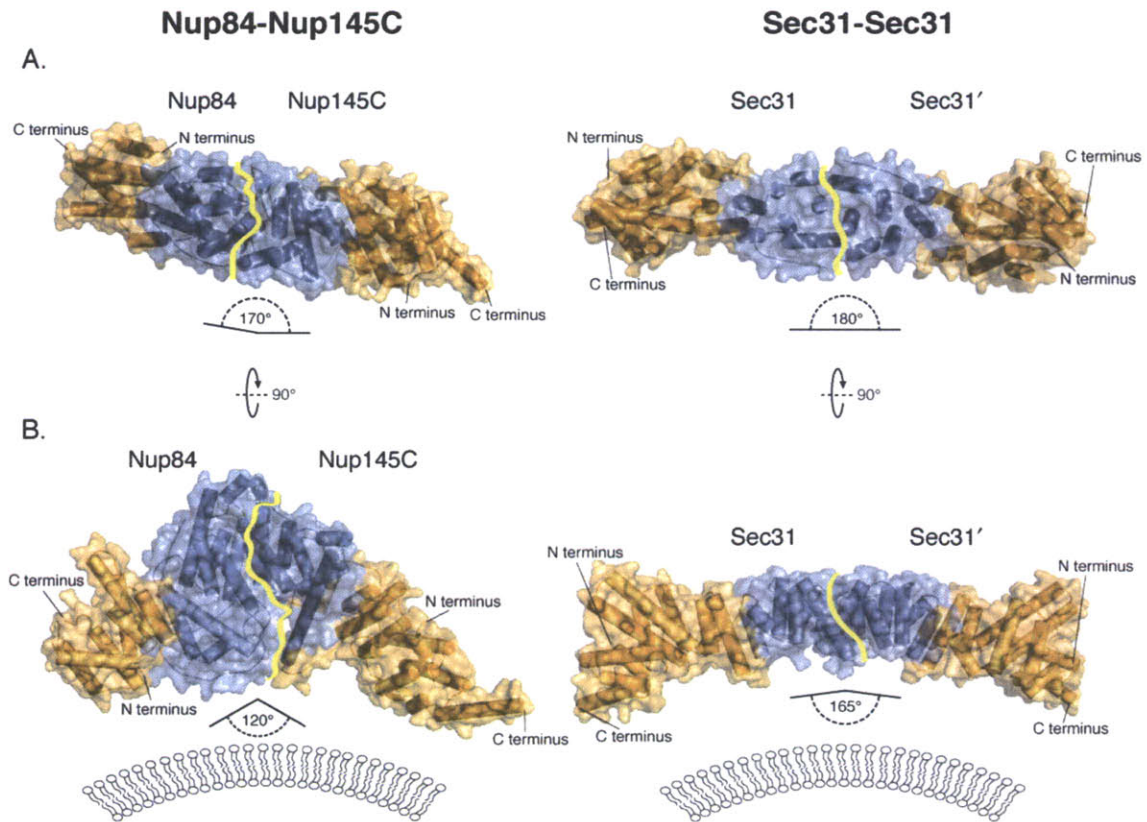
Figure 4.7 - Comparison of edge elements in the NPC and COPII lattices
(A,B) The lattice edge element Nup84•Nup145C in the NPC and Sec31•Sec31
(PDB 2PM6 (Fath et al., 2007)) in the COPII vesicle coat are shown as cartoons
in a half-transparent surface. The two ACE1 units in each edge element are
colored by module, with trunks orange and crowns blue. Yellow lines indicate the
interface between crown modules. The structures are shown from a top view in
(A) (rotated 180° from Figure 4.1) and a side view rotated by 90° in (B). The
analogous crown-crown interactions result in edge elements that share a
common architectural arrangement. Viewed from the top, the Nup84•Nup145C
edge element is bent ~10° from horizontal, whereas the Sec31•Sec31 edge is
essentially straight. Viewed from the side, the crystal structures of the edge
elements show markedly different angles, with the Nup84•Nup145C edge 45°
more acute than that of Sec31•Sec31. The angle observed in the
Nup84•Nup145C edge corresponds closely to the angle to which the
Sec31•Sec31 interface was modeled for fitting into the EM reconstructions of the
COPII cage and coat (Fath et al., 2007; Stagg et al., 2008). The proposed
position of the nuclear envelope membrane relative to the NPC edge element
shown in (B) is analogous to the known position of the COPII vesicle membrane
relative to the COPII edge element.

The insertion blade interaction between Sec13 and Nup145C or Sec31 is very similar in the two structures (data not shown). However, the different orientations of the insertion blade with respect to the ACE1 trunks result in Sec13 being positioned differently with respect to the edge elements. In the Sec31•Sec13 structure, the Sec13 propeller sits against the end of the Sec31 trunk, capping the edge element. In the Nup84•Nup145C•Sec13 structure, Sec13 is rotated ~45° forward toward the trunk and clockwise (viewed from the vertex) and rests on top of the Nup145C trunk. It remains to be determined whether additional interactions of Sec13 in the context of the entire NPC scaffold result in a conformational change from this position.

Discussion

The Nup84•Nup145C•Sec13 structure presented here, together with the previously reported structures of Nup85•Seh1, Nup84•Nup133 and Nup120 (Berke et al., 2004; Boehmer et al., 2008; Leksa et al., 2009) allows for the generation of a composite model for the majority of the Y complex at high resolution, including relative orientations of components in the long arm (Figure 4.8). The last four trunk helices of Nup84 need to be modeled to connect the Nup84•Nup145C•Sec13 and Nup84•Nup133 crystal structures. As these helices adopt identical topologies in other ACE1 structures and are predicted to be the only secondary structure elements present in Nup84 in this region, we can model their structure with high confidence (data not shown). This allows us to place the tail of Nup84 interacting with the full helical region of Nup133 relative to Nup84•Nup145C•Sec13. The position of the β-propeller of Nup133 is unknown and is probably flexible (Berke et al., 2004). The tail modules of both Nup145C and Nup85 can be confidently modeled (Brohawn et al., 2008), but the C-terminal interaction domain of Nup120 cannot and is not shown. Although the relative positions of the short arms with respect to the long arm of the Y complex cannot be assigned unambiguously, we have chosen to model the β-propellers of Sec13 and Seh1 in close proximity to one another by analogy to the interactions of β-propellers at the vertex elements in the COPII coat (Fath et al., 2007; Stagg et

al., 2008). Our positioning of the short arms is most consistent with all available data, although we cannot currently exclude alternative arrangements.

Our model is generally consistent with the recently reported EM reconstruction of the Y complex from yeast (Kampmann and Blobel, 2009). The angles of the Nup84•Nup145C and Nup84•Nup133 interfaces in our model correspond to those found in the highest-frequency EM class average. Here we incorporate ~0.5 MDa (of 0.58 MDa) of atomic models into a composite Y complex model. Most notably, the connecting Nup84•Nup145C•Sec13 structure allows for the incorporation of relative orientations of the proteins into the Y complex model. Analysis of the Y complex model reveals a number of functionally important implications.

Given the high degree of conservation between edge element structures in the NPC and COPII lattices, we predict the same inner concave surface of the edge element will face the membrane in the NPC (Figures 4.7 & 4.8). In this orientation, the N-terminal β-propeller–α-helical domain of Nup120 and the C-terminal α-helical domain of Nup133 point toward the membrane. These domains could serve as attachment sites for additional nucleoporins that could connect the Y complex to the membrane-proximal and membrane-spanning nucleoporins. Consistently, Nup120 interacts *in vitro* with Nup157, a member of the Nic96 subcomplex that can provide a link to transmembrane nucleoporins (Lutzmann et al., 2005; Onischenko et al., 2009). The ACE1-containing Nup85 is positioned away from the membrane, where it may form interactions to propagate the NPC lattice. The N terminus of Nup145C is also oriented away from the membrane, allowing its binding partner Nup145N to project its phenylalanine-glycine repeats into the pore channel (Ratner et al., 2007).
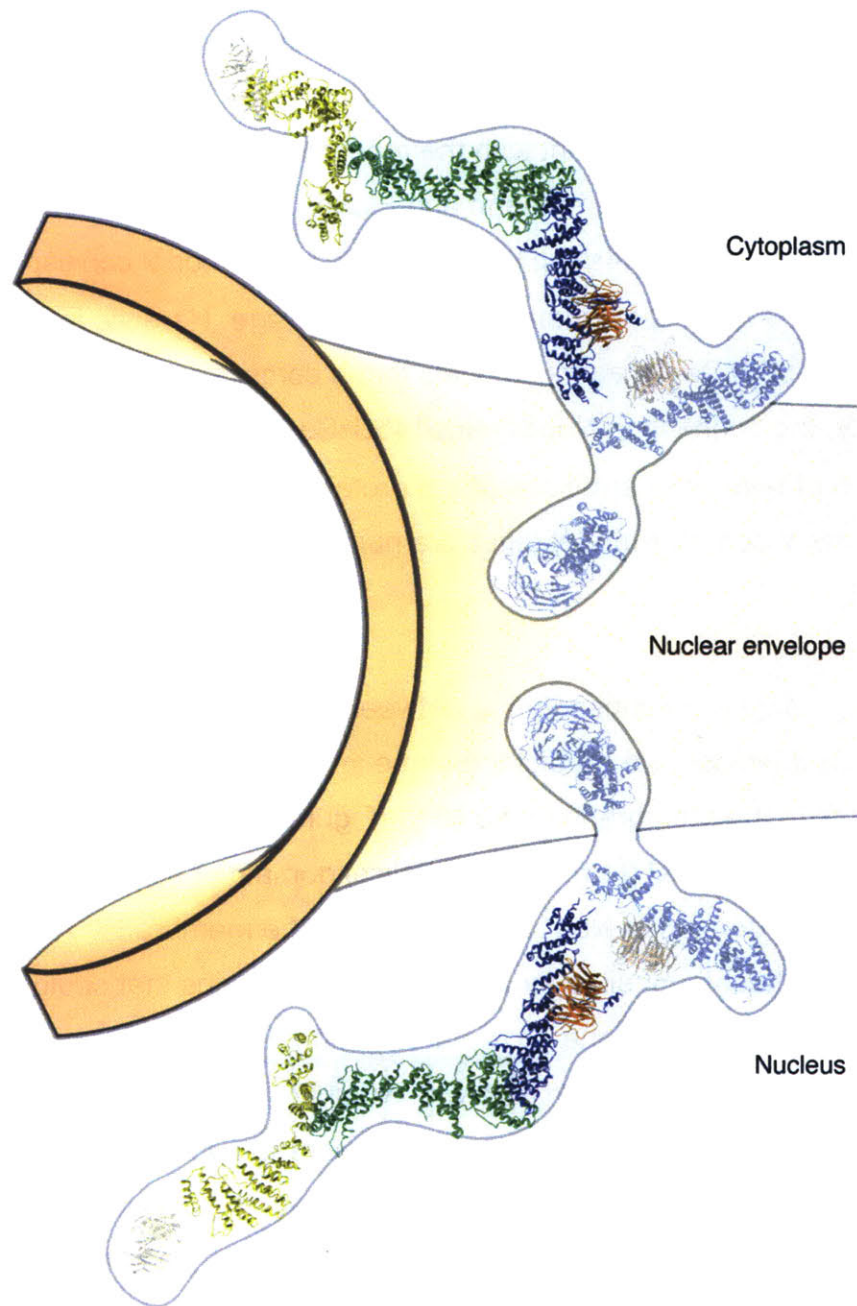
Figure 4.8 - Nup84•Nup145C is a membrane curvature-stabilizing edge element in the NPC lattice

A composite atomic model for the Y complex of the NPC, emphasizing the role of the Nup84•Nup145C edge element as a membrane curvature-stabilizing unit analogous to the Sec31•Sec31 edge element in COPII vesicle coats. The long arm of the Y complex is a composite model from crystal structures and is shown with Nup145C in blue, Sec13 in orange, Nup84 in green and Nup133 in yellow. The relative position of the N-terminal propeller of Nup133 (yellow) and the short arm components Nup120 (blue) and Nup85•Seh1 (blue–orange) are more tentatively placed and shown half-transparent (see Results for details). The long

axis of the Y complex is oriented along the positively curved nuclear envelope membrane, with the concave face of the Nup84•Nup145C edge element facing the lipid bilayer. This orientation is analogous to that of the Sec31•Sec31 edge element in the COPII coat and is consistent with the evolutionary relationship between the NPC and COPII vesicle coat lattices. Notably, although the Y complex is shown facing the membrane, it is not predicted to directly contact the nuclear envelope. Rather, other nucleoporins are predicted to have roles that correspond to adaptor complexes in other vesicle coating systems that link the membrane curvature-stabilizing coat (the Y complex) to the nuclear envelope.

