Structural and Biochemical Characterization of Nuclear Pore Complex Structural Scaffold sub-Complexes

by

Kotaro Kelley
B.A., University of California Berkeley
Berkeley, CA 2009

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

APRIL 2017 [September 2017]

Signature redacted

Signature of Author.... ....

Kotaro Kelley
Department of Biology
Julyl 1, 2017

Signature redacted

Certified by......... ....

Thomas U. Schwartz
Professor of Biology
Thesis Supervisor

Signature redacted

Accepted by........... ....

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

Structural and Biochemical Characterization of Nuclear Pore Complex Structural Scaffold sub-Complexes

Kotaro Kelley

Abstract

The nuclear pore complex (NPC) is a large, modular protein assembly that regulates nucleocytoplasmic transport in all eukaryotes. The ~60-120 MDa NPC is a modular assembly of multiple copies of ~30 distinct proteins that are arranged into biochemically distinct sub-complexes. We believe that the structural characterization of the NPC is essential for understanding its transport mechanisms and various pathologies and human diseases associated with deletions or mutations of constituents. To obtain detailed structural information of the NPC, techniques that span several resolution ranges are necessary due to its large size and complexity. For instance, recent progress in the structural characterization of the overall architecture of the NPC by cryo-electron tomography(cryo-ET) to ~23Å resolution has revealed its size, shape, and course arrangement, but lacks distinguishable protein-protein boundaries and secondary structural details. Although the entire NPC is not amenable to high resolution X-ray crystallography, we complement the cryo-ET reconstructions with a divide and conquer approach by obtaining high resolution X-ray crystal structures of individual sub-complexes. By taking advantage of the modular nature of the NPC, we can dock sub-complexes into the cryo-ET reconstructions to identify their location within the NPC. This composite structure will bridge the meso resolution cryo-ET reconstructions of the entire NPC and the incomplete but high resolution X-ray crystal structure of individual sub-complexes.

As a first step towards understanding the detailed organization of the NPC, our goal is to solve the high resolution structures of the two principal structural scaffold sub-complexes, the Y and Nic96 complexes. In this study, we present the high resolution composite X-ray crystal structure of the Y complex. Docking the composite model into previously solved random conical tilt(RCT) and tomographic reconstructions of negatively stained samples of the Y complex shows overall consistency between the three methods, yet we highlight important structural differences that constrain possible arrangements of multiple Y complexes within the NPC. By docking the composite model into the cryo-ET reconstructions of the entire NPC, we propose an arrangement of multiple Y complexes that is consistent with our composite structure. In addition, progress on structurally characterizing the Nic96 complex will be presented. Preliminary results suggest that Nup192 and Nic96 form a flexible, yet semi-ordered interface. Future directions for characterizing the rest of the Nic96 complex, including current challenges and suggestions will be discussed.

Thesis Supervisor: Thomas U. Schwartz
Title: Professor of Biology

Table of Contents

Chapter 2: Atomic structure of the Y complex of the nuclear pore

Chapter 3: Structural and Biochemical studies of the Nic96 complex

Chapter 4: Conclusions

Acknowledgments

The 6 years that I have spent at MIT have been some of the most challenging and exciting time of my life. Throughout this journey, I have been supported generously by many people. I would like to thank the following people in particular.

To my former lab mates, especially Kevin and Brian. Kevin worked with me on my main thesis project and together we shared the exciting process of publishing our results. Brian patiently trained me when I first joined the lab, a time when I had very little experience. Much of my success at bench can be attributed to my interactions with both Kevin and Brian.

To my advisor Thomas, I thank you for letting me join your lab. I appreciate that your door was always open for discussions. I've learned a great deal about crystallography and structural biology in general from you. I leave your lab as a better and more confident scientist than when I first joined. Thanks for your continued support!

Chapter 1: Introduction

Nuclear pore complex

The eukaryotic cell physically partitions chemical and mechanical processes with membrane bound compartments known as organelles. Specialization of processes within organelles allows the eukaryotic cell to have greater control over its resources by sequestering or releasing components in a temporal and spatial manner. One such partition is between the nucleus, which houses the cell's genetic information as DNA, with its associated replication and RNA transcriptional machinery, and the cytoplasm, which is the site of translation of all proteins. The separating double layer phospholipid membrane known as the nuclear envelope allows the accumulation of molecules against their concentration gradient on either side of the partition. Since the nucleus requires specialized protein molecules to carry out both replication and transcription, and the cytoplasm needs transcriptional signals generated in the nucleus for protein synthesis, there needs to be a method of communication between the two compartments. The primary conduit of regulated molecular exchange between the nucleus and the cytoplasm is the nuclear pore complex (NPC).

NPCs are located at circular openings in the nuclear envelope which are maintained by a cage like protein scaffold (figure 1; figure 3). Regulated transport is controlled by flexible phenylalanine glycine rich FG-repeat proteins that dock onto the NPC scaffold to form a porous, yet size-selective hydrogel that fills the center of the NPC. The hydrogel is formed through Van der Waals interactions between neighboring FG-repeat proteins (Hülsmann et al., 2012; Frey et al., 2006; Ribbeck and Görlich, 2001; figure 2). Small molecules and protein and RNA that are smaller than ~40 kDa are able to freely diffuse through the FG-repeat barrier (Paine et al., 1975; Bonner, 1978).

Bi-directional transport of larger proteins require a nuclear localization signal (NLS) which are bound by FG-repeat barrier penetrating nuclear transport receptors (NTRs), known as karyopherins (Cook et al., 2007; figure 1). NTRs can simultaneously bind their cargo and disrupt the self interactions of the FG-repeat hydrogel. The cargo-receptor complex is then able to diffuse through the hydrogel (Hülsmann et al., 2012; Ribbeck and Görlich, 2001; figure 2). RNA synthesized in the nucleus, destined for the cytoplasm, are exported by a similar but distinct mechanism (Cook et. al, 2007).

In order to maintain a concentration gradient of molecules across the nuclear envelope, the ratio between import and export must be biased away from thermodynamic equilibrium. In order to achieve this, an uneven pool of GTP and GDP bound GTPase Ran is maintained across the nuclear envelope with the energy to do so coming ultimately from GTP hydrolysis. GTP bound Ran (Ran-GTP) is more abundant in the nucleus, whereas the GDP bound Ran (Ran-GDP) is more abundant in the cytoplasm. In the cytoplasm, GTP hydrolysis is promoted by the NPC associating Ran GTPase activating protein (RanGAP) and in the nucleus, nucleotide exchange of GDP for GTP is facilitated by a chromatin associating guanine nucleotide exchange factor known as regulator of chromosome condensation 1 (RCC1). Directional transport is achieved by cargo release stimulated by interactions with Ran-GTP in the nucleus and Ran-GDP in the cytoplasm (Cook et al., 2007; figure 1).

The NPC has a diameter and height of ~100 nm, an estimated MW of ~60-120 MDa, and is located at circular openings where the outer and inner nuclear membrane fuse (Rout and Blobel, 1993; Reichelt et al., 1990; Bui et al., 2013; figure 3). About 30 distinct proteins, known as nucleoporins (Nups) build the NPC. More than 20 of them are universally conserved across all eukaryotes (Rout et al., 2000; Cronshaw et al., 2002). NPCs show an eightfold rotational symmetry in EM images (Akey and Radermacher, 1993; Rout and Blobel, 1993), suggesting that all nucleoporins occur in multiples of eight. The NPC has a modular design in that the ~30 distinct proteins are assigned to several biochemically distinct sub-complexes that form higher order assemblies. Stable sub-complexes, as judged by biochemical properties, are the Y complex, the Nic96 complex, the Nsp1 complex, the Nup82 complex, the Ndc1 complex, and the nuclear basket.

The NPC serves a central role in eukaryotic cell biology. In order to understand its transport mechanisms and various pathologies and human diseases associated with deletions or mutations in constituents, the structural characterization of the NPC is important. To obtain detailed structural information of the NPC, techniques that span several resolution ranges are necessary due to its large size and complexity. For instance, recent progress from several groups on the structural characterization of the overall architecture of the NPC by cryo-electron tomography (cryo-ET), ranging from ~30 to ~20 Å resolution, has revealed its size, shape, and a stacked arrangement of three distinct protein dense rings referred to as the cytoplasmic, inner, and nuclear ring, but lacks distinguishable protein-protein boundaries and secondary structural details (Bui et al., 2013; Eibauer et al., 2014; von Appen et al.,2015; figure 3). Although the entire NPC is not amenable to high-resolution X-ray crystallography, our lab and others have complemented the cryo-ET reconstructions with high resolution X-ray crystal structures of individual sub-complexes. By taking advantage of the symmetry and modularity of the NPC, we can dock crystal structures of sub-complexes into the cryo-

ET reconstructions to identify their location within the NPC. This composite structure will bridge the meso resolution cryo-ET reconstructions of the entire NPC and the incomplete but high-resolution X-ray crystal structure of individual sub-complexes.

As a first step towards understanding the detailed organization of the NPC, our lab has focused on solving the high resolution structures of its two principal structural scaffold sub-complexes, the Y and Nic96 complexes. Together, these two sub-complexes are estimated to make up ~55% of total mass of the NPC and are expected to contribute significantly to the visible density in the cryo-ET reconstructions. The following review will briefly summarize the available structural information about the two sub-complexes and serve as a foundation for the original work presented later in this thesis.

Y complex

The ~0.6 MDa Y or scNup84 complex obtained its name from the letter Y shape observed in electron micrographs of negatively stained, purified samples from *S. cerevisiae* (sc). The same study mapped the approximate connectivity of the Y complex constituents by sequentially imaging sub-assemblies (Lutzmann et al., 2002). The long arm is composed of Nup133, Nup84, Nup145C, and Sec13. Nup120 forms one of the two short arms, while Nup85 and Seh1 form the other (figure 5). Immuno-gold labeling electron microscopy suggests that the Y complex is symmetrically located at the edges of both the cytoplasmic and nuclear side of the NPC, corresponding to the cytoplasmic and nuclear rings that were revealed in the cryo-ET reconstructions (Rout et al., 2000; Bui et al., 2013).

The absolute stoichiometry of the Y complex is a hotly debated topic, but most in the field have converged on either 16 or 32 copies, split equally between the cytoplasmic and nuclear side. Quantitative immunoblotting of yeast and quantitative SDS-PAGE of rat liver nuclei estimated 16 copies of Y complexes per NPC (Rout et al., 2000; Cronshaw et al., 2002). In another study, counting photobleaching events of GFP-tagged Nups in yeast with SPEED microscopy estimated 16 copies of Y complexes per NPC (Mi et al., 2015). However, quantitative mass spectrometry and super-resolution microscopy of human cells estimated 32 copies of Y complexes per NPC (Ori et al., 2013). It is possible that stoichiometry is variable across organisms, with a corresponding difference in NPC mass and structure, but a systematic study of a single stoichiometry determination method across several organisms has yet to be reported.

Deletion, depletion, or mutation of Y complex Nups leads to severe defects in the NPC, including assembly defects and diminished protein and RNA transport. A number of human diseases can be attributed to such NPC pathologies. As an example, a heterozygous knock out of hsNup96 (scNup145N) specifically affects the nuclear transport of interferon-regulated gene transcripts, thereby disrupting the normal function of macrophages, and B and T lymphocytes (Faria et al., 2006). Perhaps the most dramatic demonstration of the Y complex's role in maintaining the integrity of the NPC as a structural scaffold comes from nuclear reconstitution assays with *Xenopus* egg extract. In the presence of the Y complex, scanning electron microscopy of the reconstituted nuclei shows distinct ring-like structures corresponding to intact NPCs. Immuno-depleting Y complex constituents result in smooth, NPC free nuclei, strongly suggesting that the Y complex is necessary for maintaining the structure of the NPC (Harel et al., 2003).

So far, the Y complex is the best structurally characterized sub-complex of the NPC. There are multiple low resolution structures from negatively-stained yeast samples, to random conical tilt (RCT) and tomographic reconstructions of negatively-stained yeast and human Y complexes, respectively (Lutzman et al., 2005; Kampmann and Blobel, 2009; Bui et al., 2013). These studies collectively suggested that the Y complex is very flexible, particularly along the long arm, such that it is averaged out in the tomographic reconstruction of the human Y complex. This would suggest that the entire Y complex is unlikely to crystallize and indeed to this date, such conditions have not been reported in the literature. However, in a concurrent effort, X-ray crystal structures of fragments and sub-assemblies of the Y complex were obtained by our lab and others.

Crystal structures of Nup120 reveal an N-terminal 7-bladed β-propeller integrated with an irregular α-helical stack that connects to the β-propeller through a helical bundle insertion between blades 6 and 7 and a C-terminal α-helical domain made up of HEAT repeats (Bilokapic and Schwartz, 2012; Leksa et al., 2009; Seo et al., 2009; Liu et al., 2012). A 90° kink between the rigid N-terminal and C-terminal domains gives Nup120 a bent, claw-like structure (Bilokapic and Schwartz, 2012). Size exclusion chromatography experiments have shown that the very C terminus of Nup120 is necessary and sufficient for binding to the C termini of both Nup145C and Nup85 (Bilokapic and Schwartz, 201). Although it has been shown that the C terminus of *S. pombe,* spNup145C, can directly interact with the C terminus of spNup85, the primary interactions at the junction of the three arms of the Y complex are mediated through Nup120 (Bilokapic and Schwartz, 2012).

Crystal structures of Nup145C, Nup85, and Nup84 share a common 65 kDa domain consisting of 28 α-helices. Because of its remarkable structural similarity to the COPII vesicle coat component Sec31, it was named the Ancestral Coatomer Element 1 (ACE1) (Brohawn et al.,2008; figure 6). The irregular α-helical stack of the ACE1 domain forms a pleated U shape and consists of a crown, trunk, and tail module. Limited proteolysis of the ACE1 domain showed that the 3 modules are connected by flexible hinges; at least partly accounting for the flexibility observed in the EM structures. Functionally, the crown and tail modules of ACE1 proteins tend to be used for binding other proteins. For example, the crown modules of Nup145C and Nup84 interact to form the long arm of the Y (Brohawn et al., 2009). The C-terminal tail module of ACE1 proteins from the NPC form an extended, irregular α-helical stack, when compared to the stubbed Sec31 C terminus (Brohawn et al.,2008; figure 6). The junction between the trunk and tail module is identified as being more flexible than the crown/trunk junction due to the tail not being part of the pleated U shape. The N-terminal ends of ACE1 proteins are often elaborated. Both Nup85 and Nup145C contain an N-terminal β-sheet that contributes a 7th blade to the open, 6-bladed β-propellers Seh1 and Sec13, respectively. The Sec31 N-terminal elaboration similarly contributes an insertion blade to Sec13, followed by its own β-propeller (Fath et al., 2007). The COPII vesicle coat forms a cage-like structure with edges formed by a domain swapped Sec31 dimer and vertices formed by 4 pair of Sec13/Sec31 β-propellers (Stagg et al., 2006; Fath et al., 2007). Sequence conservation between Nup145C, Nup85, Nup84, and Sec31 is so poor that this specific structural relationship was not revealed previous to solving their crystal structures (Brohawn et al., 2008). Despite this, comparisons between their crystal structures derived from different organisms show a remarkable amount of structural conservation. It is interesting to see if these commonalities translates to a similar design principle between higher-order assemblies of the Y complex and the cage-like scaffold structure of the COPII vesicle coat.

Nup145C is the C-terminal product of Nup145 auto proteolysis (Hodel et al., 2002). The N-terminal product, Nup145N, starts with a hydrogel forming N-terminal FG-repeat domain, followed by an unstructured region and a C-terminal globular auto-proteolytic domain (APD). After cleavage, Nup145N APD remains attached to Nup145C (Hodel et al., 2002). However, Nup82 from the Nup82 complex binds competitively to the Nup145N APD (Yoshida et al., 2011).

ACE1 protein Nup84 does not have a N-terminal elaboration, but is responsible for 1/3 of the long arm and serves as an important connector between Nup133 and the rest of the Y complex. Crystal structures of the crown and trunk of Nup84 (Brohawn and Schwartz, 2009) and the Nup133 binding tail (Boehmer et al., 2008) module have been solved separately, but there are 4 unsolved helices at the junction between the trunk and tail. Nup133 has an N-terminal 7-bladed β-propeller that connects to a large, extended, irregular α-helical domain, through a flexible linker. The C-terminal α-helical stack meets with the C-terminal tail module of Nup84 to form a large binding interface (Berke et al., 2004; Boehmer et al., 2008; figure 5). The α-helical stack domain of Nup133 is loosely connected via flexible hinges which likely contributes to the flexibility of the long arm of the Y complex (Whittle and Schwartz, 2009).

Many fragments of each arm of the Y complex have been characterized crystallographically, but there are still missing sections (figure 5). The yet unknown atomic architecture of the junction between Nup120 C terminus, Nup85 C-terminal tail module, and Nup145C C-terminal tail module, from here on collectively referred to as the Y complex "hub," holds crucial information about the three-dimensional arrangement of the three Y complex arms. This is critical not only for understanding the structure of the Y complex, but more importantly, for any further consideration of how multiple Y complexes arrange in a lattice to generate the NPC structural scaffold. Although the low resolution EM structures of the Y complex should in principle be sufficient for answering this question, there are a few reasons to seek a higher resolution crystal structure. First, how the two unsolved ACE1 tail modules look and how they incorporate with the C terminus of Nup120 would be immediately apparent from a crystal structure. Second, Y complexes adopt a preferred orientation on the carbon support of an EM grid, largely due to their extended, distinctly non-globular shape. In addition, it is unclear to what extend the grid deposition distorts the Y complex structure. Therefore, a crystal structure should provide information about the hub architecture without the limitations of EM structures.

Based on sequence conservation, the minimal Y complex has long been assumed to contain the 7 Nups mentioned above. However, it is known that the Y complex has additional, species specific Nups that dock onto one of the 3 arms, including Nup37, Nup43, and ELYS in *Homo sapiens* (hs) (Bilokapic and Schwartz, 2012; Bilokapic and Schwartz, 2013). They are not believed to affect the overall structure of the minimal Y complex. Instead, they likely decorate the Y complex to serve species specific functions. For metazoa, the 7-bladed β-propeller Nup37 binds stably to Nup120's 90° kink between its N and C-terminal domains (Bilokapic and Schwartz; 2012, figure 5). hsNup43, another 7-bladed β-propeller, attaches to the human Y complex through interactions with hsNup85 (Loïodice et al., 2004; Kim et al., 2014a). ELYS is a large ~250 kDa protein that contains an N-terminal 7-bladed β-propeller, followed by an α-helical domain that interacts with Nup120 and Nup37, and a disordered C-terminal domain (Rasala et al., 2006; Bilokapic and Schwartz, 2012; Bilokapic and Schwartz, 2013). The disordered C terminus of ELYS binds to AT-rich chromatin during open mitosis to recruit vesicles containing transmembrane anchoring Nups Pom121 and Ndc1 to initiate NPC assembly (Rasala et al., 2008). Although unicellular eukaryotes such as *S. pombe* (sp) *and M. thermophila* (mt), have an ELYS homolog, it has a much smaller molecular weight and lacks the disordered C-terminal domain and the N-terminal β-propeller (Bilokapic and Schwartz, 2013). Lower eukaryotes that undergo closed mitosis, such as *S. cerevisiae,* lack ELYS, perhaps due to mechanistic differences between open and closed mitosis. Whether or not these species specific additions to the Y complex affects the overall dimensions and assembly of the NPC is not yet known.

Akey and co-workers have estimated a vertebrate NPC diameter of ~140 nm and heigh of ~80 nm from cryo-RCT reconstructions of nuclear envelope embedded *Xenopus* NPCs (Akey and Radermacher, 1993). In a separate study, they estimated a yeast NPC diameter of ~95 nm and height of ~35 nm from RCT reconstruction of isolated and negatively stained NPCs; a ~45 nm difference in both dimensions (Yang et al., 1998). A similar height difference of ~40nm, but roughly similar diameter of ~120nm, has also been observed in cryo-ET reconstructions of intact *Dictyostelium discoideum* nuclei and native *Xenopus* nuclear envelope spreads (Elad et al., 2009). This suggests that the inconsistent diameter estimates from the earlier studies of isolated yeast NPCs may be due to technical reasons stemming from harsh purification procedures. The difference in height between vertebrate and non-vertebrate NPCs is reflective of their mass estimates. The yeast NPC has been estimated to be ~45-70 MDa whereas the *Xenopus* NPC, ~125 MDa (Rout and Blobel, 1993; Rout et al., 2000; Cronshaw et al., 2002; Reichelt et al., 1990). One tempting hypothesis is that the species specific additions to the Y complex and or different Y complex stoichiometries leads to the discrepancy in dimensions.

The higher order assembly of Y complexes, forming the structural scaffold of the NPC, is a critical structure. Unfortunately, purified Y complexes are not known to self-assemble into scaffold structures like Sec31 and Sec13 assemble into COPII vesicle coats. Consequently, reports of higher order inter-Y complex interactions have been few (Seo et al., 2009; Thierbach et al., 2013). Over the years, this has led to the proposal of many mutually exclusive models for higher order assemblies of Y complexes. Based on structural conservations between ACE1 proteins, our lab has proposed a COPII vesicle coat-like “lattice” model (Brohawn et al., 2008; figure 7, C). Based on crystal contacts observed between adjacent molecules in the unit cell, the Blobel lab has proposed the “fence-post” model. In this model, 4 rings of 8 Y complexes are stacked on top of each other to coat the membrane. Adjacent rings are connected by alternating heterooctameric poles of Nup145C/Sec13 and Nup85/Seh1 (Debler et al., 2008; Hsia et al., 2007; figure 7, A). A computational method incorporating a large amount of complementary volumetric, stoichiometric, and co-immunoprecipitation based distance restraints has led to the proposal of a nuclear and cytoplasmic ring of 8 Y complexes arranged in a head-to-tail orientation where the end of the long arm (Nup133) interacts with the short arms of an adjacent Y complex (Alber et al., 2007a; Alber et al., 2007b; figure 7, B). More recently, Beck and coworkers have attempted to computationally dock a 35 Å ET reconstruction of negatively stained human Y complexes into their approximately C2 symmetric 32 Å cryo-ET reconstruction of the entire human NPC (Bui et al., 2013; figure 3; figure 7, D). Based on their fitting procedure, the authors propose two parallel rings of 8 Y complexes in a head-to-tail arrangement, occupying both the nuclear and cytoplasmic rings of the NPC, for a total of 32 copies per NPC. The head-to-tail patterns of Ys in the parallel inner and outer rings are off register with respect to each other and predicts non-equivalent contacts between them (Bui et al., 2013). In a subsequent report, Medalia and coworkers have generated a 20 Å cryo-ET reconstruction of the entire *Xenopus* NPC. They observe that the C2 symmetry reported for the human NPC no longer holds at this improved resolution. The authors applied the same docking procedure of Beck and coworkers by fitting the 35 Å ET reconstruction of negatively stained human Y complexes into their *Xenopus* structure and conclude that their model similarly accommodates two parallel rings of Y complexes, but only for the cytoplasmic ring; noting that they were unable to reliably dock the Y complex in the nuclear ring in any configuration. The authors also observed potential discrepancies in the composition of the Y complex from the inner and outer ring (Eibauer et al., 2015; figure 3). Given the relatively low resolution of both search model and target, the docking procedures yield many putative solutions which have to be ruled out based on prior information about the assumed dimensions and shape of the Y complex. In this particular case, the human Y complex model used is missing half of its long arm and is artificially flattened at the hub where the 3 arms meet. Therefore, possible docking solutions are less restrained than for the "true" structure of the Y complex, and hence should be viewed as tentative at best. Fitting higher resolution crystal structures of fragments of the Y complex is also unreliable since distinct α-helical stacks and β-propellers are difficult to distinguish at 20 Å resolution. Complicating this analysis further is that a lot of density remains unattributed in these docking attempts (Bui et al., 2013). Two obvious improvements to the reliability of this method are a better resolved cryo-ET reconstruction of the NPC and a more accurate model of the Y complex. Better cryo-ET maps may be expected from purer sample preparations, better data collection strategies, and better instrumentation. Generating a better Y complex structure is another important goal, and will be addressed by work presented in this thesis. We were able to obtain a structure of the Nup120-Nup85-Nup145C-Sec13 Y complex hub which was combined with previously solved, overlapping fragments, to construct a complete, composite atomic structure of the entire Y. A docking analysis is then performed with the composite structure, carefully considering solutions that agree with the cyro-ET reconstruction density and flexibility intrinsic to the Y complex.

Nic96 complex

The ~0.5 MDa scNic96 (hsNup93) complex is believed to be the second structural scaffold sub-complex serving to maintain the structural integrity of the NPC. Based on immuno-gold labeling EM, it has been localized to the middle of the NPC transport axis, corresponding to the inner ring density in the cryo-ET reconstructions of the NPC (Rout et al., 2000). The stable members of the Nic96 complex are Nic96, Nup157/170, Nup53/59, Nup188 and Nup192 (Amlacher et al., 2013; figure 8). The redundant scNup157/170 and scNup53/59 paralogs are a result of a genome duplication event during yeast evolution, but exist as a single copy in metazoa (Wolfe and Shields, 1997). Despite the modest sequence identity between scNup53 and scNup59 (33%) and between scNup157 and Nup170 (44%), they are structurally very similar, although whether or not they are functionally redundant is not yet known. Nup188 and Nup192 are structurally very similar and mutually exclusively incorporate into the Nic96 complex (Amlacher et al., 2011). Although crystal structures of the folded regions of individual Nic96 complex members are available (Andersen et al., 2013; Stuwe et al, 2014; Jeudy and Schwartz, 2007; Schrader et al., 2008; Whittle and Schwartz, 2009; Seo et al., 2013; Handa et al., 2006; figure 8), no overall structure has been reported. Nic96 is another ACE1 protein with a crown, trunk, and extended C-terminal tail module (Jeudy and Schwartz, 2007; Schrader et al., 2008). Its N-terminal elaboration has two α-helices that are separated from the ACE1 domain by a long flexible linker (figure 8). The first binds to the FG-barrier forming Nsp1 complex and the second binds to the very C terminus of Nup192 or Nup188 (Stuwe et al., 2015). Thus, Nic96 functions to dock the selective FG-barrier to the structural scaffold of the NPC. Although there are conserved patches in both the crown and tail module, it is not known whether Nic96 binds to other proteins using either module, which would be analogous to the binding mode of ACE1 proteins from the Y complex and COPII vesicle coats (Jeudy and Schwartz, 2007).

Nup157/170 has an N-terminal 7-bladed β-propeller with unusual α-helical insertions in the inter-blade loops. Three of these inter-blade helices form an irregular α-helical stack with the rest of the C-terminal α-helical domain. Nup157/170 is a structural homolog of scNup133 from the Y complex (Whittle and Schwartz, 2009), yet the tight integration of its β-propeller and α-helical domains contrasts with the flexibly tethered β-propeller of Nup133 (Seo et al., 2013; Berke et al., 2004).

2-D EM class averages of negatively stained samples from *M. thermophila* Nup188, yeast Nup192, and C. *thermophila* Nup192 show a characteristic S-shape that is structurally similar to NTRs (Amlacher et al., 2011; Flemming et al., 2012; Sampathkumar et al., 2013; Andersen et al., 2013). Like NTRs, the flexible S-shape of Nup188 and Nup192 adopts several conformations. It can be roughly categorized into an 'open' or 'closed' conformation and has been observed in 2-D EM class averages and in solution by small angle X-ray scattering (SAXS) (Sampathkumar et al., 2013; Andersen et al., 2013). Partial crystal structures of the N and C termini of Nup192 and Nup188, as well as a complete structure of Nup192, show the atomic details of the S-shape fold (Sampathkumar et al., 2013; Andersen et al., 2013; Stuwe et al., 2014; Lin et al., 2016). Nup188 starts with a N-terminal irregular α-helical domain that forms a closed, right-handed superhelical ring, followed by a C-terminal regular α-helical repeat domain. Interestingly, an Src-homology-like (SH3-like) motif is inserted within the N-terminal domain of Nup188. This motif is absent from Nup192, but the overall structure is very similar to Nup188 (Andersen et al., 2013; Stuwe et al., 2014). The two copies of ctNup192-N found in the asymmetric unit of its crystal were in two different "open" conformation. Upon closer inspection, they were found to be related by a rigid body motion around hinge helices that bisect the α-helical ring into two halves (Stuwe et al., 2014).

scNup53 and scNup59 are mostly unstructured except for a homodimerization domain located in the middle (Handa et al., 2006). Nup192 and Nic96 bind to the N-terminal disordered region of Nup53 (Amlacher et al., 2011; Stuwe et al., 2014). Nup170/Nup157 and the karyopherin Kap121 bind mutually exclusively to the C-terminal disordered region (Amlacher et al., 2011; Lusk et al., 2002). At the very C terminus, an amphipathic α-helix binds Ndc1 and inserts into the nuclear membrane (Marelli et al., 2001; Onischenko et al., 2009). The dimerization domain of Nup53 suggests that the entire Nic96 assembly dimerizes. However, it is known that Nup53 from *C. thermophila* and *M. thermophila* exits as a monomer even at high concentrations (Lin et al., 2016). It is possible that the thermophilic Nic96 complex dimerizes by other means, but it is unlikely to be essential for the basic assembly and function of the NPC, given that the double deletion of Nup53 and Nup59 in yeast is viable (Marelli et al., 1998).