The branch point in the Y complex has the β-propeller proteins Sec13 and Seh1 available to generate potential vertex interactions in the NPC lattice, similar to the Sec31 β-propeller vertex interactions in COPII coats. In contrast, at the opposite side of the NPC edge element, Nup84 (unlike Nup145C, Nup85 and Sec31) does not interact with a β-propeller. The loss of a β-propeller, combined with the acquisition of the Nup133 'cap', might have evolved as a way to terminate lattice propagation in this direction of the Y complex long arm. The utility of this type of arrangement is unique to the cytoplasmic- and nucleoplasmic-facing sides of a NPC lattice, as it cannot form self-enclosed structures observed in vesicle coats.

Our model is consistent with a role for the NPC edge element in stabilizing membrane curvature. In other membrane coating systems, proteins that directly contact membranes display a positively charged surface for electrostatic interactions with membrane phospholipids (McMahon and Gallop, 2005; Shibata et al., 2009; Zimmerberg and Kozlov, 2006). Like clathrin and Sec31•Sec13, the NPC edge element does not display such a surface (Figure 4.9) and is likely to coat to stabilize, but not directly interact with, curved membranes.

To date, the NPC has been shown to be architecturally related only to COPII coats; its relationship to clathrin coats (Fotin et al., 2004) is limited to a shared fold composition of components (Devos et al., 2006). Notably, an ALPS motif in human Nup133 has been shown to associate with membranes and has been suggested to initiate membrane curvature (Drin et al., 2007), though it has not been found in *S. cerevisiae*. This site is far enough removed from the ACE1 edge

element that the Y complex could have both roles: a curvature initiator at the distal end of the long arm, and a lattice-integral stabilizer at the ACE1 edge element.

We favor a model in which the membrane-facing edge element of the Y complex is oriented parallel to the transport axis and serves to stabilize the positive membrane curvature of the nuclear envelope, consistent with the evolutionary relationship with the COPII edge element that stabilizes positive vesicle membrane curvature. As more high-resolution structures of components are solved, they could be integrated to generate a more precise overall NPC structure. Fundamental to this goal will be the elucidation of potential vertex and inter-subcomplex interactions in the NPC lattice.
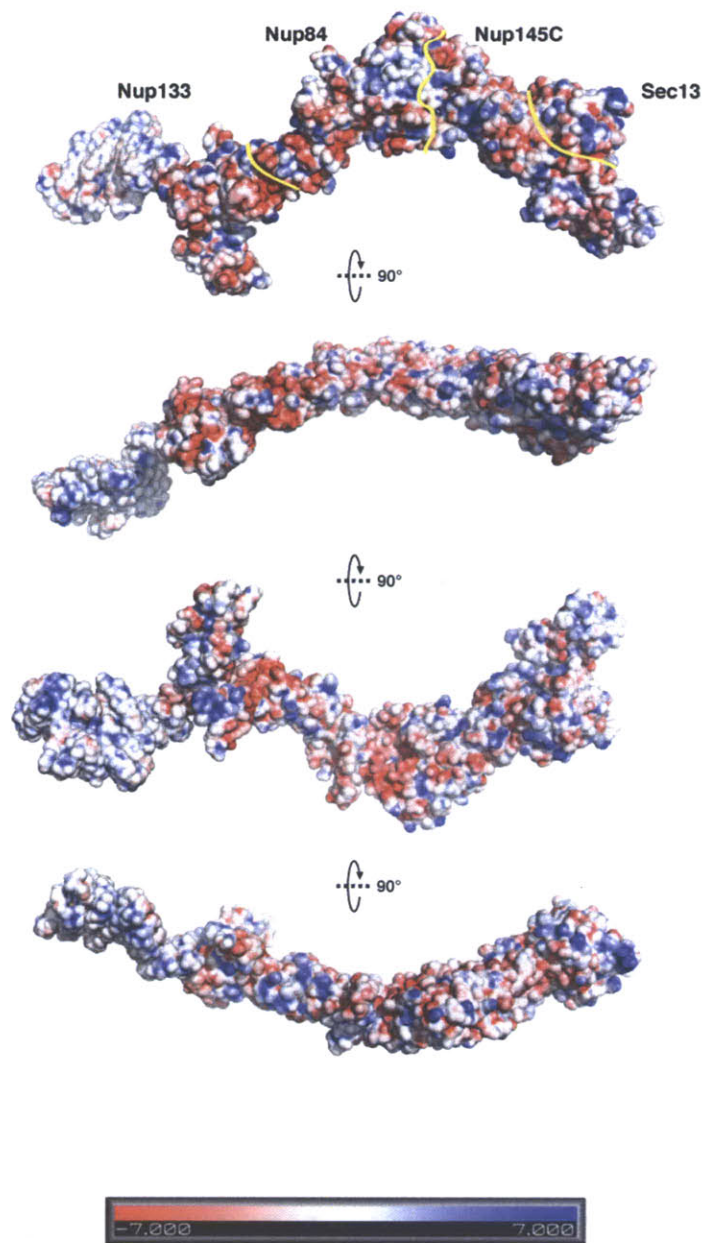
Figure 4.9 - The Y complex does not display a positively charged surface characteristic of membrane interaction

The surface of a composite structure of the long arm of the Y complex (Nup133•Nup84•Nup15C•Sec13) is shown colored according to electrostatic potential from red (-7 kT/e) to blue (+7 kT/e) in four views consecutively rotated 90°. The top view corresponds to the view in Figure 4.7B. The second view is of the concave face of the Nup84•Nup145C edge element predicted to face the membrane by analogy to the orientation of the Sec31•Sec31 edge element in the COPII lattice.

Methods

Construct generation

We cloned the trimeric complex of Nup84 (residues 1-424), Nup145C (109-555) and Sec13 from *S. cerevisiae* into a bicistronic bacterial expression vector. $Nup84_{1-424}$ was N-terminally fused with a cleavable 6xHis-tag. $Nup145C_{109-555}$ was C-terminally fused to Sec13 with a flexible 9-residue linker, to increase complex stability, without affecting chromatic behavior compared to the separate chain complex (data not shown, (Brohawn et al., 2008)). The trimeric complex is referred to as Nup84•Nup145C•Sec13 for simplicity. The completed β-propeller construct of Sec13 was generated by fusing the insertion blade of Nup145C (residues 109-179) C-terminally to full-length Sec13 via a flexible 9-residue linker. Sec13 was N-terminally fused with a cleavable 6xHis-tag.

Protein production and purification

Proteins were expressed in *E. coli* BL21-RIL(DE3) cells and purified as described (Brohawn, Leksa et al. 2008). Eluted protein was dialyzed against 50 mM Hepes-NaOH pH 7.4, 200 mM NaCl, 1 mM DTT, and 0.1 mM EDTA, the tag cleaved with protease overnight, and purified on a HiTrapS column (GE Healthcare) via a linear NaCl gradient followed by size exclusion chromatography. $Nup145C_{109-179}$•Sec13 was purified using a Superdex S75 26/60 column (GE Healthcare) run in 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT, and 0.1 mM EDTA. Nup84•Nup145C•Sec13 was purified using a Superdex S200 26/60 column (GE Healthcare) run in 10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT, and 0.1 mM EDTA. Selenomethionine-derivatized Nup84•Nup145C•Sec13 was prepared as described (Brohawn et al., 2008) and purified as the native version with reducing agent concentration at 5 mM in all buffers.

Crystallization

Small crystals of $Nup145C_{109-179}$•Sec13 grew in hanging drops of 0.5 μl protein at 85 mg ml^{-1} and 0.5 μl precipitant (0.1 M Tris-HCl pH 8.3, 26.5% (w/v) PEG 4000,

0.25 M LiCl) at 16°C in three days and were processed for seeds. Diffraction quality crystals grew as large plates (300 x 300 x 10 μm) in hanging drops of 0.2 μl seed dilution, 0.5 μl protein at 38 mg ml^{-1}, and 0.5 μl precipitant (0.1M Tris-HCl pH 8.3, 22% (w/v) PEG 4000, 0.25 M LiCl) at 16°C in three days. Crystals were cryoprotected by briefly soaking in precipitant with 25% v/v glycerol and flash frozen in liquid nitrogen.

Diffraction quality crystals of Nup84•Nup145C•Sec13 grew as half-cylinders 25-50 μm in height with a radius of 50-200 μm in hanging drops of 0.25 μl protein at 22.5 mg ml^{-1} and 0.25 μl precipitant (1.15 M sodium malonate pH 5.7) at 22°C in 2 days. Selenomethionine derivatized protein crystallized under identical conditions. $[Ta_6Br_{12}]^{2+}$-derivatized crystals were obtained by transferring crystals into a 0.5 μl drop of 1.17 M sodium malonate pH 5.7 and 200 μM $[Ta_6Br_{12}]^{2+}$ x 2Br$^-$ (Jena Biosciences) and incubating for 1-2 hours. Crystals were cryoprotected by briefly soaking in precipitant with 22.5% v/v ethylene glycol and flash frozen in liquid nitrogen.

Data collection and structure determination

iMosflm (Leslie, 1992) was used for data collection strategies, HKL2000 (Otwinowski and Minor, 1997) was used to reduce data, and model building and refinement were carried out with Coot (Emsley and Cowtan, 2004) and Phenix (Adams et al., 2002).

A 20 μm aperture beam was used to collect data from separate spot regions of the Nup145C$_{109-179}$•Sec13 crystals because diffraction quality varied over their volume. Molecular replacement was accomplished with Phaser (McCoy et al., 2007) using Sec13 (PDB ID 2PM6 (Fath et al., 2007)) as a search model. The final model is missing the first two residues of Sec13, as well as loop residues 158-167. The final model has Ramachandran plot values of 95.4% favored, 4.3% allowed, and 0.3% outliers.