The most complete assembly of the Nic96 complex to date comprising Nic96, Nup170, Nup53, and Nup192, was reported using *C. thermophila* proteins (Amlacher et al., 2011). Although individual crystal structures of the Nic96 complex members cover a large part of their folded regions, no higher order assembly structures capturing the binding interfaces have been solved. Therefore, the overall structure of the Nic96 complex remains elusive. A low resolution RCT reconstruction of the trimeric Nic96, Nup192, Nup53 complex does show extra density protruding from the S-shape of Nup192, but due to its low resolution, assignment of Nic96 is tentative at best and

Nup53 is unresolved (Amlacher et al., 2011). Based on preliminary work presented in this thesis, we demonstrate that Nic96 is located at the very C terminus of Nup192, in contrast to previous reports (Amlacher et al., 2011). In addition, to gain high resolution details of the binding interface between Nic96, Nup192, and the Nsp1 complex, crystallization attempts of various assemblies will be presented.

Nsp1 complex

The 3 membered Nsp1 complex consists of Nsp1, Nup49, and Nup57 (Finlay et al., 1991; Brohawn et al., 2009). Immuno-gold labeling EM localizes the Nsp1 complex to the inner ring of the NPC, consistent with the location of its scaffold docking partner, Nic96. The disordered N terminus of each member contains multiple FG-repeats and fills the center of the NPC with its size selective hydrogel barrier (Frey et al., 2006; Hülsmann et al., 2012). The complex self assembles through its C-terminal coiled coils in a 1:1:1 stoichiometry as demonstrated by equilibrium analytical ultracentrifugation (AUC) (Ulrich et al., 2014). Subsequent confirmation of the AUC data came from crystal structures of the entire coiled coil domains of Nup62, Nup58, and Nup54 from *X. laevis* and *C. thermophila* (Chug et al., 2015; Stuwe et al., 2015). The Nsp1 complex C-terminal structure is formed by a parallel, heterotrimeric coiled coil, followed by a second coiled coil region that folds back onto the first to form a compact six-helix bundle that is reminiscent of the number 4 (Chug et al., 2015; Stuwe et al., 2015; figure 9). The very N-terminal helix of Nic96 is sandwiched by the fold back architecture of the coiled coil regions of the Nsp1 complex. All 3 members of the Nsp1 complex are necessary for it to bind to Nic96 (Stuwe et al., 2015). Despite its distinct shape, the location of the Nsp1 complex coiled coil region has not been identified in the cryo-ET reconstructions of the NPC. This may be due to several reasons: its relatively small size (~80kDa) may make it hard to distinguish from any surrounding density, flexible attachment to the structural scaffold may cause it to be averaged out, or it could interact with yet unknown binding partners to form a different overall structure.

Nup82 complex

The Nup82 complex is localized to the cytoplasmic ring of the NPC and is comprised of Nup159, Nup82, and Nsp1. Nup159 has an N-terminal β-propeller followed by a long disordered FG-repeat rich region and an α-helical C-terminal domain (CTD) (Weirich et al., 2004). In *S. cerevisiae*, multiple dynein light chains (Dyn2) bind to the Nup159 FG-repeat region, suggesting that the Nup82 complex tethers the NPC to microtubules, likely for nucleus positioning within the cell (Stelter et al., 2007). Nup82 has a N-terminal β-propeller that binds to the APD of Nup145N, followed by a C-terminal coiled domain (Yoshida et al., 2011). The FG-repeats of Nup159, Nsp1, and Nup145N are thought to protrude radially from the NPC, and together with Nup82, serves as a platform for proteins involved in the last step of mRNA export (Fernandez-Martinez et al., 2016). Once mRNA is packaged into an export competent messenger ribonucleoprotein particle (mRNP) and reaches the cytoplasmic face of the NPC, it is remodeled by DEAD-box RNA helicase Dbp5, nucleoporin Gle1, and the N-terminal β-propeller of Nup159 (Folkmann et al., 2011; Montpetit et al., 2011). This process removes the associated transport receptor proteins from the mRNA, preventing it from traveling backwards through the transport channel (Lund and Guthrie, 2005). Similar to the Nsp1 complex, the Nup82 complex is held together primarily by coiled coil interactions between their C-termini (Belgareh et al., 1998; Gaik et al., 2015). According to size exclusion chromatography coupled to multi-angle laser-light scattering (SEC-MALLS) of an over expressed scNup82 complex with a truncated scNup159, the relative stoichiometry is 1:1:1, but after considering that Nup159 dimerizes through its CTD, the absolute stoichiometry is 2:2:2 (1998; Gaik et al., 2015). A recent computational study of a more complete assembly which incorporated EM data and cross-linking distance restraints, confirms a 2:2:2 stoichiometry (Fernandez-Martinez et al., 2017). A structural model for the interacting regions, generated in the same report, predicts mainly parallel-heterotrimeric coiled coils connected by flexible linkers for each unit. An asymmetric dimer structure is formed by two halves, each consisting of a single copy of Nup82, Nup159, and Nsp1, that adopt different configurations in a shape reminiscent of the letter D (Fernandez-Martinez et al.; figure 10). The Nup82 holo-complex docks to the rest of the structural scaffold by binding to Y complex member Nup85 (Fernandez-Martinez et al., 2017). It is exciting to see if experimental results can confirm the Nup82 holo-complex structure and its binding interface with Nup85, since the added mass to the short arm of the Y complex should further constrain possible docking solutions into the cryo-ET reconstructions of the NPC. Nevertheless, the authors attempt to dock a computationally generated Y/Nup82 complex model, based on a composite Y complex presented in this thesis, into the cryo-ET reconstruction. With the additional steric considerations, they conclude that only a single ring of Y complexes in a head-to-tail arrangement can be accommodated (Fernandez-Martinez et al., 2017).

Ndc1 complex

The Ndc1 complex has 4 members, each having predicted transmembrane α-helices. The Ndc1 complex is localized to the membrane proximal center of the transport axis and is assumed to be adjacent to the central inner ring. It is believed to act as a membrane adaptor for the structural scaffold of the NPC. Two of the four members, scNdc1 and scPom33 are conserved between yeast and metazoa. Two additional transmembrane Nups that interact with Ndc1 are scPom152 and scPom34 for yeast and the sequence unrelated Nups hsPom121 and hsGp210 for metazoa (Rothballer and Kutay, 2013). One possible consequence of the poor conservation of the Ndc1 complex between yeast and metazoa could be tied to their difference in open vs. closed mitosis. To date, only Pom152 has been structurally characterized experimentally. It consists of a N-terminal structural scaffold interacting head unit that is held proximal to the membrane by an adjacent transmembrane helix, followed by a chain of 9 Ig-like globular domains that form a semi-flexible tail sticking into the inner nuclear space (Upla et al., 2017). The authors go on to predict that 16 copies of the Pom121 Ig-like domain wrap around the luminal side of the nuclear membrane with 8 pairs of anti-parallel dimers. Ndc1 is predicted to have 6 N-terminal transmembrane helices and a soluble, C-terminal globular domain. Both ends of Ndc1 point into the transport axis of the NPC, where they can interact with the structural scaffold of the NPC (Rothballer and Kutay, 2013). Future work solving high resolution structures of the Ndc1 holo-complex and in complex with its known scaffold binding partners, Nup53 and

Nup170 of the Nic96 complex, can reveal the details of how the NPC is tethered to the membrane (Marelli et al., 2001; Onischenko et al., 2009).

Nuclear basket

The nuclear basket is a basketball hoop shaped structure that protrudes into the nucleus from the NPC structural scaffold seen in scanning electron micrographs of intact NPCs. Functionally, the basket is thought to interact with a number of nuclear proteins including the desumoylating enzyme Ulp1 in yeast, mRNPs, the proteasome, and telomere silencing factors (Li and Hochstrasser, 2000; Niepel et al., 2013). There are likely multiple proteins that are part of the basket structure, but it is assumed that hsTpr (scMlp1/scMlp2) is the primary contributor (Brohawn et al., 2009).

Tpr is a large, ~2000 residue protein that is predicted to form mainly coiled coils. Furthermore, based on structure prediction and limited proteolysis experiments, it is known that Tpr and Mlp1/Mlp2 form a series of flexibly tethered coiled coil domains, similar to beads-on-a-string (Niepel et al., 2013; Krull et al., 2004). To date, there have not been any reports on the atomic structure of Tpr or Mlp1/Mlp2. Due to its proximity to chromatin and interactions with important cellular machinery, the nuclear basket is a very interesting structure to study, although it is also perhaps the least understood NPC sub-complex.

References

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

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) Determining the architectures of macromolecular assemblies. Nature 450, 683–694.

Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D.P., Arumugam, M., Bork, P., and Hurt, E. (2011). Insight into Structure and Assembly of the Nuclear Pore Complex by Utilizing the Genome of a Eukaryotic Thermophile. Cell 146, 277.

Andersen, K.R., Onischenko, E., Tang, J.H., and Kumar, P. (2013). Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. Elife 2, e00745.

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, 7944-55-7955.

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. Molecular Biology of the Cell 9, 3475-3492.

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. The Journal Of Cell Biology 167, 591.

Bilokapic, S., and Schwartz, T.U. (2012). Molecular basis for Nup37 and ELY5/ELYS recruitment to the nuclear pore complex. Proceedings Of The National Academy Of Sciences 109, 15241.

Bilokapic, S., and Schwartz, T.U. (2013). Structural and Functional Studies of the 252 kDa Nucleoporin ELYS Reveal Distinct Roles for Its Three Tethered Domains. Structure 21, 572.

Boehmer, T., Jeudy, S., Berke, I.C., and Schwartz, T.U. (2008). Structural and functional studies of Nup107/Nup133 interaction and its implications for the architecture of the nuclear pore complex. Mol. Cell 30, 721-31.

Bonner W.M. (1978). Protein migration and accumulation in nuclei. In Busch, H. (ed.), *The Cell Nucleus.* Vol. 6, Part C. Academic Press, New York, NY, 97-148

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, 1369-73-1373.

Brohawn, S.G., and Schwartz, T.U. (2009). Molecular architecture of the Nup84-Nup145C-Sec13 edge element in the nuclear pore complex lattice. Nature Structural and Molecular Biology *16*, 1173-7.

Brohawn, S.G., Partridge, J.R., and Schwartz, T.U. (2009). The Nuclear Pore Complex Has Entered the Atomic Age. Structure 17, 1156.

Bui, K.H., Appen, von, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., et al. (2013). Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. Cell 155, 1233.

Capelson, M., Hetzer, M.W. (2009) The role of nuclear pores in gene regulation, development and disease. EMBO Rep. 10, 697-705

Chug, H., Trakhanov, S., Hülsmann, B., Pleiner, T., and Gorlich D. (2015). Crystal structure of the metazoan Nup62• Nup58• Nup54 nucleoporin complex. Science express.

Cook, A., Bono, F., Jinek, M., and Conti, E. (2007). Structural Biology of Nucleocytoplasmic Transport. Annual Review Of Biochemistry 76, 647.

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

Debler, E.W., Ma, Y., Seo, H., Hsia, K., Noriega, T.R., Blobel, G., and Hoelz, A. (2008). A fence-like coat for the nuclear pore membrane. Mol. Cell 32, 815-26.

Eibauer, M., Pellanda, M., Turgay, Y., Dubrovsky, A., Wild, A., and Medalia, O. (2015). Structure and gating of the nuclear pore complex. Nature Communications 6, 7532.

Elad, N., Maimon, T., Frenkiel-Krispin, D., and Medalia, O. (2009). Structural analysis of the nuclear pore complex by integrated approaches. Curr. Opin. Struct. Biol. 19, 226-32.

Fath, S., Mancias, J.D., Bi, X., and Goldberg, J. (2007). Structure and organization of coat proteins in the COPII cage. Cell 129, 1325-36.

Faria, A.M., Levay, A., Wang, Y., Kamphorst, A.O., Rosa, M.L., Nussenzveig, D.R., Balkan, W., Chook, Y.M., Levy, D.E., Fontoura, B.M. (2006). The nucleoporin Nup96 is required for proper expression of interferon-regulated proteins and functions. Immunity 24, 295–304

Fath, S., Mancias, J.D., Bi, X., and Goldberg, J. (2007). Structure and organization of coat proteins in the COPII cage. Cell 129, 1325-36.

Fernandez-Martinez, Z., Kim, S.J., Shi. Y., Upla, P., Pellarin, R., Gagnon, M., Chemmama, I.E., Wang, J., Nudelman, I., Zhang, W., Williams, R., Rice, W.J., Stokes, D.L., Zenklusen, D., Chait, B.T., Sali, A., Rout, M.P. (2016) Structure and Function of the Nuclear Pore Complex Cytoplasmic mRNA Export Platform. Cell 167, 1159-1160.

Wang, J., Nudelman, I., Zhang, W., Williams, R., Rice, W.,
Stokes, D.L., Zenklusen, D., Chait, B.T., Sali, A., and Rout M.P. (2017) Structure and Function of the Nuclear Pore Complex Cytoplasmic mRNA Export Platform. Cell 167, 1215-1228.

Finlay, D.R., Meier, E., Bradley, P., Horecka, J., and Forbes, D.J. (1991). A complex of nuclear pore proteins required for pore function. The Journal Of Cell Biology 114, 169-83.

Flemming, D., Devos, D.P., Schwarz, J., Amlacher, S., Lutzmann, M., and Hurt, E. (2012). Analysis of the yeast nucleoporin Nup188 reveals a conserved S-like structure with similarity to karyopherins. Journal Of Structural Biology 177, 99.

Folkmann, A.W., Noble, K.N., Cole, C.N., and Wente, S.R. (2011). Dbp5, Gle1-IP6 and Nup159: a working model for mRNP export. Nucleus 2, 540-8.

Frey, S., Richter, R.P., and Görlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science 314, 815-7.

Gaik, M., Flemming, D., Appen, von, A., Kastritis, P., Mucke, N., Fischer, J., Stelter, P., Ori, A., Bui, K.H., Bassler, J., et al. (2015). Structural basis for assembly and function of the Nup82 complex in the nuclear pore scaffold. The Journal Of Cell Biology 208, 283.

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, 114-24.