The structure of Nup84•Nup145C•Sec13 was solved with multiple isomorphous replacement with anomalous scattering (MIRAS) using selenomethionine and $[Ta_6Br_{12}]^{2+}$ derivatives. 12 selenium sites (out of 13) and 4 $[Ta_6Br_{12}]^{2+}$ sites were found with SHELXC/D/E (Sheldrick, 2008) and refined in SHARP (Vonrhein et al., 2007). $Nup145C_{109-179}$•Sec13 (this work) and $Nup145C_{180-555}$ (PDB accession code 3BG1 (Hsia et al., 2007)) were placed into the solvent-flattened map from SHARP with BrutePTF (Strokopytov et al., 2005). PhaserEP (McCoy et al., 2007) was used to refine selenium sites and the partial model. Discussion refers to the final selenomethionine crystal as it had a lower B factor and more interpretable maps than native crystals without any appreciable differences in the overall structure (data not shown). Residues 3-7 of Sec13, 554-555 of Nup145C, and 1-32 of Nup84 are not modeled, in addition to residues missing in $Nup145C_{109-179}$•Sec13. The absence of observed density for α1 of Nup84 may be due to the absence of ACE1 helix α17 in the crystal construct, which typically interacts with helix α1 in ACE1 proteins. The final model has Ramachandran plot values of 87.9% allowed, 10.9% allowed, and 1.2% outliers.

The high-resolution structure fragments superimpose well with the corresponding regions in the complete Nup84•Nup145C•Sec13 structure. $Nup145C_{109-179}$•Sec13 aligns with the same region in the trimeric structure with an average rmsd of 0.75 Å. The major difference is the orientation of the N-terminal 11 residues of Sec13. In the $Nup145C_{109-179}$•Sec13 structure, this region is extended away from the molecule and amino acids 3-8 from Sec13 form a strand E zipper closure with strand D from blade 2 of a neighboring Sec13 molecule. A short strand is formed from part of the loop connecting Sec13 and Nup145C and forms strand F. In the Nup84•Nup145C•Sec13 crystal, this interaction is not possible as there is not a symmetry related Sec13 molecule in an equivalent position. Instead, the N-terminus of Sec13 extends towards the N-terminal helix of Nup145C, though it is not modeled. We presume that the zipper interaction of two Sec13 molecules and the linker in the $Nup145C_{109-179}$•Sec13 structure is a crystal-packing artifact. Nup145C•Sec13 in the trimeric structure overlays well with the reported *S.*

cerevisiae / *H. sapiens* hybrid structure with an average rmsd of 1.2 Å. The major difference between the two structures is a rotation of the Nup145C insertion blade/Sec13 unit of ~5-10° about the propeller axis. Whether this is a relevant movement of the molecules remains to be determined. Some rearrangement in the crown of Nup145C is observed in the current structure that is accounted for by reordering to form the interaction site for Nup84.

Structure analysis

Structure figures were made in Pymol (http://www.pymol.org). Interface calculations were performed using the PISA server (Krissinel and Henrick, 2004). Alignments were made with MUSCLE (Edgar, 2004), analyzed in Jalview (Waterhouse et al., 2009), and figures produced with Aline (Bond and Schuttelkopf, 2009). Structural superpositions were performed with Coot (Krissinel and Henrick, 2004) and Cealign (Jia et al., 2004).

Acknowledgements

Chapter five

Extending the lattice model of the NPC

A remarkable amount of insight into the structure and evolutionary origins of the nuclear pore complex (NPC) has been gained in the past several years. Though a relatively small percentage of the NPC has been structurally characterized in that time, it has provided us not only with firm evidence of a common ancestry of the NPC and COPII vesicle coats by virtue of the shared ancestral coatomer element 1 (ACE1) fold, but a means to create a conceptual framework for the entire NPC structure. Here I discuss future directions of research into the structure of the NPC, expand upon potential parallels to other vesicle coating systems, and hypothesize how other components of the NPC could be incorporated into the assembly based on our working model of the NPC lattice.

From A to Y: a complete Y complex structure

The most direct extension of work described in this thesis is the complete structural characterization of the Y complex. To date, we have x-ray structures for ~500 kDa of the 575 kDa complex (~87%) combined from yeast and human (Berke et al., 2004; Boehmer et al., 2008; Brohawn et al., 2008; Brohawn and Schwartz, 2009b; Leksa et al., 2009; Whittle and Schwartz, 2009). What remains is a short 4-helix stretch between the ACE1 trunk and tail of Nup84 and the "hub" of the Y complex formed by the interaction of the C-terminal (~35 kDa) domain of Nup120 with the tail modules Nup145C and Nup85 (~20 kDa each).

The portion of Nup84 that is structurally uncharacterized can be confidently predicted to be a group of 4 ACE1 helices that form the transition from trunk to tail modules. This region bridges the structures of the *S. cerevisiae* Nup84 trunk and crown modules and the *H. sapiens* Nup107 (Nup84 homolog) tail module. It is clear from proteolysis experiments that the region is flexible in *H. sapiens* Nup107 (Boehmer et al., 2008), and therefore is likely to be difficult to crystallize. Several constructs of *S. cerevisiae* Nup84•Nup145C•Sec13 that include the missing section were made and evaluated, but did not crystallize in initial screening attempts. Perhaps the most likely to succeed approach would be to design *S. cerevisiae* Nup84 tail constructs in complex with Nup133 by analogy to

the Nup107•Nup133 structure that include the missing region in Nup84. If successful, this would have the additional benefit of providing structural information for *S. cerevisiae* Nup133, which we currently model based on the *H. sapiens* structure. Another approach would be to design constructs of Nup84 that splice out the crown module and contain just tail and trunk. While the structure of this region will show the orientation of the Nup84 modules with respect to one another, it is almost certain to verify our predicted model (Brohawn and Schwartz, 2009b) and is thus somewhat of a lower priority for future efforts.

In contrast, the uncharacterized hub of the Y complex is an immediately important next structural target. The hub structure will likely prove to be important for several reasons. First, it is still unclear from bioinformatic prediction methods what structure the C-terminal domain of Nup120 will adopt, apart form it being predominantly helical. Second, while we can confidently predict the overall fold of the tail modules of Nup145C and Nup85 (Brohawn et al., 2008), the manner in which they will interact with Nup120 is unclear. The interaction could be arranged similarly to that between the ACE1 tail module of Nup84 and Nup133 (Boehmer et al., 2008) in which the terminal helices of Nup84 bind Nup133. Consistently, we have shown that the C-terminal helix of Nup145C is involved in the interaction with Nup120 (Figure 2.7). However, Nup120 and Nup133 do not appear to be related proteins: we have shown that their N-terminal regions adopt distinct folds (Berke et al., 2004; Boehmer et al., 2008; Leksa et al., 2009; Whittle and Schwartz, 2009) and their C-terminal domains do not have convincing sequence homology. Still, this does not exclude the possibility of Nup120 adopting a similar fold to the Nup84 binding site of Nup133 in its C-terminal region, perhaps arranged as a tandem array in order to provide binding sites to both Nup85 and Nup145C. Third, the structure of the hub would also answer the remaining questions about the relative orientations of the proteins in the Y complex. We have speculated that the β-propellers of Sec13 and Seh1 could be in close proximity to one another in the Y complex, potentially serving as a vertex element in the higher order assembly of the NPC (Chapter 4). In contrast, a model

produced by fitting crystal structures into an EM reconstruction of the Y complex showed the reverse orientation for Nup85•Seh1 (Kampmann and Blobel, 2009). The structure of the hub is certain to resolve this discrepancy.

Beyond the Y: incorporation of other nucleoporins into the NPC lattice

Insights gained from structures of the majority of the Y complex and pieces of the Nic96 complex have led us to propose a lattice model for the NPC. Of course, at this point the model is far from complete and raises a number of questions to be answered in the future. The most exciting lines of inquiry will extend from this preliminary model and build up to the higher order structure of the NPC. Cleary, the general approach taken to study the Y complex will be applied to the remainder of the NPC. Dissection of the Nic96 complex, the Ndc1 membrane-spanning complex, and the peripheral complexes structures can likely be achieved in much the same manner as for the Y complex. However, insight gained from our work to date can also be used to ask more targeted biological questions in route to a complete atomic level description of the NPC. Specifically, one can ask whether other Nups will have functions similar to vesicle-associated proteins such as those that form lattice vertices or act as adaptors or tethers. Below, hypotheses for the manner of integration and function of scaffold nucleoporins in the NPC lattice are discussed.

β-propellers and vertex elements in the NPC lattice

One prediction from our lattice model for the NPC is that vertex elements will be used to connect the edge elements we have identified in the Y complex. Identifying the nucleoporins involved in putative vertex elements in the NPC lattice and the manner of their interaction is thus a major goal of future research.

The vertex arrangement in the NPC and COPII lattices might prove to be similar. The presence of Sec13 in both the NPC and COPII coats lends circumstantial support to this idea. In the COPII coat, 4 edge elements interact to form a vertex. Each edge contributes two interacting β-propellers, a proximal Sec31 7-bladed β-

propeller and a distal Sec13 6-bladed β-propeller (Chapter 2, (Fath et al., 2007; Stagg et al., 2008)). Four interactions between the β-propellers seem responsible for forming the vertex: two are invariable contacts (one between two Sec31 propellers and one between Sec13 and Sec31) and two are variable (both between neighboring Sec31 propellers) (Stagg et al., 2008). In the NPC, one molecule of Sec13 per edge element could be used in a vertex, but what additional propellers could be involved? Seh1 seems likely to be for several reasons. First, it is highly homologous to Sec13 and like Sec13 it interacts with an ACE1 protein, Nup85. Second, it is probably in close proximity to Sec13 in the Y complex (Chapter 4). Lastly, recent evidence suggests it may interact with other proteins and play a role in a separate coating complex away from the NPC. Neither Nup145C nor Nup85 have a 7-bladed β-propeller like Sec31, so if an additional propeller is used in a NPC vertex it would have to be contributed from another Nup. Candidates include the other conserved β-propellers in the Y complex (in Nup120 or Nup133) or the Nic96 complex (Nup157/Nup170). Three additional β-propellers interact with the Y complex in metazoans (Nup37, Nup43, and ELYS) and though these have yet to be identified in *S. cerevisiae*, two are present in *A. nidulans* (Nup37 and ELYS) (Brohawn et al., 2009). The functions of these propellers have yet to be investigated.

While it seems reasonable to expect that other β-propellers in the NPC will play a role in vertex interactions, evidence of propeller-propeller interactions is currently lacking. This is perhaps indicative of weak interactions at the expected vertices, which would thus not have been likely to be indentified in studies probing direct interactions of Nups to date. Consistent with this notion is the fact that vertex interactions in the COPII lattice have only been characterized in coat reconstructions by EM (Fath et al., 2007; Stagg et al., 2008). With the expectation that vertex contacts in the NPC will be of a weak and/or transient nature, future studies can be designed to specifically probe for these types of interactions. It may be important to not only test interactions between known β-propellers of the NPC, but to revisit global approaches to ascertain whether

another β-propeller-containing protein is involved that had been missed in previous NPC inventorying experiments.