Harel, A., Orjalo, A.V., Vincent, T., Lachish-Zalait, A., Vasu, S., Shah, S., Zimmerman, E., Elbaum, M., Forbes, D.J. (2003) Removal of a single pore subcomplex results in vertebrate nuclei devoid of nuclear pores. Mol. Cell, 11, 853-64

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, 347-58.

Hsia, K., Stavropoulos, P., Blobel, G., and Hoelz, A. (2007). Architecture of a coat for the nuclear pore membrane. Cell 131, 1313-26-1326.

Hülsmann, B.B., Labokha, A.A., and Görlich, D. (2012). The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. Cell 150, 738-51.

Jeudy, S., and Schwartz, T.U. (2007). Crystal structure of nucleoporin Nic96 reveals a novel, intricate helical domain architecture. Journal Of Biological Chemistry 282, 34904-12-34912.

Kabachinski, G., and Schwartz, T.U. (2015). The nuclear pore complex – structure and function at a glance. Journal Of Cell Science 128, 423.

Kampmann, M., and Blobel, G. (2009). Three-dimensional structure and flexibility of a membrane-coating module of the nuclear pore complex. Nature Structural and Molecular Biology 16, 782.

Kim, D.I., Birendra, K.C., Zhu, W., Motamedchaboki, K., Doye, V., and Roux, K.J. (2014a). Probing nuclear pore complex architecture with proximity-dependent biotinylation. Proceedings Of The National Academy Of Sciences 111, E2453-61-E2461.

Krull, S., Thyberg, J., and Björkroth, B. (2004). Nucleoporins as components of the nuclear pore complex core structure and Tpr as the architectural element of the nuclear basket. Molecular Biology of the Cell 15, 4261-4277.

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, 2367-77.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., Davenport, A.M., Jeon, Y.E., Hoelz, A. (2016) Architecture of the symmetric core of the nuclear pore. Science 353, 308.

Liu, X., Mitchell, J.M., Wozniak, R.W., Blobel, G., and Fan, J. (2012). Structural evolution of the membrane-coating module of the nuclear pore complex. Proceedings Of The National Academy Of Sciences 109, 16498-503-16503.

Loïodice, I., Alves, A., Rabut, G., Van Overbeek, M., Ellenberg, J., Sibarita, J., and Doye, V. (2004). The entire Nup107-160 complex, including three new members, is targeted as one entity to kinetochores in mitosis. Molecular Biology Of The Cell 15, 3333-44-3344.

Lund, M.K., Guthrie, C. (2005) The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. Molecular Cell 20, 645-51.

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. The Journal Of Cell Biology 159, 267-78-278.

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

Lutzmann, M., Kunze, R., Stangl, K., Stelter, P., Tóth, K.F., Böttcher, B., and Hurt, E. (2005). Reconstitution of Nup157 and Nup145N into the Nup84 complex. Journal of Biological Chemistry 280, 18442-51.

Marelli, M., Lusk, C.P., Chan, H., Aitchison, J.D., and Wozniak, R.W. (2001). A link between the synthesis of nucleoporins and the biogenesis of the nuclear envelope. The Journal Of Cell Biology 153, 709-24-724.

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

Mi, L., Goryaynov, A., Lindquist, A., Rexach, M., and Yang, W. (2015). Quantifying Nucleoporin Stoichiometry Inside Single Nuclear Pore Complexes In vivo. Scientific Reports 5, 9372.

Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature 472, 238-42.

Niepel, M., Molloy, K.R., Williams, R., Farr, J.C., Meinema, A.C., Vecchietti, N., Cristea, I.M., Chait, B.T., Rout, M.P., and Strambio-De-Castillia, C. (2013). The nuclear basket proteins Mlp1p and Mlp2p are part of a dynamic interactome including Esc1p and the proteasome. Molecular Biology Of The Cell 24, 3920.

Ori, A., Banterle, N., Iskar, M., Andres-Pons, A., Escher, C., Bui, H.K., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., et al. (2013). Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. Molecular Systems Biology 9, 648.

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. The Journal Of Cell Biology 185, 475-91-491.

Paine, P.L., Moore, L.C., and Horowitz, S.B. (1975). Nuclear envelope permeability. Nature 254, 109–114.

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. Proceedings Of The National Academy Of Sciences 103, 17801-6-17806.

Rasala, B.A., Ramos, C., Harel, A., and Forbes, D.J. (2008). Capture of AT-rich chromatin by ELYS recruits POM121 and NDC1 to initiate nuclear pore assembly. Molecular Biology Of The Cell 19, 3982-96-3996.

Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. The Journal Of Cell Biology 110, 883-94.

Ribbeck, K., and Görlich, D. (2001). Kinetic analysis of translocation through nuclear pore complexes. The EMBO Journal 20, 1320-30.

Rothballer, A., and Kutay, U. (2013). Poring over pores: nuclear pore complex insertion into the nuclear envelope. Trends Biochem. Sci. 38, 292-301.

Rout, M.P., and Blobel, G. (1993). Isolation of the yeast nuclear pore complex. The Journal Of Cell Biology 123, 771-83-783.

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. The Journal Of Cell Biology 148, 635-51-652.

Sampathkumar, P., Kim, S.J., Upla, P., Rice, W.J., Phillips, J., Timney, B.L., Pieper, U., Bonanno, J.B., Fernandez-Martinez, J., Hakhverdyan, Z., et al. (2013). Structure, Dynamics, Evolution, and Function of a Major Scaffold Component in the Nuclear Pore Complex. Structure 21, 560.

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

Seo, H., Blus, B.J., Jankovic, N.Z., and Blobel, G. (2013). Structure and nucleic acid binding activity of the nucleoporin Nup157. Proceedings Of The National Academy Of Sciences 110, 16450-5-16455.

Seo, H., Ma, Y., Debler, E.W., Wacker, D., Kutik, S., Blobel, G., and Hoelz, A. (2009). Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. Proceedings Of The National Academy Of Sciences 106, 14281-6-14286.

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

Stelter, P., Kunze, R., Flemming, D., Höpfner, D., Diepholz, M., Philippsen, P., Böttcher, B., and Hurt, E. (2007). Molecular basis for the functional interaction of dynein light chain with the nuclear-pore complex. Nature Cell Biology 9, 788-96.

Stuwe, T., Lin, D.H., Collins, L.N., Hurt, E., and Hoelz, A. (2014). Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. Proceedings Of The National Academy Of Sciences 111, 2530.

Stuwe, T., Bley, C.J. , Thierbach, K., Petrovic, S., Schilbach, S., Mayo, D.J., Perriches, T., Rundlet, E.J., Jeon, Y.J., Collins, L.N., Huber, F.M., Lin, D.H., Paduch, M., Koide, A., Lu, V., Fischer, J., Hurt, E., Koide, S., Kossiakoff, A., Hoelz, A. (2015). Architecture of the fungal nuclear pore inner ring complex. Science, 350, 56-64

Thierbach, K., von Appen, A., Thoms, M., Beck, M., Flemming, D., Hurt, E. (2013) Protein interfaces of the conserved Nup84 complex from Chaetomium thermophilum shown by crosslinking mass spectrometry and electron microscopy. Structure 21, 1672-82.

Upla, P., Kim, S.J., Sampathkumar, P., Dutta, K., Cahill, S.M., Chemmama, I.E., Williams, R., Bonanno, J.B., Rice, W.J., Stokes, D.L., Cowburn, D., Almo, S.C., Sali, A., Rout, M.P., Fernandez-Martinez, J. (2017) Molecular Architecture of the Major Membrane Ring Component of the Nuclear Pore Complex. Structure 25, 1-12.

Ulrich, A., Partridge, J.R., and Schwartz, T.U. (2014). The stoichiometry of the nucleoporin 62 subcomplex of the nuclear pore in solution. Molecular Biology Of The Cell 25, 1484.

von Appen, A., Kosinski, J., Sparks L., Ori, A., DiGuilio, A. L., Vollmer, B., Mackmull, M.T., Banterle, N., Parc, L., Kastritis, P., Buczak, K., Mosalaganti, S., Hagen, W., Andres-Pons, A., Lemke, E. A., Bork, P., Antonin, W., Glavy, J. S., Bui, K. H., and Beck, M. (2005) In *situ* structural analysis of the human nuclear pore complex. Nature 526, 140-143.

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, 749-60.

Whittle, J.R.R., and Schwartz, T.U. (2009). Architectural Nucleoporins Nup157/170 and Nup133 Are Structurally Related and Descend from a Second Ancestral Element. Journal Of Biological Chemistry 284, 28442.

Whittle, J., and Schwartz, T.U. (2010). Structure of the Sec13–Sec16 edge element, a template for assembly of the COPII vesicle coat. The Journal Of Cell Biology.

Wolfe, K.H., and Shields, D.C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387, 708-13.

Yoshida, K., Seo, H., Debler, E.W., Blobel, G., and Hoelz, A. (2011). Structural and functional analysis of an essential nucleoporin heterotrimer on the cytoplasmic face of the nuclear pore complex. Proceedings Of The National Academy Of Sciences 108, 16571-6.

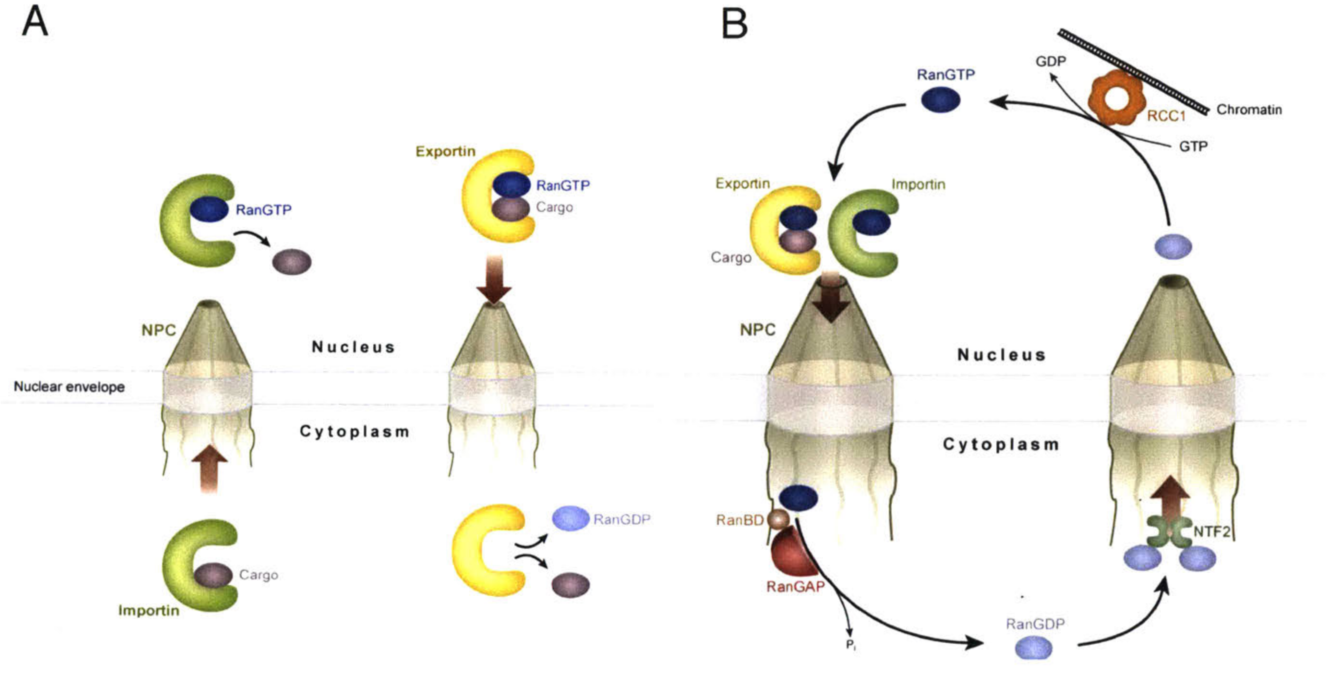


Figure 1. Schematic drawing of transport through the NPC and associated Ran cycle. (A) An importin binds a cargo in the cytoplasm and releases it upon binding RanGTP in the nucleus. An exportin binds both cargo and RanGTP in the nucleus and releases them upon conversion of RanGTP into RanGDP. **(B)** The high concentration of RanGDP in the cytosol is maintained by RanGAP, which is bound to the cytoplasmic filaments of the nuclear pore complex. With the help of other factors, it acts on the RanGTP that enters the cytoplasm (via binding to exportins and importins). The high concentration of RanGTP in the nucleus is maintained by regulator of chromosome condensation 1 (RCC1), which acts on the RanGDP that enters the nucleus [with its dedicated transport factor nuclear transport factor 2 (NTF2)].

Figure borrowed and modified from Cook et al., 2007.

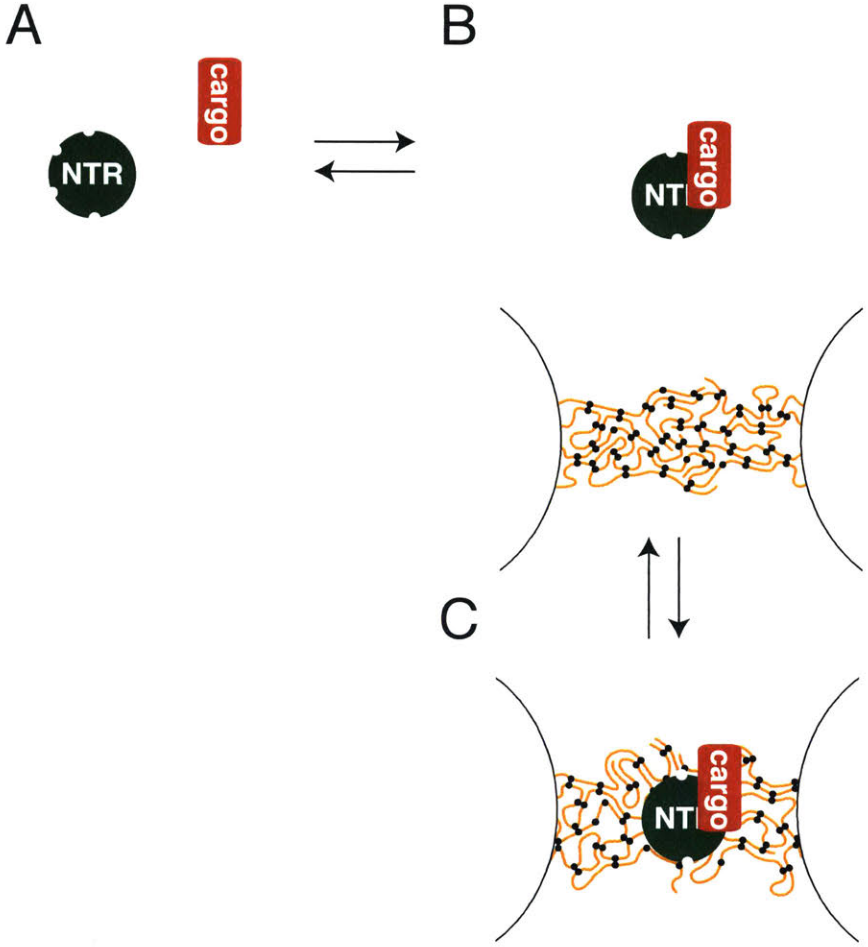


Figure 2. Transport of cargo through the NPC FG-repeat hydrogel.
(A) Cargo(NLS containing) protein larger than ~40kDa and transport receptor (importins) on the cytoplasmic side. **(B)** Receptor-cargo complex is formed. NPC transport barrier is formed by a FG-repeat protein hydrogel. **(C)** Receptor-cargo complex diffusing through barrier by disrupting the hydrogel.