What might be the role of β-propellers in the NPC be if they are not used in COPII-type vertex interactions? One possibility is that some may have roles similar to that demonstrated for the β-propeller in clathrin coats. The seven-bladed β-propeller at the N-terminus of the clathrin heavy chain is not involved in lattice vertex interactions, but rather projects inward toward the vesicle membrane (Fotin et al., 2004). The β-propeller binds to a wide variety of cargo adaptors and membrane-interacting proteins and thus serves to bridge the outer coat to the membrane and contents of the vesicle (Owen et al., 2004; Young, 2007). This is accomplished through multiple protein binding sites on the clathrin β-propeller, present both on the top face of the propeller and in grooves on the outer face between β-blades (ter Haar et al., 2000). One candidate β-propeller in the NPC for a similar role is in Nup120. In our current model for the arrangement of the Y complex in the NPC, Nup120 extends its N-terminal β-propeller towards the nuclear envelope (NE) membrane (Chapter 4). In this position, it is ideally situated to bind to other nucleoporins that would link the outer edge element in the Y complex to the NE. Interestingly, the β-propeller surface of Nup120 displays two regions of high sequence conservation: one on the top surface and one on the outer face between blades 3 and 4 (Chapter 3). Nup120 has been shown to interact with Nup157 (Lutzmann et al., 2005), which could play a role in linking the Y complex to the membrane (see below). Future work taking a targeted approach to testing this interaction as well as identifying other potential Nup120 β-propeller interacting Nups should prove interesting.

β-propellers on the periphery of the NPC are likely to be involved in scaffolding and recruiting other proteins with functional roles at the NPC. This has been shown to be the case for the interaction of the Nup214 β-propeller with the RNA helicase Ddx19 at the cytoplasmic face of the NPC (Napetschnig et al., 2007; von Moeller et al., 2009) any may also be true for Nup82 (Xu and Powers, 2009).

α-helical nucleoporins in the NPC lattice

α-helical domains make up ~50% of the mass of the NPC (Chapter 1). Early bioinformatic approaches predicted that all of the non coiled-coil helical domains would adopt regular α-solenoid like folds (Devos et al., 2006). Surprisingly, each helical domain to be structurally characterized has so far been found to fall into one of three distinct, non α-solenoid folds: ACE1, the Nup133/Nup170 fold, or the Nup120 fold (Brohawn et al., 2008; Brohawn and Schwartz, 2009b; Leksa et al., 2009; Whittle and Schwartz, 2009). The role that the Nup120 helical domain serves in the NPC lattice in bridging Nup145C and Nup85 in the Y complex was discussed above. Possible functions of the ACE1 proteins not known to form edge elements and of the Nup133/Nup170 fold in the lattice are discussed below. It is sill not clear whether the two largest helical Nups, Nup188 and Nup192, will adopt a more regular helical repeat arrangement, one of the three helical folds seen to date in the NPC, or present yet another surprise.

ACE1 nucleoporins

Four nucleoporins (Nic96, Nup85, Nup145C, and Nup84) have been shown to adopt ACE1 folds. As discussed previously (Chapter 4), Nup145C and Nup84 are proposed to form an edge element in the NPC lattice. However, it remains to be determined what exactly the roles of the Nic96 and Nup85 ACE1s in the NPC lattice are. It is tempting to speculate that the ACE1 crown modules of Nic96 and Nup85 will be involved in crown-crown interactions either homo- or hetero-typically to form additional edges. However, the crowns of Nic96 and Nup85 are distinct from other ACE1s in that their surface is neither noticeably hydrophobic nor well conserved (Brohawn et al., 2008; Jeudy and Schwartz, 2007). This suggests that if they do participate in crown-crown interactions, they might be weaker in nature than the edge interactions so far observed. It also is unclear what the role of the tail module of Nic96 is, though it was predicted to form an interaction site from structural analysis before it was recognized as an ACE1 (Jeudy and Schwartz, 2007). Structural homology between Nup133 (which interacts with the Nup84 ACE1 tail) and Nup157/Nup170 (see below) suggest

Nup157/Nup170 could interact with the Nic96 tail, though this has yet to be observed.

It may be that Nic96 and Nup85 are used differently in the NPC lattice than the other known ACE1 proteins. Interestingly, in both crystals, pairs of molecules form antiparallel homotypic interactions along their long axes. Nic96 has an especially striking surface charge distribution that results in a dipole moment due to a generally positively charged crown and negatively charged tail that may contribute to the interaction seen in the crystal (Jeudy and Schwartz, 2007; Schrader et al., 2008b). While these interactions are not observed in solution, they may be relevant within the confines of the assembled NPC, with its correspondingly high protein concentrations. Future work should elucidate the role of these two ACE1 proteins in the NPC lattice.

Nup133/Nup170 and other helical nucleoporins – adaptors/tethers in the NPC?

A second α-helical fold indentified in Nups is shared between the Y complex member Nup133 and the Nic96 complex components Nup157/Nup170 (Whittle and Schwartz, 2009). The architecture of this second ancestral element is quite distinct from either ACE1 or regular α-solenoid like folds. While the relationship between Nup133 and Nup157/Nup170 is more distant than between ACE1s, the proteins clearly share a stretched, multipartite helical arrangement with conserved topology. The role for these Nups in the NPC architecture is still unclear, but the presence of a membrane binding ALPS motif and Nup84 interaction site in human Nup133 and evidence for Nup157/Nup170 being linked to transmembrane Nups (possibly through Nup53/Nup59), Nup120, and anchoring other Nups to the NPC suggests they may both act as connectors in the lattice (Drin et al., 2007; Flemming et al., 2009; Lutzmann et al., 2005; Makio et al., 2009; Onischenko et al., 2009).

A recent structure of the Dsl1 vesicle-tethering complex has provided a basis for speculation that the Nup133/170 fold is related to tethering complex proteins and may play a functionally similar role in the NPC (Ren et al., 2009). Vesicle tethering complexes are multiprotein assemblies that bridge vesicles to target membranes, typically over large distances. The trimeric Dsl1 complex is involved in tethering COPI vesicles to the ER membrane (Sztul and Lupashin, 2009). Vesicle interaction is accomplished by one subunit (Dsl1) binding to COPI coatomers with an extended loop. Dsl1 also binds to the other two components, Sec39 and Tip20, which are in turn anchored to the ER membrane via interaction with ER associated SNAREs (Ren et al., 2009). Interestingly, while Sec39 of the Dsl1 complex was described as a novel fold, its C-terminal region is reminiscent of the Nup133/170 domain. Preliminary structural superpositions show that Sec39 and Nup133/170 indeed share the same fold topology in this domain. Despite the many differences in the systems, at a simplistic level Nup133, Nup157/Nup170, and Sec39 may have a common function in linking membrane-coating proteins to membrane-binding proteins. It is thus not unreasonable to hypothesize they evolved from an ancestor with that same capability. Whether this relationship is meaningful or not remains to be addressed in future work.

Regardless, it seems likely that at least some nucleoporins will play an adaptor-like role in connecting the peripheral lattice of the NPC including the edge element in the Y complex to the pore membrane. One or more Nups of the Nic96 complex (including Nup53/Nup59, Nup188, and Nup192) seem most likely to play such a role. It is possible that these adaptor interactions will be more complex and difficult to elucidate than the strong binary interactions observed in the Y complex. As has been seen in vesicle coat adaptor complexes, they may involve tertiary or higher order interactions, be relatively weak or transient, or might involve unstructured regions of peptide outside of well-ordered domains (Edeling et al., 2006; McMahon and Mills, 2004).

Overall features

As structural knowledge of the NPC architecture progresses, we are able to begin to ask questions relating to the overall NPC structure. While many issues are likely to be resolved in the coming years, two in particular that may be addressed by extension of the lattice model are elaborated on below.

Flexibility of the NPC

It will be interesting to see if a lattice model is able to account for structural plasticity and deviations from eightfold symmetry within NPCs. Deviations from eight fold symmetric NPCs have been consistently observed (Akey, 1995; Beck et al., 2007; Frenkiel-Krispin et al., 2009) and in fact extend to the presence of supernumerary complexes of 9 or 10 scaffold segments (Hinshaw and Milligan, 2003). Construction of the NPC as a lattice akin to vesicle coatomers suggests a possible mechanism for these observations. In both COPII and clathrin lattices, significant heterogeneity in coat architecture is observed, though the lattices accomplish flexibility of assembly in fundamentally different ways. In clathrin coated vesicles, a variety of lattice organizations are made possible by restriction of the hub (or vertex) structure while allowing for variable interactions along the legs (edges) of the clathrin triskelion (Fotin et al., 2004). In COPII lattices, variation of vertex interactions has been shown to be able to account for different coat geometries, while the edge interactions remain relatively fixed (Stagg et al., 2008). In both cases, fixing the geometry of the edge or vertex interaction and allowing for variability in the other results in flexibility in lattice assembly. Once it is determined how NPC vertex interactions are arranged, the question of whether this principle is recapitulated in the NPC could be addressed.

Lattice architecture and membrane curvature

The nuclear pore membrane is unique in that it defined by both positive and negative curvature. Highly curved membranes are generally unstable and so presumably there will be mechanisms to generate and stabilize both positive and negative membrane curvature in the pore membrane (McMahon and Gallop,

2005; Shibata et al., 2009; Zimmerberg and Kozlov, 2006). While we are beginning to develop a model for membrane stabilization in the NPC, major questions still to be resolved include how pore membrane curvature is generated and how the negative membrane curvature is stabilized.

In the direction parallel to the transport axis, the pore membrane displays positive curvature. In the COPII lattice, the Sec31•Sec31 edge element is arranged parallel to and stabilizes the positive curvature of the vesicle membrane (Fath et al., 2007; Stagg et al., 2008). Based on the equivalent architecture of the edge element in the Y complex of the NPC, we propose it will likewise be arranged parallel to the positive membrane curvature of the pore membrane and will function similarly (Chapter 4). A recent cryo-ET reconstruction of Xenopus NPCs has defined densities arranged parallel to the transport axis that are roughly the expected size of two Y complexes (Frenkiel-Krispin et al., 2009). Although it has not yet been possible to fit high resolution models of the Y complex into the tomographic reconstructions, these data are consistent with our model of Y complexes arranged as struts in the NPC lattice. The surface of the Y complex proposed to face towards the membrane is not observed to display a noticeable positively charged surface or other characteristic consistent with direct membrane binding (Figure 4.9). Thus, it (like the COPII outer coat) is proposed to play a membrane scaffolding and curvature-organizing role.

Interaction between proteins and lipid bilayers can generate curvature in different ways that may be important in the NPC (McMahon and Gallop, 2005; Shibata et al., 2009; Zimmerberg and Kozlov, 2006). Amphipathic helix insertion into membrane leaflets is one mechanism that produces local regions of positive curvature. For instance, in COPI, COPII, and clathrin-coated vesicles, positive curvature is initiated by helix insertion of the GTPases Arf1, Sar1, and dynamin, respectively (Pucadyil and Schmid, 2009). In the NPC, the presence of an amphipathic, membrane-inserting helix ('ALPS'-motif) in human Nup133 has been demonstrated to sense positive curvature in vitro (Drin et al., 2007), though

the importance of this domain *in vivo* and its generality across species in still unclear. Recent studies have also implicated membrane insertion by reticulons in NPC assembly (Dawson et al., 2009). Integral membrane proteins of the NPC might directly induce membrane curvature if the transmembrane domains are funnel shaped and add surface area preferentially to one leaflet. Alternatively, oligomerization of either the NE luminal or NPC facing domains of transmembrane Nups could impart and stabilize curvature. In fact, the NE luminal domains of the yeast transmembrane Nups have been suggested to oligomerize into a ring structure (Alber et al., 2007a). This type of structure could be responsible for imparting and stabilizing the negative curvature of the pore membrane in the plane of the nuclear envelope.