Figure borrowed and modified from Hülsmann et al., 2012.

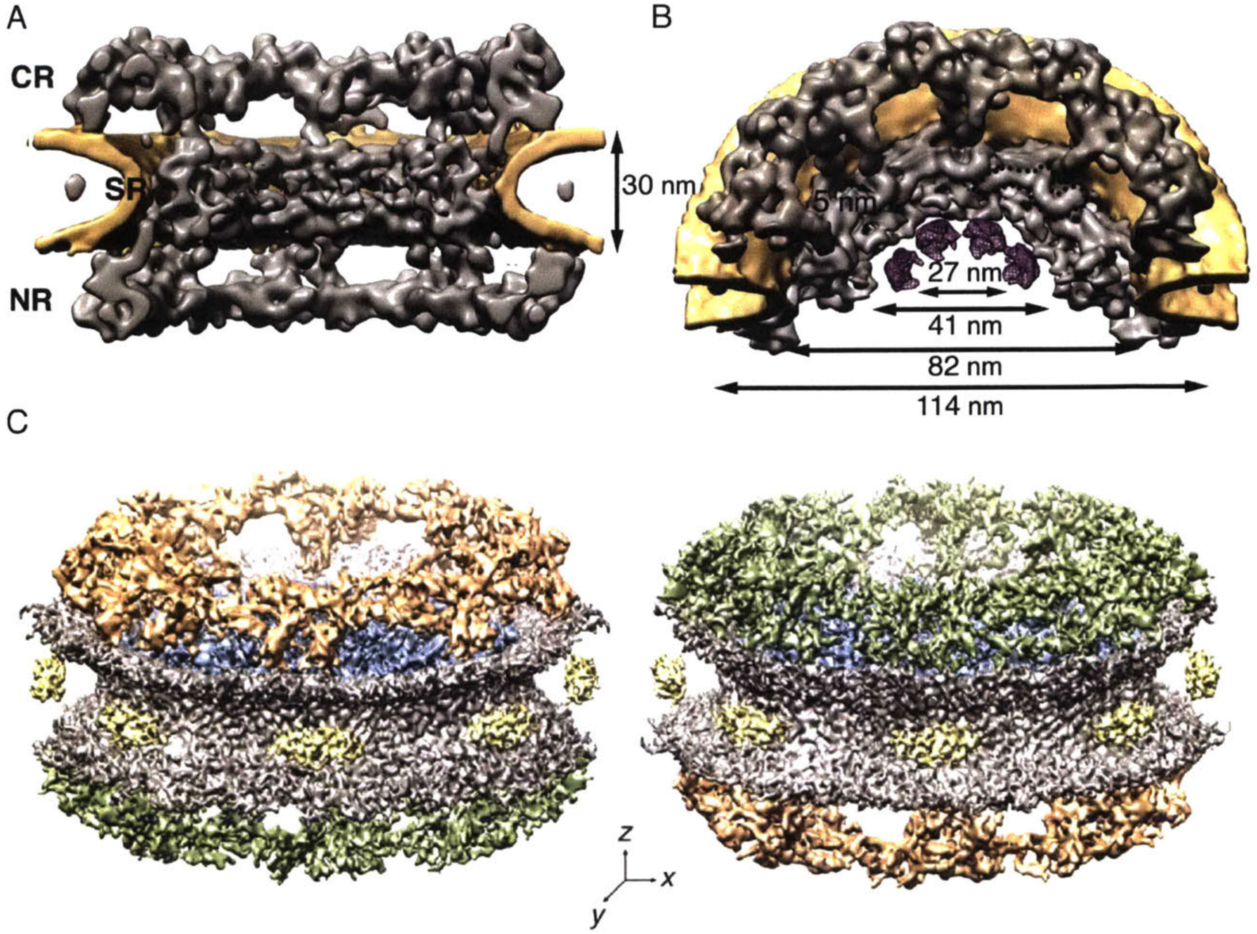


Figure 3. Cryo-ET reconstructions of the metazoan NPC. (**A**) Side-on cut-away view of the 32 Å cryo-ET reconstruction of the human NPC from Bui et al., 2013. CR, cytoplasmic ring. SR, spoke ring (In this thesis we will use IR, inner ring, instead of SR). NR, nucleoplasmic ring. Protein density (grey) and the membrane (brown). (**B**) Top-down view of the human NPC from Bui et al., 2013. Dimensions of the NPC are labeled. (**C**) 20 Å cryo-ET reconstruction of the Xenopus NPC from Eibauer et al., 2015. The two side-on views are rotated 180° w.r.t. each other. The CR (gold), IR (blue), NR (green), membrane (grey), and periplasmic density (yellow).

Figures borrowed from Bui et al., 2013 and Eibauer et al., 2015.

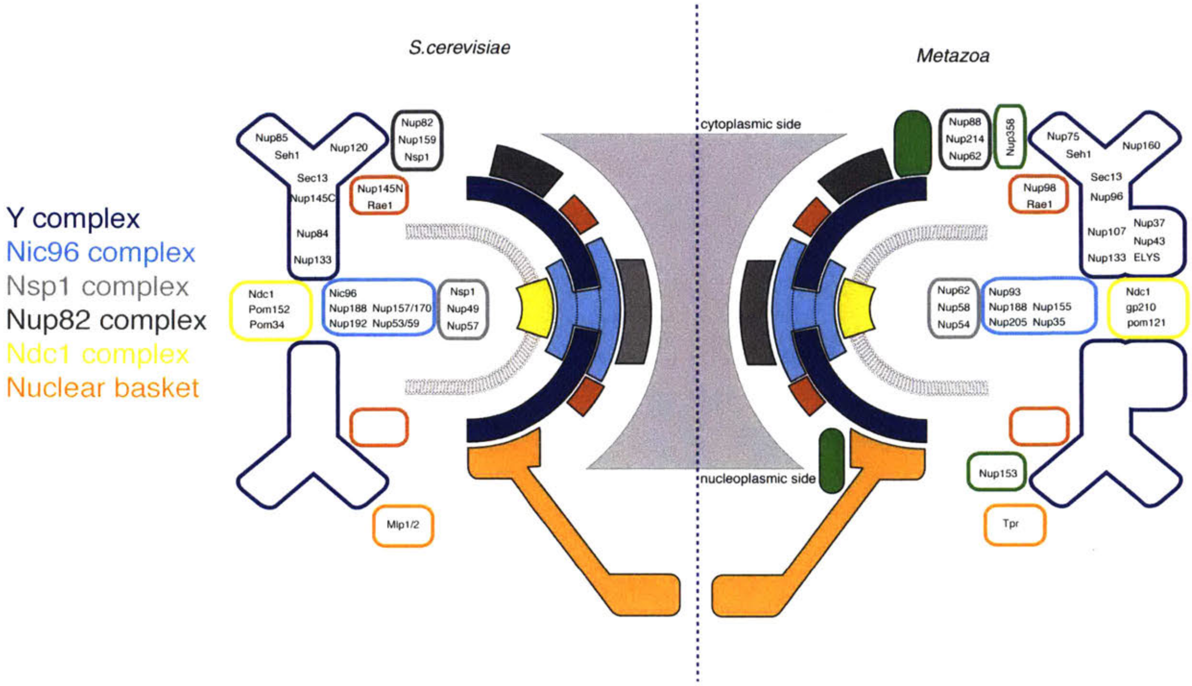


Figure 4. Schematic organization of NPC. NPC viewed perpendicular to its transport axis. The ~30 distinct nucleoporins are organized into biochemically distinct sub-complexes. Multiple copies of each sub-complex organize around an eight-fold rotational symmetry axis parallel to the transport axis. Each sub-complex is color coded and placed to reflect their approximate localization based on immuno-gold labeling EM studies.

Figure borrowed and modified from Brohawn et al., 2009.

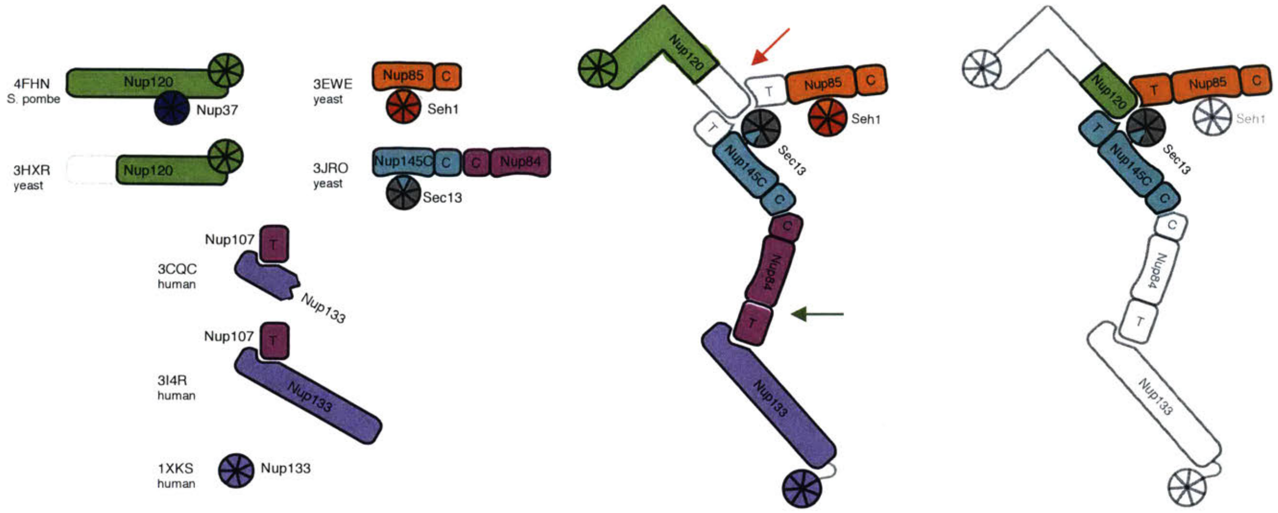


Figure 5. Schematic arrangement of the Y complex. (left) Previously solved Y-complex crystal structures. **(middle)** Schematic arrangement of the Y complex. Previously solved fragments have been colored in. Unsolved fragments are shown in transparent grey (red arrow, the "hub" and green arrow, 4 unsolved helices of Nup84). **(right)** Cartoon of the "hub" crystal construct (colored) presented in this thesis.

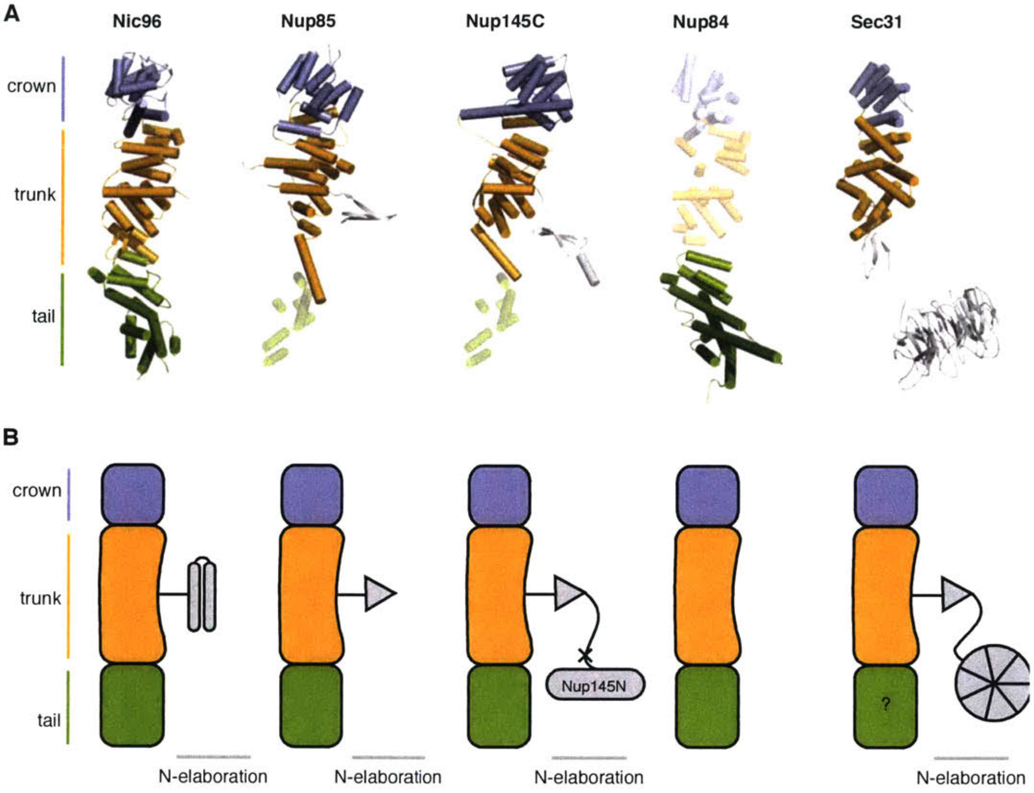


Figure 6. The ACE1 fold. (A) Crystal structures of ACE1 proteins. Unsolved modules have been computationally modeled and shown transparent. Blue, crown module; yellow, trunk module; green, tail module. Several Nups share this common fold with COPII vesicle coat edge element Sec31. **(B)** Cartoon representation of the crystal structures above illustrates the similarity of the ACE1 elements.

Figure borrowed and modified from Brohawn et al., 2008.