Proteins that bind to the surface of lipid bilayers can also generate membrane shape. This manner of interaction is typified by the BAR domain. BAR domains share the characteristics of forming crescent shaped helical assemblies that bind membranes and impart curvature. Variations include canonical BAR domains that bind membrane surfaces with their concave face and induce positive curvature, N-BAR domains that have an additional amphipathic helices that insert into the membrane to further effect curvature, and inverse-BAR domains that bind lipids with their convex face and induce negative curvature (Suetsugu et al., 2009). Unrelated proteins also use a similar strategy: the Bro1 domain has a convex membrane-binding site and induces negative curvature and the COPII adaptor complex Sec23•34 has a concave membrane-binding surface that stabilizes positive curvature initiated by Sar1 (Bi et al., 2002; Kim et al., 2005). While sequence analysis of Nups has not revealed homology to these types of domains, clearly a number of systems have evolved to utilize similar structures. It will be very interesting to see if a Nup or complex of Nups will similarly present a curved lipid-binding surface to impart and/or stabilize curvature.

Additionally, the lipid composition at the pore membrane could have effects on membrane organization and structure and its contributions remain unexplored.

As the architecture of the NPC becomes clear, significant interplay between different mechanisms of membrane deformation and stabilization is likely to be revealed.

Summary

Structural insight in to the NPC has progressed rapidly over the past several years. While a complete atomic level structure is still far in the future, work described in this thesis has allowed us to produce a model for the NPC architecture that is lattice-like based on an evolutionary relationship with the COPII vesicle coat. Importantly, this insight and a working model of the NPC lattice allows for the generation of a number of specific hypotheses for the manner of incorporation and function of uncharacterized Nups and the NPC assembly as a whole. Much exciting work on the path to a full structural description of the NPC is clearly still to come.

Appendix

A lattice model for the nuclear pore complex

Abstract

The nuclear pore complex (NPC) is one of the largest protein machines in the cell and forms the sole conduit for nucleocytoplasmic transport in eukaryotes. The NPC is composed of an eightfold radially symmetric scaffold of architectural proteins that anchor a set of phenylalanine-glycine (FG) repeat proteins that form the transport barrier. As a step toward elucidating the molecular architecture of the NPC, we solved the structure of nucleoporin 85 (Nup85) in complex with Seh1, a module in the heptameric Nup84 or Y subcomplex. We define a new tripartite protein element, the ancestral coatomer element ACE1, which Nup85 specifically shares with several other nucleoporins and vesicle coat proteins. We predicted and verified functional sites on nucleoporin ACE1 members based on analogy to ACE1 interactions that propagate the COPII vesicle coat. Thus, we provide the first experimental evidence for evolution of the NPC and vesicle coats from a common ancestor. We propose that the NPC structural scaffold, like vesicle coats, is a polygonal network composed of vertex and edge elements that forms a molecular lattice upon which additional nucleoporins assemble. Here we further discuss our findings and elaborate on our lattice model of the nuclear pore complex.

Discussion

All nucleocytoplasmic transport in a cell proceeds through nuclear pore complexes (NPCs). NPCs are composed of an eight-fold radially symmetric structural scaffold that anchors a group of FG-repeat-containing proteins that form the transport barrier (D'Angelo and Hetzer, 2008; Tran and Wente, 2006; Weis, 2003). Elucidating the three dimensional structure of the NPC is critical for understanding its roles in nucleocytoplasmic transport and cellular homeostasis. A path towards an atomic resolution structure of the 40-60 MDa NPC is made possible by the realization that the NPC is a modular assembly (Schwartz, 2005). The NPC is composed of ~30 proteins (Nups) that are arranged into distinct subcomplexes, each present in multiple copies (Rout et al., 2000). The structural scaffold contains the most stably attached nups and comprises two

subcomplexes in yeast: the heptameric Y complex (composed of Nup133, Nup84, Nup145C, Sec13, Nup85, Seh1, and Nup120) and the heteromeric Nic96 subcomplex (likely composed of Nic96, Nup192, Nup188, Nup157/170, Nup59, and Nup53) (Rabut et al., 2004).

In an attempt to better understand the architecture of the NPC, we solved the crystal structure of Nup85 in complex with Seh1 (Brohawn et al., 2008). Nup85 interacts with Seh1 via insertion of an N-terminal insertion blade into the open 6-bladed β-propeller of Seh1. The remainder of Nup85 forms a uniquely arrayed J-shaped α-helical block with two distinct units we term "crown" and "trunk". The fold is notably different from the regular α-helical stack that was predicted (Devos et al., 2004; Devos et al., 2006).

We found that this fold is shared in four other proteins of known structure: the nucleoporins Nup145C, Nup84, and Nic96 and the COPII vesicle coatomer Sec31 (Boehmer et al., 2008; Fath et al., 2007; Hsia et al., 2007; Jeudy and Schwartz, 2007). While a related architecture between the NPC and vesicle coats has been proposed based on similar fold composition (Devos et al., 2004), we provide the first experimental and structural evidence of a common ancestry (Brohawn et al., 2008). Comparison of the structures shows a shared core composed of three modules: the crown, trunk and C-terminal tail. We termed this tripartite fold the ancestral coatomer element 1 (ACE1). While the overall organization and topology is identical, there are significant differences in the relative orientation of the modules between members. This suggests the boundaries between ACE1 modules may function as hinges. This relationship was not previously predicted due to low sequence conservation and was initially obscured at the structural level by differences in relative orientations of the modules (Hsia et al., 2007).

We used characterized interactions of ACE1 proteins to predict functional sites on other members. First, in COPII vesicle coats, Sec31•Sec13 dimers form edge

elements through Sec31 crown•crown homodimerization (Figure A.1A) (Fath et al., 2007). We predicted and demonstrated that Nup145C likewise interacts crown•crown with its binding partner in the Y complex, Nup84. Second, we predicted and verified that Nup85 and Nup145C tails interact with Nup120 as the Nup84 tail module interacts with Nup133 (Boehmer et al., 2008). These and other data were used to construct an improved model of the Y complex (Figure A.1B).

We propose that the nuclear pore complex scaffold has a lattice structure assembled from vertex and edge elements similar in principle to vesicle coats. We can envision at least two alternative models for this lattice. In the first (Figure A.1B(i)), two rings of the Y complex sandwich an inner ring of the Nic96 subcomplex. This arrangement would generate a scaffold of ~50 nm diameter, consistent with the observed pore size in yeast (Kiseleva et al., 2004; Yang et al., 1998). Alternatively (Figure A.1B(ii)), the scaffold may consist of two stacked rings of the Y complex without an intervening Nic96 subcomplex ring, which may be sufficient to traverse the ~30-50 nm pore height. Uncertainty about the exact arrangement arises from still incompletely understood stoichiometries of components and limited information about overall NPC size. Future clarification of the connectivity between scaffold subcomplexes will additionally help to discern the possible lattice arrangements. Homology to COPII coatomers suggests ACE1-containing subcomplexes will be edge elements in the NPC lattice. The nature of the vertex elements in the NPC is less clear, though it may well also involve β-propeller-β-propeller interactions.
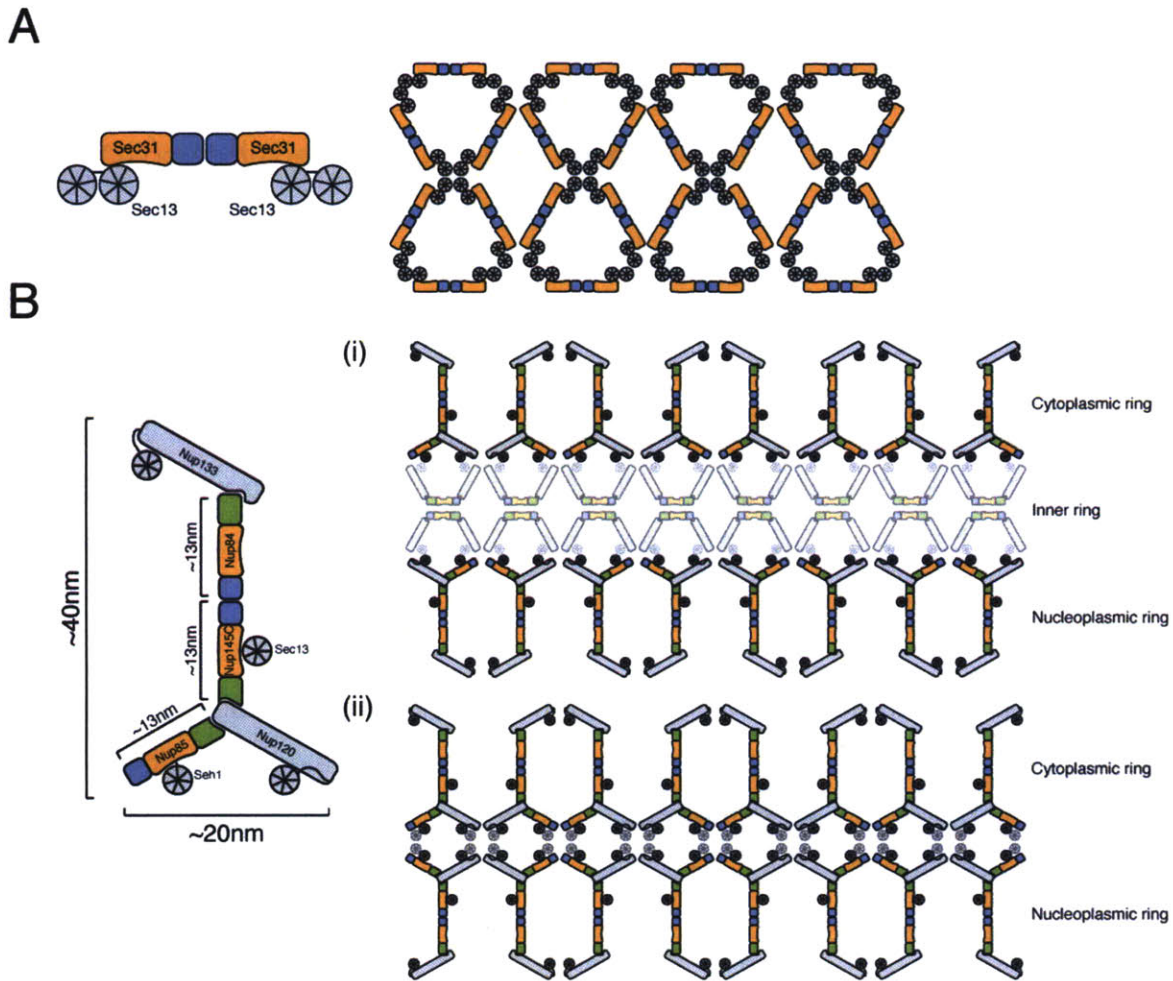
Figure A.1 - A lattice model of the NPC
Sec31, Nup85, Nup145C, Nup84, and Nic96 ACE1 proteins are colored with crowns blue, trunks orange, and tail modules green. Other protein folds are shown in grey. (A) Schematic organization of the COPII outer vesicle coat. On the left, an edge element consisting of two Sec31•Sec13 heterodimers is shown. Two Sec31 molecules interact crown•crown. On the right, an entire COPII cuboctahedron coat composed of 24 edge elements is shown unwrapped and laid flat (Fath et al., 2007). Vertex elements are formed where four Sec31 and four Sec13 β-propellers interact. (B) Alternative organizations of the NPC lattice. On the left, the Y complex is shown in schematic fashion illustrating how ACE1 interactions organize the Y-shaped structure. Nup145C and Nup84 also interact crown•crown. On the right, the entire NPC structural scaffold is shown unwrapped and laid flat. Two rings of the Y complex form the lattice of the NPC scaffold either with an intervening ring of the Nic96 subcomplex (i) or alone (ii). Both the identity and organization of the vertex elements and the Nic96 subcomplex in the pore lattice are unknown and are shown half-transparent. The presented organization is not meant to predict relative positions of proteins or the structure per se, but rather emphasizes the principally similar lattice organization of NPCs and vesicle coats.

Our lattice model of the NPC prompts a number of additional potential parallels to vesicle coats. First, vesicle coatomers do not directly contact membranes, but use adaptor protein complexes to span the ~8 nm gap and recruit cargo (Owen et al., 2004). Consistently, a ~8 nm gap has been observed between the structural scaffold of the NPC and the nuclear membrane (Beck et al., 2007). Conceivably other nups fill corresponding adaptor complex roles by linking the lattice to transmembrane nups, or transmembrane nups could act directly as adaptors. Second, COPII vesicle coats size flexibility is made possible largely by hinges at coat vertices (Stagg et al., 2008). It may be that analogous hinges as well as those between ACE1 modules in the NPC lattice confer plasticity that may be used for pore dilation or NPC (dis)assembly.

Recent work has produced two conflicting models of the molecular organization of the NPC. A computationally generated model that integrates a wealth of localization, interaction, and other primary data similarly places the Y complex in two peripheral NPC rings flanking an inner ring composed of the Nic96 subcomplex (Alber et al., 2007a; Alber et al., 2007b). In contrast, a model based on crystal packing interactions in Nup145C•Sec13 places the Y complex in four stacked rings organized by hetero-octameric poles of Nup145C•Sec13 and Nup85•Seh1 units (Hsia et al., 2007). A tube of 32 Y complexes was proposed to envelope inner cylinders of the Nic96 subcomplex and FG Nups generating a "concentric cylinder" model of the NPC (Debler et al., 2008; Hsia et al., 2007).

Our model is incompatible with the "concentric cylinder" model for the NPC. Specifically, the demonstrated crown•crown interaction between Nup84 and Nup145C overlaps with Nup145C crystal contacts necessary for the propagation of the models hetero-octameric poles. Exposed hydrophobic surfaces tend to form crystal-packing contacts; whether or not they are physiologically relevant needs to be addressed by additional experiments (Kobe et al., 2008). Careful analysis of packing interactions in crystals can unveil biologically important protein interfaces, especially if the crystallized proteins are part of a higher-order

assembly *in vivo*, as nups are. Crystal contacts with at least some hydrophobic character are also observed in the structures of Nup85•Seh1, Nup107•Nup133 and Nic96 (Boehmer et al., 2008; Brohawn et al., 2008; Jeudy and Schwartz, 2007; Schrader et al., 2008b), which are all fragments of larger assemblies. While some of these interactions can likely be ruled out as crystal artifacts, i.e. because they involve surfaces created by the use of truncated proteins that would otherwise be buried in the hydrophobic core of the protein, others may be indicative of real functional sites. Weak interactions observed in crystals may point to inter-subcomplex contact areas that are important for self-assembly and to date have not been observed in solution.

We have provided the first structural evidence of a common ancestry of vesicle coats and the nuclear pore complex and provide a lattice model of the NPC based on this commonality. Our lattice model is generally consistent with the computational model of the NPC (Alber et al., 2007b), though the absence of additional structural knowledge precludes a detailed comparison. Our model provides a framework upon which further structural and cell biological studies can be placed in an effort to more fully understand the assembly principles and function of the NPC.

Acknowledgements

References

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr *58*, 1948-1954.

Aitchison, J.D., Blobel, G., and Rout, M.P. (1995). Nup120p: a yeast nucleoporin required for NPC distribution and mRNA transport. J Cell Biol *131*, 1659-1675.

Akey, C.W. (1995). Structural plasticity of the nuclear pore complex. J Mol Biol *248*, 273-293.

Akey, C.W., and Radermacher, M. (1993). Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. J Cell Biol *122*, 1-19.

Akhtar, A., and Gasser, S.M. (2007). The nuclear envelope and transcriptional control. Nat Rev Genet *8*, 507-517.

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., *et al.* (2007a). Determining the architectures of macromolecular assemblies. Nature *450*, 683-694.

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

Alber, F., Forster, F., Korkin, D., Topf, M., and Sali, A. (2008). Integrating diverse data for structure determination of macromolecular assemblies. Annu Rev Biochem *77*, 443-477.

Antonin, W., Ellenberg, J., and Dultz, E. (2008). Nuclear pore complex assembly through the cell cycle: regulation and membrane organization. FEBS Lett *582*, 2004-2016.

Aravind, L., Iyer, L.M., and Koonin, E.V. (2006). Comparative genomics and structural biology of the molecular innovations of eukaryotes. Curr Opin Struct Biol *16*, 409-419.

Aris, J.P., and Blobel, G. (1989). Yeast nuclear envelope proteins cross react with an antibody against mammalian pore complex proteins. J Cell Biol *108*, 2059-2067.

Bailer, S.M., Balduf, C., and Hurt, E. (2001). The Nsp1p carboxy-terminal domain is organized into functionally distinct coiled-coil regions required for assembly of nucleoporin subcomplexes and nucleocytoplasmic transport. Mol Cell Biol *21*, 7944-7955.

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

Beck, M., Lucic, V., Forster, F., Baumeister, W., and Medalia, O. (2007). Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. Nature *449*, 611-615.

Belgareh, N., Snay-Hodge, C., Pasteau, F., Dagher, S., Cole, C.N., and Doye, V. (1998). Functional characterization of a Nup159p-containing nuclear pore subcomplex. Mol Biol Cell *9*, 3475-3492.

Bennett-Lovsey, R.M., Herbert, A.D., Sternberg, M.J., and Kelley, L.A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. Proteins *70*, 611-625.

Berke, I.C., Boehmer, T., Blobel, G., and 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.

Bhattacharya, A. (2009). Protein structures: Structures of desire. Nature *459*, 24-27.

Bi, X., Corpina, R.A., and Goldberg, J. (2002). Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. Nature *419*, 271-277.

Boehmer, T., Enninga, J., Dales, S., Blobel, G., and 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 U S A *100*, 981-985.

Boehmer, T., Jeudy, S., Berke, I.C., and Schwartz, T.U. (2008). Structural and Functional Studies of Nup107/Nup133 Interaction and Its Implications for the Architecture of the Nuclear Pore Complex. Mol Cell.

Bond, C.S., and Schuttelkopf, A.W. (2009). ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr D Biol Crystallogr *65*, 510-512.

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

Brohawn, S.G., Partridge, J.R., Whittle, J.R., and Schwartz, T.U. (2009). The nuclear pore complex has entered the atomic age. Structure *17*, 1156-1168.

Brohawn, S.G., and Schwartz, T.U. (2009a). A lattice model of the nuclear pore complex. Commun Integr Biol *2*, 205-207.

Brohawn, S.G., and Schwartz, T.U. (2009b). Molecular architecture of the Nup84-Nup145C-Sec13 edge element in the nuclear pore complex lattice. Nat Struct Mol Biol *16*, 1173-1177.

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.

157

Brunger, A.T., DeLaBarre, B., Davies, J.M., and Weis, W.I. (2009). X-ray structure determination at low resolution. Acta Crystallogr D Biol Crystallogr *65*, 128-133.

Carmody, S.R., and Wente, S.R. (2009). mRNA nuclear export at a glance. J Cell Sci *122*, 1933-1937.

Chaudhuri, I., Soding, J., and Lupas, A.N. (2008). Evolution of the beta-propeller fold. Proteins *71*, 795-803.

Cheng, Y., Boll, W., Kirchhausen, T., Harrison, S.C., and Walz, T. (2007). Cryo-electron tomography of clathrin-coated vesicles: structural implications for coat assembly. J Mol Biol *365*, 892-899.

Chiu, W., Baker, M.L., and Almo, S.C. (2006). Structural biology of cellular machines. Trends Cell Biol *16*, 144-150.

Clamp, M., Cuff, J., Searle, S.M., and Barton, G.J. (2004). The Jalview Java alignment editor. Bioinformatics *20*, 426-427.

Conti, E., Muller, C.W., and Stewart, M. (2006). Karyopherin flexibility in nucleocytoplasmic transport. Curr Opin Struct Biol *16*, 237-244.

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

Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., and Matunis, M.J. (2002). Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol *158*, 915-927.

D'Angelo, M.A., and Hetzer, M.W. (2008). Structure, dynamics and function of nuclear pore complexes. Trends Cell Biol *18*, 456-466.

D'Angelo, M.A., Raices, M., Panowski, S.H., and Hetzer, M.W. (2009). Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. Cell *136*, 284-295.

Daigle, N., Beaudouin, J., Hartnell, L., Imreh, G., Hallberg, E., Lippincott-Schwartz, J., and Ellenberg, J. (2001). Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells. J Cell Biol *154*, 71-84.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S., *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res *35*, W375-383.

Dawson, T.R., Lazarus, M.D., Hetzer, M.W., and Wente, S.R. (2009). ER membrane-bending proteins are necessary for de novo nuclear pore formation. J Cell Biol *184*, 659-675.

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

DeGrasse, J.A., DuBois, K.N., Devos, D., Siegel, T.N., Sali, A., Field, M.C., Rout, M.P., and Chait, B.T. (2009). Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. Mol Cell Proteomics 8, 2119-2130.

DeLaBarre, B., and Brunger, A.T. (2006). Considerations for the refinement of low-resolution crystal structures. Acta Crystallogr D Biol Crystallogr 62, 923-932.

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

Devos, D., Dokudovskaya, S., Williams, R., Alber, F., Eswar, N., Chait, B.T., Rout, M.P., and Sali, A. (2006). Simple fold composition and modular architecture of the nuclear pore complex. Proc Natl Acad Sci U S A 103, 2172-2177.

Doublie, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods Enzymol 276, 523-530.

Doye, V., Wepf, R., and Hurt, E.C. (1994). A novel nuclear pore protein Nup133p with distinct roles in poly(A)+ RNA transport and nuclear pore distribution. Embo J 13, 6062-6075.

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

Drummond, S.P., Rutherford, S.A., Sanderson, H.S., and Allen, T.D. (2006). High resolution analysis of mammalian nuclear structure throughout the cell cycle: implications for nuclear pore complex assembly during interphase and mitosis. Can J Physiol Pharmacol 84, 423-430.

Dujon, B. (2006). Yeasts illustrate the molecular mechanisms of eukaryotic genome evolution. Trends Genet 22, 375-387.

Dultz, E., Zanin, E., Wurzenberger, C., Braun, M., Rabut, G., Sironi, L., and Ellenberg, J. (2008). Systematic kinetic analysis of mitotic dis- and reassembly of the nuclear pore in living cells. J Cell Biol 180, 857-865.

Edeling, M.A., Smith, C., and Owen, D. (2006). Life of a clathrin coat: insights from clathrin and AP structures. Nat Rev Mol Cell Biol 7, 32-44.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792-1797.