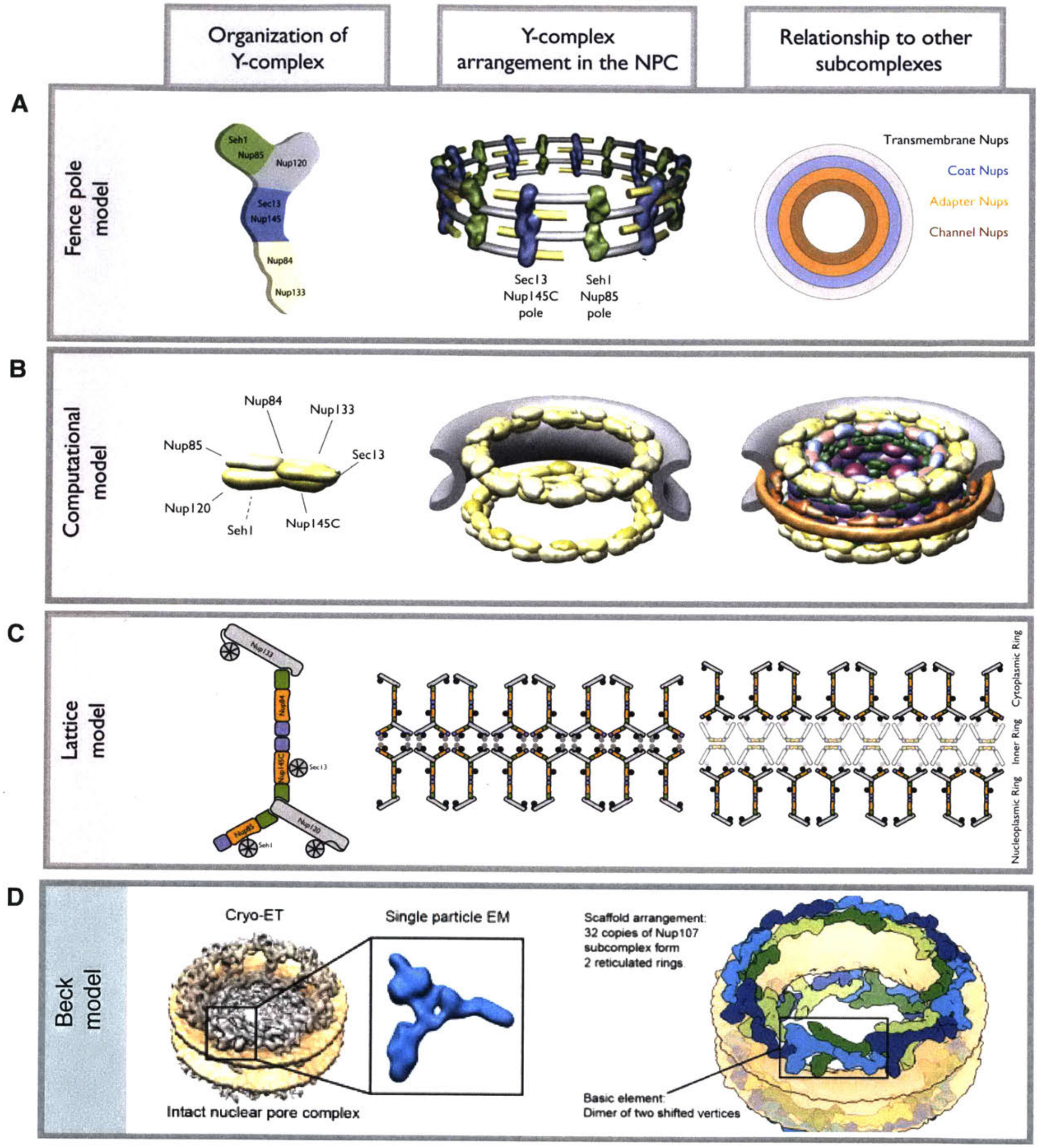


Figure 7. Higher order assembly models of the Y complex. (A) The fence pole model. Hsia et al., 2007) **(B)** Computationally generated model. (Alber et al., 2007) **(C)** Lattice model. (Brohawn et al., 2008) **(D)** Beck model. (Bui et al., 2013).

Figure borrowed and modified from Brohawn et al., 2010 and Bui et al., 2013.

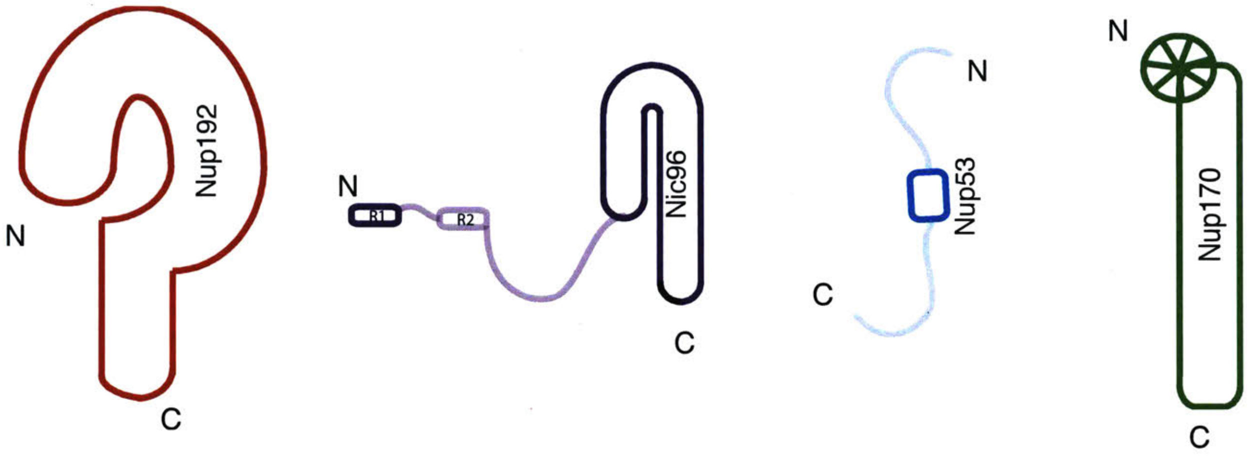


Figure 8. Cartoon representation of Nic96 complex nucleoporins. Regions for which crystal structures are available are colored. Yet unsolved regions are transparent. Nup192/Nup188 has a characteristic S-shape. ACE1 protein Nic96 binds to the C terminus of Nup192/Nup188 through its N-terminal R2 helix and to the Nsp1 complex through its R1 helix. Nup53 binds both Nup192/Nup188 and Nic96 through its N-terminal unstructured region and to Nup170 through its C-terminal unstructured region. Nup170 has a N-terminal β-propeller followed by an α-helical stack which shares structural holomoly with Y-complex member Nup133.

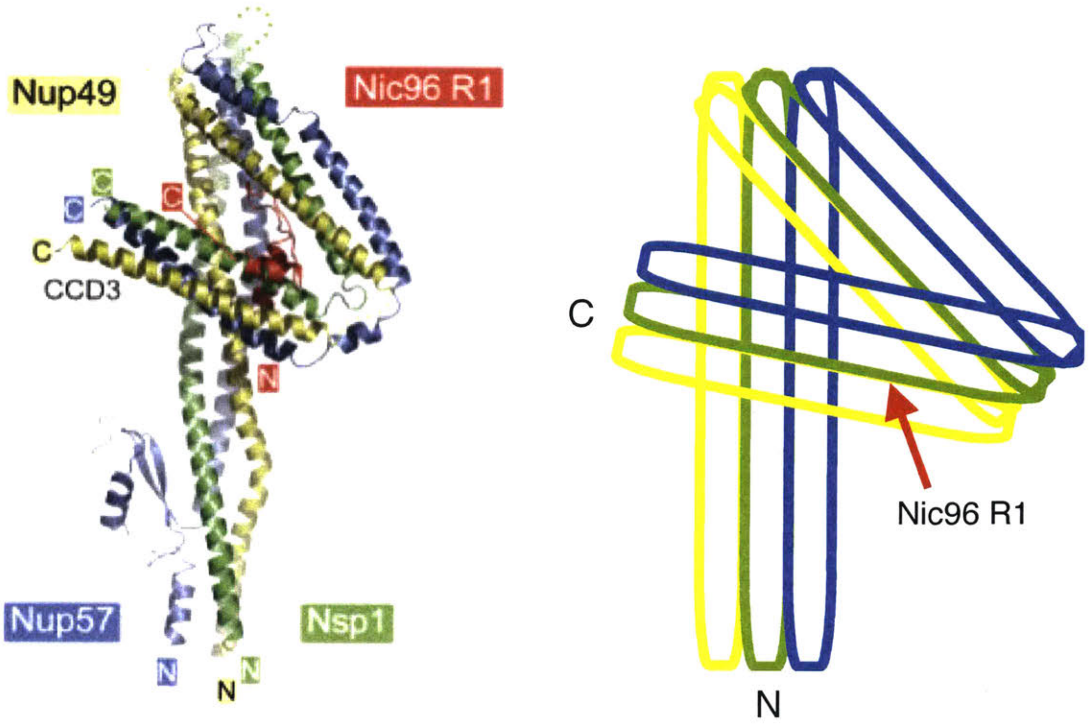


Figure 9. Crystal structure of the Nsp1 complex bound to the R1 helix of Nic96. The Nsp1 complex C-terminal structure is formed by a parallel, heterotrimeric coiled coil followed by a second coiled coil region that folds back onto the first to form a compact six helix bundle that is reminiscent of the number 4. The very N-terminal R1 helix of Nic96 is sandwiched by the fold back architecture of the coiled coil regions of the Nsp1 complex. Right, cartoon representation of the crystal structure.

Figure borrowed and modified from Stuwe et al., 2015

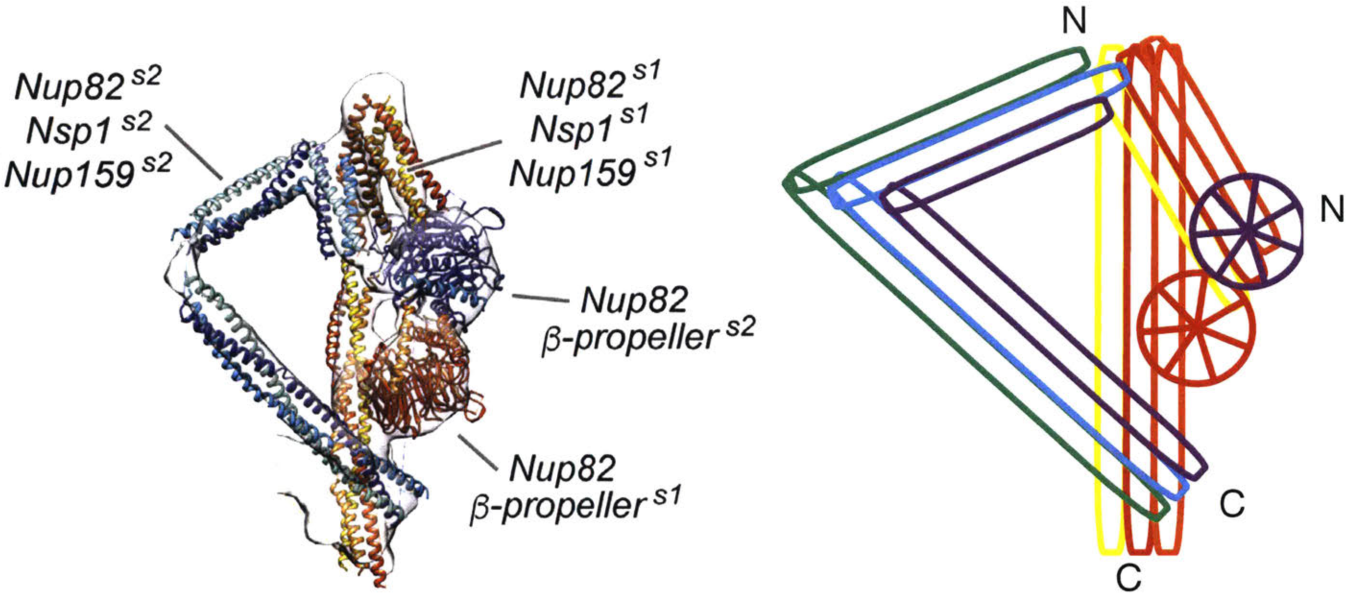


Figure 10. Computational structure of the Nup82 complex. A structural model for the interacting C-terminal region of the Nup82 complex predicts mainly parallel hetero-trimeric coiled-coils connected by flexible linkers for each unit. An asymmetric dimer structure is formed by two halves (s1 and s2), each consisting of a single copy of Nup82, Nup159, and Nsp1, that adopt different configurations in a shape reminiscent of the letter D. Right, cartoon representation of the crystal structure.

Figure borrowed and modified from Fernandez-Martinez et al., 2016.

Chapter 2: Atomic structure of the Y complex of the nuclear pore

The material presented in this chapter was adapted, with permission, from the following publication:

Kelley, K.*, Knockenhauer, K.E.*, Kabachinski, G., and Schwartz, T.U. (2015). Atomic structure of the Y complex of the nuclear pore. Nature Structural & Molecular Biology *22*, 425-431.

*Equal Contribution

Author contributions: T.U.S., K.K. and K.E.K. designed the study. K.K. and K.E.K. performed the experiments. K.K., K.E.K. and T.U.S. analyzed the data. G.K. performed and analyzed the fitness tests. K.K., K.E.K., G.K. and T.U.S. interpreted the structure and wrote the manuscript.

Introduction

In all eukaryotic cells transcription and translation are physically separated between the nucleus and cytoplasm. This allows for distinct gene expression control mechanisms, for example in cell differentiation and development, unavailable to prokaryotes. Nuclear pore complexes (NPCs), which perforate the nuclear envelope and act as the main transport gate, therefore have a fundamental role in cellular homeostasis (Hetzer and Wente, 2009; Strambio-de-Castillia et al., 2010). The NPC is a modular, donut-shaped assembly of ~30 different proteins (nucleoporins or nups), arranged in multiples of eight around a central axis that is aligned with the main transport channel (Brohawn et al., 2009). Nups can be classified as (i) architectural nups, which form the stable scaffold of the NPC, (ii) peripheral nups with various degrees of mobility, and (iii) nups with the characteristic phenylalanine-glycine (FG)-repeat elements in disordered extensions that form the permeability barrier.

To gain mechanistic insight into NPC function, considerable effort has been undertaken to determine the NPC structure at high resolution. Due to its enormous size of ~40–120 MDa (Reichelt et al., 1990; Rout et al., 2000; Yang et al., 1998), this will ultimately only be possible with a combination of different visualization techniques, notably X-ray crystallography and cryo-electron microscopy.