Elad, N., Maimon, T., Frenkiel-Krispin, D., Lim, R.Y., and Medalia, O. (2009). Structural analysis of the nuclear pore complex by integrated approaches. Curr Opin Struct Biol 19, 226-232.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-2132.

Fabre, E., and Hurt, E. (1997). Yeast genetics to dissect the nuclear pore complex and nucleocytoplasmic trafficking. Annu Rev Genet *31*, 277-313.

Fahrenkrog, B., Aris, J.P., Hurt, E.C., Pante, N., and Aebi, U. (2000). Comparative spatial localization of protein-A-tagged and authentic yeast nuclear pore complex proteins by immunogold electron microscopy. J Struct Biol *129*, 295-305.

Fath, S., Mancias, J.D., Bi, X., and Goldberg, J. (2007). Structure and Organization of Coat Proteins in the COPII Cage. Cell *129*, 1325-1336.

Feuerbach, F., Galy, V., Trelles-Sticken, E., Fromont-Racine, M., Jacquier, A., Gilson, E., Olivo-Marin, J.C., Scherthan, H., and Nehrbass, U. (2002). Nuclear architecture and spatial positioning help establish transcriptional states of telomeres in yeast. Nat Cell Biol *4*, 214-221.

Flemming, D., Sarges, P., Stelter, P., Hellwig, A., Bottcher, B., and Hurt, E. (2009). Two structurally distinct domains of the nucleoporin Nup170 cooperate to tether a subset of nucleoporins to nuclear pores. J Cell Biol *185*, 387-395.

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

Franz, C., Walczak, R., Yavuz, S., Santarella, R., Gentzel, M., Askjaer, P., Galy, V., Hetzer, M., Mattaj, I.W., and Antonin, W. (2007). MEL-28/ELYS is required for the recruitment of nucleoporins to chromatin and postmitotic nuclear pore complex assembly. EMBO Rep *8*, 165-172.

Frenkiel-Krispin, D., Maco, B., Aebi, U., and Medalia, O. (2009). Structural Analysis of a Metazoan Nuclear Pore Complex Reveals a Fused Concentric Ring Architecture. J Mol Biol.

Frey, S., and Gorlich, D. (2007). A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. Cell *130*, 512-523.

Fribourg, S., Braun, I.C., Izaurralde, E., and Conti, E. (2001). Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. Mol Cell *8*, 645-656.

Galy, V., Mattaj, I.W., and Askjaer, P. (2003). Caenorhabditis elegans nucleoporins Nup93 and Nup205 determine the limit of nuclear pore complex size exclusion in vivo. Mol Biol Cell *14*, 5104-5115.

Garcia De La Torre, J., Huertas, M.L., and Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys J *78*, 719-730.

Gillespie, P.J., Khoudoli, G.A., Stewart, G., Swedlow, J.R., and Blow, J.J. (2007). ELYS/MEL-28 chromatin association coordinates nuclear pore complex assembly and replication licensing. Curr Biol *17*, 1657-1662.

Grandi, P., Doye, V., and Hurt, E.C. (1993). Purification of NSP1 reveals complex formation with 'GLFG' nucleoporins and a novel nuclear pore protein NIC96. EMBO J *12*, 3061-3071.

Grandi, P., Schlaich, N., Tekotte, H., and Hurt, E.C. (1995). Functional interaction of Nic96p with a core nucleoporin complex consisting of Nsp1p, Nup49p and a novel protein Nup57p. Embo J *14*, 76-87.

Grant, R.P., Neuhaus, D., and Stewart, M. (2003). Structural basis for the interaction between the Tap/NXF1 UBA domain and FG nucleoporins at 1A resolution. J Mol Biol *326*, 849-858.

Guttinger, S., Laurell, E., and Kutay, U. (2009). Orchestrating nuclear envelope disassembly and reassembly during mitosis. Nat Rev Mol Cell Biol *10*, 178-191.

Handa, N., Kukimoto-Niino, M., Akasaka, R., Kishishita, S., Murayama, K., Terada, T., Inoue, M., Kigawa, T., Kose, S., Imamoto, N., et al. (2006). The crystal structure of mouse Nup35 reveals atypical RNP motifs and novel homodimerization of the RRM domain. J Mol Biol *363*, 114-124.

Harel, A., Orjalo, A.V., Vincent, T., Lachish-Zalait, A., Vasu, S., Shah, S., Zimmerman, E., Elbaum, M., and Forbes, D.J. (2003). Removal of a single pore subcomplex results in vertebrate nuclei devoid of nuclear pores. Mol Cell *11*, 853-864.

Hawryluk-Gara, L.A., Platani, M., Santarella, R., Wozniak, R.W., and Mattaj, I.W. (2008). Nup53 is required for nuclear envelope and nuclear pore complex assembly. Mol Biol Cell *19*, 1753-1762.

Hawryluk-Gara, L.A., Shibuya, E.K., and Wozniak, R.W. (2005). Vertebrate Nup53 interacts with the nuclear lamina and is required for the assembly of a Nup93-containing complex. Mol Biol Cell *16*, 2382-2394.

Heath, C.V., Copeland, C.S., Amberg, D.C., Del Priore, V., Snyder, M., and Cole, C.N. (1995). Nuclear pore complex clustering and nuclear accumulation of poly(A)+ RNA associated with mutation of the Saccharomyces cerevisiae RAT2/NUP120 gene. J Cell Biol *131*, 1677-1697.

Heessen, S., and Fornerod, M. (2007). The inner nuclear envelope as a transcription factor resting place. EMBO Rep *8*, 914-919.

Herzog, F., Primorac, I., Dube, P., Lenart, P., Sander, B., Mechtler, K., Stark, H., and Peters, J.M. (2009). Structure of the anaphase-promoting complex/cyclosome interacting with a mitotic checkpoint complex. Science *323*, 1477-1481.

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

Hinshaw, J.E., and Milligan, R.A. (2003). Nuclear pore complexes exceeding eightfold rotational symmetry. J Struct Biol *141*, 259-268.

Hodel, A.E., Hodel, M.R., Griffis, E.R., Hennig, K.A., Ratner, G.A., Xu, S., and Powers, M.A. (2002). The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. Mol Cell *10*, 347-358.

Holm, L., and Sander, C. (1995). Dali: a network tool for protein structure comparison. Trends Biochem Sci *20*, 478-480.

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

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., *et al.* (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947-962.

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

Jeudy, S., and Schwartz, T.U. (2007). Crystal structure of nucleoporin Nic96 reveals a novel, intricate helical domain architecture. J Biol Chem *282*, 34904-34912.

Jia, Y., Dewey, T.G., Shindyalov, I.N., and Bourne, P.E. (2004). A new scoring function and associated statistical significance for structure alignment by CE. J Comput Biol *11*, 787-799.

Kalverda, B., and Fornerod, M. (2007). The nuclear life of nucleoporins. Dev Cell *13*, 164-165.

Kampmann, M., and Blobel, G. (2009). Three-dimensional structure and flexibility of a membrane-coating module of the nuclear pore complex. Nat Struct Mol Biol.

Kilmartin, J.V., and Adams, A.E. (1984). Structural rearrangements of tubulin and actin during the cell cycle of the yeast Saccharomyces. J Cell Biol *98*, 922-933.

Kim, J., Sitaraman, S., Hierro, A., Beach, B.M., Odorizzi, G., and Hurley, J.H. (2005). Structural basis for endosomal targeting by the Bro1 domain. Dev Cell *8*, 937-947.

Kiseleva, E., Allen, T.D., Rutherford, S., Bucci, M., Wente, S.R., and Goldberg, M.W. (2004). Yeast nuclear pore complexes have a cytoplasmic ring and internal filaments. J Struct Biol *145*, 272-288.

Kiseleva, E., Goldberg, M.W., Cronshaw, J., and Allen, T.D. (2000). The nuclear pore complex: structure, function, and dynamics. Crit Rev Eukaryot Gene Expr *10*, 101-112.

Kobe, B., Guncar, G., Buchholz, R., Huber, T., Maco, B., Cowieson, N., Martin, J.L., Marfori, M., and Forwood, J.K. (2008). Crystallography and protein-protein interactions: biological interfaces and crystal contacts. Biochem Soc Trans *36*, 1438-1441.

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.

Kohler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol 8, 761-773.

Kosova, B., Pante, N., Rollenhagen, C., and Hurt, E. (1999). Nup192p is a conserved nucleoporin with a preferential location at the inner site of the nuclear membrane. J Biol Chem 274, 22646-22651.

Kramer, A., Ludwig, Y., Shahin, V., and Oberleithner, H. (2007). A pathway separate from the central channel through the nuclear pore complex for inorganic ions and small macromolecules. J Biol Chem 282, 31437-31443.

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr 60, 2256-2268.

Leksa, N.C., Brohawn, S.G., and Schwartz, T.U. (2009). The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture. Structure 17, 1082-1091.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No 26.

Li, O., Heath, C.V., Amberg, D.C., Dockendorff, T.C., Copeland, C.S., Snyder, M., and Cole, C.N. (1995). Mutation or deletion of the Saccharomyces cerevisiae RAT3/NUP133 gene causes temperature-dependent nuclear accumulation of poly(A)+ RNA and constitutive clustering of nuclear pore complexes. Mol Biol Cell 6, 401-417.

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

Lim, R.Y., Fahrenkrog, B., Koser, J., Schwarz-Herion, K., Deng, J., and Aebi, U. (2007). Nanomechanical basis of selective gating by the nuclear pore complex. Science 318, 640-643.

Lim, R.Y., Ullman, K.S., and Fahrenkrog, B. (2008). Biology and biophysics of the nuclear pore complex and its components. Int Rev Cell Mol Biol 267, 299-342.

Liu, H.L., De Souza, C.P., Osmani, A.H., and Osmani, S.A. (2009). The three fungal transmembrane nuclear pore complex proteins of Aspergillus nidulans are dispensable in the presence of an intact An-Nup84-120 complex. Mol Biol Cell 20, 616-630.

Liu, S.M., and Stewart, M. (2005). Structural basis for the high-affinity binding of nucleoporin Nup1p to the Saccharomyces cerevisiae importin-beta homologue, Kap95p. J Mol Biol 349, 515-525.

Liu, Y., and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. Protein Sci 11, 1285-1299.

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

Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953-961.

Lusk, C.P., Makhnevych, T., Marelli, M., Aitchison, J.D., and Wozniak, R.W. (2002). Karyopherins in nuclear pore biogenesis: a role for Kap121p in the assembly of Nup53p into nuclear pore complexes. J Cell Biol *159*, 267-278.

Lutzmann, M., Kunze, R., Buerer, A., Aebi, U., and Hurt, E. (2002). Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins. Embo J *21*, 387-397.

Lutzmann, M., Kunze, R., Stangl, K., Stelter, P., Toth, K.F., Bottcher, B., and Hurt, E. (2005). Reconstitution of Nup157 and Nup145N into the Nup84 complex. J Biol Chem *280*, 18442-18451.

Makio, T., Stanton, L.H., Lin, C.C., Goldfarb, D.S., Weis, K., and Wozniak, R.W. (2009). The nucleoporins Nup170p and Nup157p are essential for nuclear pore complex assembly. J Cell Biol *185*, 459-473.