One of the main architectural elements of the NPC is the Y-shaped complex, which is essential for NPC formation (Harel et al., 2003; Walther et al., 2003). In *S. cerevisiae*, it is a 7-membered 575 kDa complex composed of Nup84, Nup85, Nup120, Nup133, Nup145C, Sec13 and Seh1 (Lutzmann et al., 2002) . Homologs for these 7 members are also found in humans, suggesting conservation within all opisthokonts (Belgareh et al., 2001; Neumann et al., 2010; Vasu et al., 2001). It is widely believed that the Y-complex, together with the heteromeric Nic96 complex, forms the principal NPC scaffold (Alber et al., 2007; Bui et al., 2013; Vollmer and Antonin, 2014). Thus, the atomic structures of the assembled Y- and Nic96 complexes are important milestones toward understanding the NPC at the highest resolution.

All 7 members of the Y-complex have been studied crystallographically, though only individually or as heterodimeric or trimeric complexes, and typically truncated (Berke et al., 2004; Bilokapic and Schwartz, 2012; Boehmer et al., 2008; Brohawn and Schwartz, 2009; Brohawn et al., 2008; Debler et al., 2008; Hsia et al., 2007; Leksa et al., 2009; Nagy et al., 2009; Sampathkumar et al., 2011; Seo et al., 2009; Whittle and Schwartz, 2009). The branch point or 'hub' of the Y-complex, where the two short arms and the long stalk meet, is structurally the least understood even though it is arguably the most important element (Fig. 1a).

Here we now report the structure of the Y-complex hub, which enables us to combine all the additional fragmentary structures into a highly accurate assembled structure of the Y-complex. We can show that the Y-complex structure is widely conserved among all eukaryotes. Species-specific additions to the Y-complex decorate, but do not principally alter the overall structure.

Results

Structure of the Y-complex hub

We generated a series of structure-based expression constructs containing the elements of Nup120, Nup145C, and Nup85 that were known to directly interact with each other at the Y-complex hub (Bilokapic and Schwartz, 2012; Brohawn et al., 2008; Fernandez-Martinez et al., 2012; Leksa et al., 2009). In addition, these constructs were designed to overlap at least partially with the already structurally characterized Y-complex fragments. We succeeded in obtaining crystals of a heterotetrameric construct containing $\text{Nup85}_{257\text{-}1181}$, $\text{Nup120}_{952\text{-}1241}$, $\text{Nup145C}_{233\text{-}791}$, and Sec13 from the thermophilic fungus *Myceliophthora thermophila (mt)*, which diffracted to 4.1 Å resolution (Table 1). The structure was solved with a combination of molecular replacement and single anomalous dispersion (SAD) using selenomethionine-derivatized protein (see Methods for details). The crystals have a high solvent content (68%) and the structure has a comparatively high degree of positional disorder (Wilson B factor 144 Å^2). Despite the high Wilson B factor, we were able to properly assign the sequence to all four proteins within the assembled complex based on model building guided by selenomethionine positions, homology models, and phylogenetic considerations (Table 2).

The overall structure of the heterotetramer is roughly V-shaped, composed of three helical units, Nup85, Nup120, and Nup145C, and a laterally attached β-propeller, Sec13 (Fig. 1b, Fig. 2). Nup85 and Nup145C form the long sides of the 'V', whereas Nup120 is sandwiched between the two sides and acts as the main connector. Sec13 is bound to Nup145C as previously described in the Nup145C-Sec13-Nup84 structure from *Saccharomyces cerevisiae (sc)* (Brohawn and Schwartz, 2009; Nagy et al., 2009), namely by the insertion of a seventh blade into its open, 6-bladed β-propeller. Viewed from the side, it is noticeable that the heterotetramer is substantially bent rather than flat (Fig. 1c). The overall dimensions of the complex are ~125 Å x ~140 Å x ~50 Å.

Nup85 and Nup145C belong to the ancestral coatomer element 1 (ACE1)-class of proteins. These proteins are characterized by a ~65kDa, tripartite helical segment composed of a crown, a trunk and a tail element, which adopts a characteristic fold-back structure involving ~30 α-helices (Brohawn et al., 2008; 2009). ACE1 proteins are exclusively found in the NPC scaffold and the COPII vesicle coat (Whittle and Schwartz, 2010). The scNup85 and scNup145C fragments solved previously both lack the tail elements, which are present in this structure. These tail elements are in direct contact with Nup120. Even though the identity between the Mt and Sc sequences is low (14% for Nup85, 20% for Nup145C) (Table 3), the structures superpose well (Fig. 3). Therefore, we assigned secondary structure elements in the Mt proteins in accord with the published Sc fragments (Brohawn and Schwartz, 2009; Brohawn et al., 2008). Nup145C has 27 helices in total. Helices α1–α3 and α12–α20 form the trunk, α4–α11 the crown, and α21 to α27 the tail. The entire helical stack of Nup145C has a crescent shape. To stabilize the Sec13 interaction, we fused Sec13 N-terminally to 145C, similar to what was previously done for scSec13-Nup145C (Brohawn and Schwartz, 2009). For this reason, our Nup145C construct is lacking 232 N-terminal, predicted-to-be disordered residues.

Nup85 has 33 helices in total. Helices α1–α3 and α12–α20 form the trunk, α4–α11 the crown, and α21 to α30 the tail. Compared to mtNup145C, the mtNup85 tail domain is longer, and contains 4 additional helices. In *S. cerevisiae*, Nup85 binds Seh1 very similarly to Nup145C binding Sec13, i.e. via an insertion blade that closes the open, 6-bladed β-propeller. Although Sordariomycota, including *M. thermophila* and *C. thermophilum*, also contain a recognizable Seh1 homolog, it does not bind to Nup85 (Thierbach et al., 2013). The structure now reveals that mtNup85 lacks the essential Seh1-binding site, i.e. the insertion blade. Instead, it contains an additional N-terminal helix α0 (Fig. 1, Fig. 4), incompatible with Seh1 binding. In light of this observation, we suggest that Seh1 is not a member of the conserved core of the Y-complex. This core is a heterohexamer rather than a heteroheptamer. The presence of an Seh1 homolog in Sordariomycota is likely due to its function in the GATOR(SEA) complex (Algret et al., 2014; Bar-Peled et al., 2013).

The C-terminal fragment of Nup120 contains 10 helices (α23–α29,α29a,α29b,α30), 8 of which form a regular stack. This stack superimposes very well (RMSD 2.2 Å over 135 Cα positions) with the C terminus of Nup120 from *S. pombe* (Bilokapic and Schwartz, 2012). The C-terminal mtNup120 element is wedged between the tail domains of Nup85 and Nup145C, and forms the core of the hub. The interface between Nup120 and Nup85 buries a ~940 $Å^2$ surface, whereas the Nup120-Nup145C interface measures ~1700 $Å^2$. The Nup120-Nup85 interface includes helices α27 and α29 from Nup120 and helices α28 and α30 from Nup85. This arrangement generates a tightly packed four-helix bundle between the two proteins, likely explaining the high affinity of both proteins for each other. In comparison, the Nup120-Nup145C interface is more complicated, with two major elements contributing contacts. One part of the binding interface is generated by the terminal helix α27 of Nup145C that runs along the narrow ridge of the Nup120 C-terminal helical stack. Helices α29a and α29b of Nup120 are outside of the C-terminal stack and clamp down Nup145C, as they are positioned on the outward-facing side of the molecule, thereby forming a second interaction site. Helices α29a+b are not conserved, but appear to be specific for Sordariomycota.

Sec13 is placed right in the neck of the V-shaped molecule, with contacts to all three helical proteins. The contact area with mtNup145C is largest and very similar to the homologous structure of scNup145C (Brohawn and Schwartz, 2009; Nagy et al., 2009). In addition to the close contacts with the insertion blade, the top surface of the Sec13 β-propeller binds the trunk of the Nup145C-ACE1 element. The Sec13-Nup120 contacts are established through interactions between blades 1 and 2 of Sec13, and the very C-terminal helix α30 of Nup120. Finally, contacts between Sec13 and Nup85 are few, but nonetheless likely important. Taken together, it appears as if the Sec13 position helps to rigidify the interaction between the three helical elements.

Fitness analysis

To evaluate the importance of the hub interface *in vivo,* we used a fitness test in *S. cerevisiae.* C-terminal truncations of the last helix of Nup85, Nup120, and Nup145C, were designed to selectively disrupt the mapped interfaces between the three helical proteins. The Nup85Δα30 mutant had the most severe phenotype and showed drastically reduced growth (Fig. 5). Nup145CΔα27 and Nup120Δα30 have progressively milder phenotypes. Nup85Δα30 almost phenocopies the lethal Nup85 knockout (Siniossoglou et al., 1996), suggesting that the Nup85-Nup120 interaction is critical for NPC assembly. For Nup120 and Nup145C, it is likely that the mapped interfaces are not the exclusive elements that integrate these proteins into the NPC, but that additional contacts exist. The N-terminal extension of Nup145C, past the Sec13 insertion blade and not part of our structure, is likely to play a role in this. However, contacts to other scaffold nucleoporins need to be considered as well. Additionally, although we did not formally quantify the protein levels or test the fold retention of the individual truncated proteins, based on previous *in vivo* (Fernandez-Martinez et al., 2012) and *in vitro* (Bilokapic and Schwartz, 2012; Brohawn et al., 2008; Leksa et al., 2009; Stuwe et al., 2015) experiments, we can assume that our specific truncations are folded correctly and expressed at wild type levels.

Composite high-resolution structure of the Y-complex

With the heterotetrameric hub assembly in hand we set out to build a complete high-resolution, composite structure of the entire Y-complex. The structures of full-length Nup120 (Bilokapic and Schwartz, 2012) and Nup84-Nup145C-Sec13 (Brohawn and Schwartz, 2009) contain overlapping elements with the hub structure, 19 kDa and 85 kDa in size, respectively, and were superposed with high confidence (Fig. 6). This generated a heteropentameric complex, in which Nup84 is the preliminary terminal fragment of the long stalk of the Y. Nup84 is an ACE1 protein, but the published structure lacks the tail domain (residues 443–724) that interacts with Nup133. However, the structure of most of the tail domain of Nup107, the human Nup84 homolog, in complex with Nup133 is known (Boehmer et al., 2008). Therefore, we were able to model full-length Nup84 based on the experimentally known N- and C-terminal fragments, and the homology-modeled structure of the intervening 84-residue segment using other ACE1 domains (see Methods for details). After Nup84 was positioned, Nup133 could be docked based on the Nup107-Nup133 complex structure (Whittle and

Schwartz, 2009). Finally, the last element on the long stalk of the Y-complex is the N-terminal Nup133 β-propeller, which is loosely tethered to the C-terminal, α-helical stack element (Berke et al., 2004; Kim et al., 2014). The resulting composite structure constitutes the conserved, heterohexameric core of the Y-complex (Fig. 7a).

Because the composite model is built from structural elements of four different organisms, we examined to what extent this might affect the overall structure. Therefore, we also built models for the heteroheptameric Y-complex in *S. cerevisiae*, including Seh1, as well as the heterodecameric Y-complex in *H. sapiens*, including Seh1 and Nup37 (Fig. 7b, c). The Seh1 position can be deduced from the Seh1-Nup85 complex structure (Brohawn et al., 2008; Debler et al., 2008), whereas the Nup37 position is known from the Nup37-Nup120 complex structure (Bilokapic and Schwartz, 2012; Liu et al., 2012). The positions of Nup43 and ELYS within the human complex are not well understood and were therefore excluded. By comparison, we observed that the conserved, heterohexameric core changes only in local areas between the three models and that the overall shape and dimensions of the Y-complex appear conserved in all opisthokonts. Organism-specific proteins decorate the Y-core, but do not significantly influence its overall structure otherwise.

Our composite structure reveals that the Y-complex when viewed from the front measures about 20 nm wide and 40 nm high (Fig. 7). This is in good agreement with published EM structures and computational models (Bui et al., 2013; Fernandez-Martinez et al., 2012; Kampmann and Blobel, 2009). However, the principal angles between the three extensions from the hub deviate significantly between our X-ray based structure and the EM structures, which is evident by superimposing the different sets of data (Fig. 8). Strikingly, when viewed from the side or the top, the composite crystal structure reveals a distinct, three-dimensional shape (Fig. 7e, Fig. 8b, f), whereas the previously reported structures were essentially flat (Fig. 8d, h). At the hub, the three extensions, namely the two arms and the stalk, exhibit strong curvature and form a dome-shaped structure. As a result, we measure a thickness of ~8 nm for the Y-complex (Fig. 7e), compared to ~4 nm reported in previous EM analyses. Theoretically, this difference could be the result of flexibility within the Y-complex, which is well documented. In our composite structure we can now specify the main hinge regions and flexible areas (Fig. 9). These flexible regions are detected by general considerations regarding protein structure, domain boundaries established by limited proteolysis, and flexibility observed in previous crystallographic and EM studies (Berke et al., 2004; Bilokapic and Schwartz, 2012; Boehmer et al., 2008; Bui et al., 2013; Kampmann and Blobel, 2009; Whittle and Schwartz, 2009). The hub itself is rather rigid because four proteins engage in a tightly coordinated interface. To flatten the Y-complex, the helical Nup120 C-terminal domain or the Nup85 ACE1 element would have to bend by nearly 90° with respect to each other. Helical stacks have the propensity to bend, as best exemplified by various nuclear transport receptors (Cook et al., 2007; Süel et al., 2006), but the direction of bending is determined by the helical orientation. For the flattening of the Y, the predicted-to-be bendable elements are, however, oriented in an unfavorable way. This means that the necessary distortions in Nup120 or Nup85 would be energetically costly because of the disruption to the hydrophobic core that they would generate, which we consider to be rather unlikely. Therefore, the simplest explanation for the discrepancy between our composite crystal structure and previous random-conical tilt (RCT) negative stain EM structures is that the latter were artificially flattened in the direction normal to the EM grid, a well-known phenomenon.