Marelli, M., Aitchison, J.D., and Wozniak, R.W. (1998). Specific binding of the karyopherin Kap121p to a subunit of the nuclear pore complex containing Nup53p, Nup59p, and Nup170p. J Cell Biol *143*, 1813-1830.

Matsuoka, Y., Takagi, M., Ban, T., Miyazaki, M., Yamamoto, T., Kondo, Y., and Yoneda, Y. (1999). Identification and characterization of nuclear pore subcomplexes in mitotic extract of human somatic cells. Biochem Biophys Res Commun *254*, 417-423.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

McMahon, H.T., and Gallop, J.L. (2005). Membrane curvature and mechanisms of dynamic cell membrane remodelling. Nature *438*, 590-596.

McMahon, H.T., and Mills, I.G. (2004). COP and clathrin-coated vesicle budding: different pathways, common approaches. Curr Opin Cell Biol *16*, 379-391.

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

Napetschnig, J., Blobel, G., and Hoelz, A. (2007). Crystal structure of the N-terminal domain of the human protooncogene Nup214/CAN. Proc Natl Acad Sci U S A *104*, 1783-1788.

Nehrbass, U., Rout, M.P., Maguire, S., Blobel, G., and Wozniak, R.W. (1996). The yeast nucleoporin Nup188p interacts genetically and physically with the core structures of the nuclear pore complex. J Cell Biol *133*, 1153-1162.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D.G., and Notredame, C. (2004). 3DCoffee: combining protein sequences and structures within multiple sequence alignments. J Mol Biol *340*, 385-395.

Onischenko, E., Stanton, L.H., Madrid, A.S., Kieselbach, T., and Weis, K. (2009). Role of the Ndc1 interaction network in yeast nuclear pore complex assembly and maintenance. J Cell Biol *185*, 475-491.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods in Enzymology *276*, 307-326.

Owen, D.J., Collins, B.M., and Evans, P.R. (2004). Adaptors for clathrin coats: structure and function. Annu Rev Cell Dev Biol *20*, 153-191.

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

Paoli, M. (2001). Protein folds propelled by diversity. Prog Biophys Mol Biol *76*, 103-130.

Partridge, J.R., and Schwartz, T.U. (2009). Crystallographic and biochemical analysis of the Ran-binding zinc finger domain. J Mol Biol *391*, 375-389.

Pemberton, L.F., and Paschal, B.M. (2005). Mechanisms of receptor-mediated nuclear import and nuclear export. Traffic *6*, 187-198.

Pemberton, L.F., Rout, M.P., and Blobel, G. (1995). Disruption of the nucleoporin gene NUP133 results in clustering of nuclear pore complexes. Proc Natl Acad Sci U S A *92*, 1187-1191.

Peters, R. (2009). Translocation through the nuclear pore: Kaps pave the way. Bioessays *31*, 466-477.

Powell, L., and Burke, B. (1990). Internuclear exchange of an inner nuclear membrane protein (p55) in heterokaryons: in vivo evidence for the interaction of p55 with the nuclear lamina. J Cell Biol *111*, 2225-2234.

Presgraves, D.C., Balagopalan, L., Abmayr, S.M., and Orr, H.A. (2003). Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature *423*, 715-719.

Pryer, N.K., Salama, N.R., Schekman, R., and Kaiser, C.A. (1993). Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum in vitro. J Cell Biol *120*, 865-875.

Pucadyil, T.J., and Schmid, S.L. (2009). Conserved functions of membrane active GTPases in coated vesicle formation. Science *325*, 1217-1220.

Rabut, G., Doye, V., and Ellenberg, J. (2004). Mapping the dynamic organization of the nuclear pore complex inside single living cells. Nat Cell Biol *6*, 1114-1121.

Rasala, B.A., Orjalo, A.V., Shen, Z., Briggs, S., and Forbes, D.J. (2006). ELYS is a dual nucleoporin/kinetochore protein required for nuclear pore assembly and proper cell division. Proc Natl Acad Sci U S A *103*, 17801-17806.

Ratner, G.A., Hodel, A.E., and Powers, M.A. (2007). Molecular determinants of binding between Gly-Leu-Phe-Gly nucleoporins and the nuclear pore complex. J Biol Chem *282*, 33968-33976.

Reichelt, R., Holzenburg, A., Buhle, E.L., Jr., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. J Cell Biol *110*, 883-894.

Ren, Y., Yip, C.K., Tripathi, A., Huie, D., Jeffrey, P.D., Walz, T., and Hughson, F.M. (2009). A structure-based mechanism for vesicle capture by the multisubunit tethering complex Dsl1. Cell *139*, 1119-1129.

Rost, B., Yachdav, G., and Liu, J. (2004). The PredictProtein server. Nucleic Acids Res *32*, W321-326.

Rout, M.P., Aitchison, J.D., Magnasco, M.O., and Chait, B.T. (2003). Virtual gating and nuclear transport: the hole picture. Trends Cell Biol *13*, 622-628.

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol *148*, 635-651.

Rout, M.P., and Blobel, G. (1993). Isolation of the yeast nuclear pore complex. J Cell Biol *123*, 771-783.

Schrader, N., Koerner, C., Koessmeier, K., Bangert, J.A., Wittinghofer, A., Stoll, R., and Vetter, I.R. (2008a). The crystal structure of the Ran-Nup153ZnF2 complex: a general Ran docking site at the nuclear pore complex. Structure *16*, 1116-1125.

Schrader, N., Stelter, P., Flemming, D., Kunze, R., Hurt, E., and Vetter, I.R. (2008b). Structural basis of the nic96 subcomplex organization in the nuclear pore channel. Mol Cell *29*, 46-55.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J *78*, 1606-1619.

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

Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr A *64*, 112-122.

Shibata, Y., Hu, J., Kozlov, M.M., and Rapoport, T.A. (2009). Mechanisms Shaping the Membranes of Cellular Organelles. Annu Rev Cell Dev Biol *25*, 14.11-14.26.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19-27.

Siniossoglou, S., Lutzmann, M., Santos-Rosa, H., Leonard, K., Mueller, S., Aebi, U., and 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., and 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., LaPointe, P., Razvi, A., Gurkan, C., Potter, C.S., Carragher, B., and Balch, W.E. (2008). Structural basis for cargo regulation of COPII coat assembly. Cell *134*, 474-484.

Stewart, M. (2007). Ratcheting mRNA out of the nucleus. Mol Cell *25*, 327-330.

Stoffler, D., Feja, B., Fahrenkrog, B., Walz, J., Typke, D., and Aebi, U. (2003). Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport. J Mol Biol *328*, 119-130.

Strambio-de-Castillia, C., Blobel, G., and Rout, M.P. (1999). Proteins connecting the nuclear pore complex with the nuclear interior. J Cell Biol *144*, 839-855.

Strokopytov, B.V., Fedorov, A., Mahoney, N.M., Kessels, M., Drubin, D.G., and Almo, S.C. (2005). Phased translation function revisited: structure solution of the cofilin-homology domain from yeast actin-binding protein 1 using six-dimensional searches. Acta Crystallogr D Biol Crystallogr *61*, 285-293.

Suetsugu, S., Toyooka, K., and Senju, Y. (2009). Subcellular membrane curvature mediated by the BAR domain superfamily proteins. Semin Cell Dev Biol.

Sun, Y., and Guo, H.C. (2008). Structural constraints on autoprocessing of the human nucleoporin Nup98. Protein Sci *17*, 494-505.

Sztul, E., and Lupashin, V. (2009). Role of vesicle tethering factors in the ER-Golgi membrane traffic. FEBS Lett *583*, 3770-3783.

Tang, S., and Presgraves, D.C. (2009). Evolution of the Drosophila nuclear pore complex results in multiple hybrid incompatibilities. Science *323*, 779-782.

ter Haar, E., Harrison, S.C., and Kirchhausen, T. (2000). Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin. Proc Natl Acad Sci U S A *97*, 1096-1100.

Terry, L.J., and Wente, S.R. (2007). Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. J Cell Biol *178*, 1121-1132.

Tran, E.J., and Wente, S.R. (2006). Dynamic nuclear pore complexes: life on the edge. Cell *125*, 1041-1053.

Vetter, I.R., Nowak, C., Nishimoto, T., Kuhlmann, J., and 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.

von Moeller, H., Basquin, C., and Conti, E. (2009). The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. Nat Struct Mol Biol *16*, 247-254.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol Biol *364*, 215-230.

Walther, T.C., Alves, A., Pickersgill, H., Loiodice, I., Hetzer, M., Galy, V., Hulsmann, B.B., Kocher, T., Wilm, M., Allen, T., *et al.* (2003). The conserved Nup107-160 complex is critical for nuclear pore complex assembly. Cell *113*, 195-206.

Wang, H., Kakaradov, B., Collins, S.R., Karotki, L., Fiedler, D., Shales, M., Shokat, K.M., Walther, T.C., Krogan, N.J., and Koller, D. (2009). A complex-based reconstruction of the Saccharomyces cerevisiae interactome. Mol Cell Proteomics *8*, 1361-1381.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Watson, M.L. (1959). Further observations on the nuclear envelope of the animal cell. J Biophys Biochem Cytol *6*, 147-156.

Weirich, C.S., Erzberger, J.P., Berger, J.M., and Weis, K. (2004). The N-terminal domain of Nup159 forms a beta-propeller that functions in mRNA export by tethering the helicase Dbp5 to the nuclear pore. Mol Cell *16*, 749-760.

Weis, K. (2003). Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. Cell *112*, 441-451.

Whittle, J.R., and Schwartz, T.U. (2009). Architectural nucleoporins Nup157/170 and Nup133 are structurally related and descend from a second ancestral element. J Biol Chem *284*, 28442-28452.

Winey, M., Hoyt, M.A., Chan, C., Goetsch, L., Botstein, D., and Byers, B. (1993). NDC1: a nuclear periphery component required for yeast spindle pole body duplication. J Cell Biol *122*, 743-751.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., *et al.* (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901-906.

Xu, S., and Powers, M.A. (2009). Nuclear pore proteins and cancer. Semin Cell Dev Biol *20*, 620-630.

Xu, Y., Xing, Y., Chen, Y., Chao, Y., Lin, Z., Fan, E., Yu, J.W., Strack, S., Jeffrey, P.D., and Shi, Y. (2006). Structure of the protein phosphatase 2A holoenzyme. Cell *127*, 1239-1251.

Xylourgidis, N., and Fornerod, M. (2009). Acting out of character: regulatory roles of nuclear pore complex proteins. Dev Cell *17*, 617-625.

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

Young, A. (2007). Structural insights into the clathrin coat. Semin Cell Dev Biol *18*, 448-458.

Yu, H., Braun, P., Yildirim, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., *et al.* (2008). High-quality binary protein interaction map of the yeast interactome network. Science *322*, 104-110.

Zabel, U., Doye, V., Tekotte, H., Wepf, R., Grandi, P., and Hurt, E.C. (1996). Nic96p is required for nuclear pore formation and functionally interacts with a novel nucleoporin, Nup188p. J Cell Biol *133*, 1141-1152.

Zimmerberg, J., and Kozlov, M.M. (2006). How proteins produce cellular membrane curvature. Nat Rev Mol Cell Biol *7*, 9-19.

Zuleger, N., Korfali, N., and Schirmer, E.C. (2008). Inner nuclear membrane protein transport is mediated by multiple mechanisms. Biochem Soc Trans *36*, 1373-1377.