Implications for NPC assembly models

Next, we tested whether our composite human Y-complex could be positioned into the recently published 3.2 nm cryo-ET density map (Bui et al., 2013) of the human NPC, which predicted a staggered two-ring, head-to-tail orientation of Y-complexes, symmetrically positioned on the nucleoplasmic and the cytoplasmic face of the NPC. We were able to recapitulate the published results of docking the flat, Y-complex structure determined by RCT negative stain EM (EMD code 2443) (Bui et al., 2013) into the cryo-ET map (Bui et al., 2013). We then tried the same procedure using our highly curved Y-complex structure. We searched with the human Y-complex model omitting Nup133 and the tail domain (residues 658–925) of Nup107, presumably the most flexible Y-complex elements. Nup133 has an N-terminal β-propeller flexibly connected to a C-terminal helical stack domain (Berke et al., 2004). The helical domain of Nup133 is tripartite, with hinges connecting the three helical segments (Whittle and Schwartz, 2009). A flexible hinge in the Nup107 ACE1 element between the trunk-crown, and tail segments can be postulated based on independent structure-mapping studies performed on the human and yeast homologs (Boehmer et al., 2008; Brohawn and Schwartz, 2009; Nagy et al., 2009). Using this stubbed Y-complex, we found three top numerical solutions (see Methods for details). Two solutions roughly coincided with the outer Y-complex ring postulated by Bui et al. (Bui et al., 2013) (Fig. 10), and the third solution coincided with the inner ring of that study (Fig. 11). While our work was in review, Stuwe et al. reported the docking of a similarly stubbed Y-complex from *S. cerevisiae* based on a 7.4 Å crystal structure, which shows a curved topology consistent with our structure, and arrived at a similar solution (Stuwe et al., 2015). However, when we added the Nup107 tail and Nup133 structures back to the docked Y-complex model of our third solution, in any topologically reasonable way, we observed extensive steric clashes with the neighboring Y-complex that seem highly implausible (Fig. 11). Thus, we did not consider this solution further. Regarding the two top solutions, they are rotated around the hub by approximately 20° relative to each other. In each solution the long stalk could be fitted reasonably well to two different regions in the EM density. Both solutions result in a seemingly closed ring when Nup133 is added, albeit the head-to-tail contact would be different in each case. To fit each solution, the long stalk needs to adopt different conformations, largely by adjusting Nup133, which seems realistic based on the expected flexibility around distinct hinge points (Fig. 10c). Obviously, both solutions cannot coexist due to excessive steric clashes. Therefore, the easiest way to explain our docking results is to suggest that the Y-complex ring is a single rather than a double ring, but that it can adopt at least two conformations. We argue that because of subtomogram-averaging, we might see an overlay of the two, equally and most populated, states of the Y-complex ring in the cryo-ET density.

Discussion

As reasonable as our docking attempts may appear, we would like to caution about the interpretation of these results. First, the available cryo-ET map (EMD-2444) (Bui et al., 2013) is calculated based on assumptions that we still do not know to be necessarily correct. For example, a strict eightfold rotational symmetry is applied, which may be appropriate at nm resolution, but possibly not at atomic resolution. If this symmetry is not true on the atomic level, the calculated map could be intrinsically flawed. Due to the similarity of various scaffold nups on a nanometer scale this is particularly troublesome. Second, docking at ~3 nm resolution is at best tentative and only reasonable to attempt because of the distinct and large size of the Y-complex. It is possible that an entire portion of the Y-complex is so flexible in the NPC assembly that it could simply be averaged out in the cryo-ET study. This could in principal be true for the long stalk as well as most of the Nup120 arm, which can bend perpendicular to the long axis of its C-terminal helical stack domain (Bilokapic and Schwartz, 2012) . Many additional docking solutions would need to be considered, if such possibilities were entertained.

One way of independently confirming a specific assembly model initially derived from docking is to map the assumed contacts between neighboring complexes. Bui et al. attempted this by crosslinking experiments (Bui et al., 2013). Surprisingly, only two such crosslinks (Nup107-Nup133 and Nup43-Nup96) were found. However, these interactions could not be confirmed by additional experiments and, therefore, are insufficient to distinguish between the models. A two-ring model in particular would generate many inter-Y contact sites, thus the paucity of detected crosslinks is unexpected. Taking additional studies into account does not resolve the discrepancy between the reticulated two ring model and a flexible one ring model. For example, the radial distribution of fluorescently labeled scaffold nucleoporins was determined by super-resolution microscopy (Szymborska et al., 2013). However, the raw localization precision of 20-30 nm attained in that study is insufficient to distinguish between the different ring models. Stoichiometric considerations, which have major model implications, are also not definitive. For example, a recent study concluded there are 32 copies of Y-complexes per NPC in HeLa cells (Ori et al., 2013), a number supporting a two ring model, whereas earlier studies argued for 16 copies (Cronshaw et al., 2002; Rout et al., 2000), which would favor a one ring model. Recent advancements in super-resolution microscopy should allow for the direct counting individual nucleoporins within an NPC, putting this controversy to rest soon (Ries et al., 2012). Finally, mass estimates for the NPC range from 40 MDa to 120 MDa, depending on the species and technique used (Reichelt et al., 1990; Rout et al., 2000; Yang et al., 1998). Again, it is premature to use this information to definitively validate a specific model. In conclusion, to confirm any model, additional experimental data that have an appropriate resolution are needed. Interaction data that would confirm contacts between neighboring subcomplexes would likely be the most helpful tool. Regardless, the Y-complex structure presented here at least provides a benchmark that any reasonable model needs to be consistent with.

Methods

Construct generation

Nup120$_{952\text{-}1241}$, Nup145C$_{233\text{-}791}$, Nup85$_{257\text{-}1181}$, and full-length Sec13 were cloned from *Myceliophthora thermophila (Mt).* Nup120$_{952\text{-}1241}$ was C-terminally fused with a 10xHis tag and cloned into a kanamycin resistant, T7-promoter-based bacterial expression vector. To increase stability, full-length Sec13 was fused C-to-N-terminally to

Nup145C$_{147-705}$ using a flexible four-residue linker, in analogy to a previously described *S. cerevisiae* construct (Brohawn and Schwartz, 2009). The Sec13-Nup145C fusion was N-terminally tethered to a 3C protease cleavable SUMO tag and cloned into an ampicillin resistant, T7-promoter-based bacterial expression vector. Nup85$_{257-1181}$ was N-terminally fused with a 3C protease cleavable 10xHis-8xArg-SUMO tag. The tetrameric complex is referred to as Nup120-Sec13-Nup145C-Nup85 for simplicity.

Protein production and purification

Nup120 and Sec13-Nup145C vectors were co-transformed into *Escherichia coli* LOBSTR-RIL(DE3) (Kerafast) (Andersen et al., 2013) cells and protein production was induced with 0.2 mM IPTG at 18 °C for 12-14 h. Nup85 was expressed separately under identical conditions. Cells were collected by centrifugation at 6,000g, resuspended in lysis buffer (50 mM potassium phosphate (pH 8.0), 500 mM NaCl, 30 mM imidazole, 3 mM β-mercaptoethanol (βME), 1 mM PMSF) and lysed using a cell disruptor (Constant Systems). The lysate was cleared by centrifugation at 12,500g for 15min. The soluble fraction was incubated with Ni Sepharose 6 Fast Flow beads (GE Healthcare) for 30min on ice. After washing the beads with lysis buffer the protein was eluted (250 mM imidazole, pH8.0, 150 mM NaCl, 3 mM βME). The Nup85 Ni-eluate was diluted 1:1 with 20 mM Hepes·KOH (pH 7.4), 0.1 mM EDTA, 1 mM dithiothreitol (DTT) and subjected to cation exchange chromatography on a HiTrapS column (GE Healthcare) using a linear NaCl gradient. The Nup120-Sec13-Nup145C Ni-eluate was incubated with 3C and dialyzed overnight at 4 °C (20 mM Hepes·KOH (pH 7.4), 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT). Both samples were further purified separately via size exclusion chromatography on a Superdex S200 16/60 column (GE Healthcare) equilibrated in running buffer (10 mM Tris·HCl (pH 8.0), 150 mM NaCl, 1 mM NaCl, 1mM DTT). Purified 10xHis-8xArg-SUMO-Nup85 was incubated with an excess of purified Nup120-Sec13-Nup145C and the assembled quaternary complex was separated via cation-exchange chromatography. Nup120-Sec13-Nup145C-Nup85 complex was incubated with 3C overnight at 4 °C and subjected to a final purification by size exclusion chromatography on Superdex S200 (GE Healthcare). Selenomethionine (SeMet) derivatized Nup85 and Nup120-Sec13-Nup145C-Nup85 was prepared as described previously (Brohawn et al., 2008) and purified as the native version, except that the reducing agent concentration (βME, DTT) was 5 mM in all buffers.

Crystallization

Initial crystals of Nup120-Sec13-Nup145C-Nup85 were obtained at 18 °C in 9 days in sitting drops over a reservoir containing 50 mM Tris·HCl (pH 8.5), 1 M NH_4SO_4. 10 mM EDTA as an additive reduced nucleation and improved the shape of the crystals. Hanging drops of 1 μl protein at 3 mg/ml and 1 μl of precipitant (50 mM Tris·HCl (pH 8.23), 0.7 M $(NH_4)_2SO_4$, 20 mM EDTA) were supplemented on the third day of incubation with 0.2 μl seed stock to yield diffraction quality, thin triangular plates. Selenomethionine-derivatized protein crystallized under identical conditions. Crystals were cryo-protected in mother liquor supplemented with 20 % (v/v) ethylene glycol.

Data collection and structure determination

Data collection was performed at the Advanced Photon Source end station 24-IDC. All data processing steps were carried out with programs provided through SBgrid (Morin et al., 2013). Data reduction was performed using HKL2000 (Otwinowski and Minor, 1997). Sec13 from *S. cerevisiae* (3MZK, chain A) (Whittle and Schwartz, 2010) was used as a search model for molecular replacement (MR). One unique solution was found in a Nup85-only SeMet derivative (space group $P2_1$, 1 copy per asymmetric unit). Partial MR phases were then used to find 12 out of 20 possible SeMet positions (8 were in disordered regions) with *Phaser* within the PHENIX suite (Adams et al., 2010) using the MR-single anomalous dispersion (MR-SAD) protocol. An interpretable 4.1 Å experimental electron density map was obtained after solvent modification using *Parrot* from the CCP4 suite (Winn et al., 2011). The backbone of most of the model was traced and the sequence for Nup85 was assigned, aided by the selenium positions as markers and the partial Nup85 models available (PDB codes: 3EWE, 3F3F) (Brohawn et al., 2008; Debler et al., 2008) Next, this initial model was used as a search model for MR with the dataset obtained from fully SeMet-derivatized protein (space group P1, two copies per asymmetric unit). Two unambiguous solutions were readily found using *Phaser.* Again, MR-SAD was used to find 36 selenium sites (out of 74 possible, with 26 in disordered regions), which served as markers to assign the sequence for Nup120 and Nup145C. To build *mt*Sec13, a starting model was generated using *Rosetta* (Terwilliger et al., 2012) and Sec13 from *S. cerevisiae* (62% sequence identity). Model building was carried out with *Coot* (Emsley et al., 2010) and *phenix.refine* was used for refinement. NCS and secondary structure restraints were applied throughout the refinement process. For the final refined structure, Ramachandran values were: 91% favored, 8.2% allowed, and 0.8% outlier. The *MolProbity* (Chen et al., 2010) clashscore was 8.33 and the *MolProbity* percentile was 97%.

Generation of Y-complex composite structures

A composite structure of the minimal, hexameric Y-complex was generated in *Coot* using secondary structure match (SSM) superposition of solved, overlapping crystal structure fragments. Nup85, Sec13, and the Nup120 tail coordinates are the *M. thermophila* orthologs and are from this study. Nup120$_{1\text{-}948}$ from *S. pombe* (sp) (PDB code: 4FHM) (Bilokapic and Schwartz, 2012) was used to generate a full-length hybrid Nup120 structure together with the Mt fragment. Nup145C is a hybrid structure with the trunk and tail domains from Mt (this study) and the crown domain from *S. cerevisiae*, which allowed for the confident modeling of the Nup145C-Nup84 junction previously solved in *S. cerevisiae* (PDB code: 3JRO, 3IKO) (Brohawn and Schwartz, 2009; Nagy et al., 2009). Nup84 is a hybrid structure: with the crown and trunk domains from *S. cerevisiae* (PDB code: 3JRO) (Brohawn and Schwartz, 2009) and the tail domain from the *H. sapiens* Nup107 ortholog (PDB code: 3I4R) (Whittle and Schwartz, 2009). Four helices between the trunk and tail domains on Nup84 were modeled from the corresponding helices solved in Nic96 (PDB code: 2QX5, 2RFO) (Jeudy and Schwartz, 2007; Schrader et al., 2008) using Phyre2 one-to-one threading (Kelley and Sternberg, 2009). Nic96 was chosen as the template because superposition between its trunk and tail domains and the corresponding regions on Nup84 yielded the lowest RMSD of all solved ACE1 domain proteins. Additionally, the *BackPhyre* (Kelley and Sternberg, 2009) structure prediction server suggests that Nic96 is the closest Nup homolog to Nup84 and vice versa, despite low sequence identity of 8%. The Nup133 stacked α-helical domain and its junction with Nup107 are from *H. sapiens* (PDB code: 3I4R) (Whittle and

Schwartz, 2009). The interface between the Nup133 β-propeller (PDB code: 1XKS) (Berke et al., 2004) and the stacked α-helical domain is not known but is expected to be flexible, based on limited proteolysis data (data not shown).

Composite structures of the heptameric *S. cerevisiae* and the octameric *H. sapiens* Y-complexes were generated in *Coot* using SSM superposition onto the minimal composite structure of solved and modeled structure fragments generated by one-to-one threading with the program *Phyre2*. For the *S. cerevisiae* Y-complex composite, scNup120$_{714\text{-}1036}$ was modeled with spNup120 (PDB code: 4FHM) (Bilokapic and Schwartz, 2012) as the template. scNup85$_{495\text{-}744}$ was modeled with mtNup85. scNup145C$_{534\text{-}712}$ was models with mtNup145C. scNup133 was modeled with hsNup133 (PDB code: 1XKS, 3I4R) (Berke et al., 2004; Whittle and Schwartz, 2009). scNup84$_{425\text{-}726}$ was modeled with hsNup107 and scNic96 as described above. For the human Y-complex composite, hsNup160 was modeled with spNup120 (PDB code: 4FHM) (Bilokapic and Schwartz, 2012) as the template. hsNup85$_{1\text{-}411}$ was modeled with scNup85 (PDB codes: 3EWE, 3F3F) (Brohawn et al., 2008; Debler et al., 2008). hsNup85$_{415\text{-}652}$ was modeled with mtNup85. hsNup96$_{1\text{-}732}$ was modeled with scNup145C (PDB code: 3JRO, 3IKO) (Brohawn and Schwartz, 2009; Nagy et al., 2009). hsNup96$_{733\text{-}937}$ was modeled with mtNup145C. hsNup107$_{1\text{-}657}$ was modeled with scNup84 and scNic96 as described above. The composite structures are available upon request.

Fitting composite Y-complex structures into single particle 3-D reconstruction EM maps of Y-complexes

Composite *H. sapiens* (omitting Nup133 and Nup107$_{658\text{-}925}$) *and S. cerevisiae* Y-complex structures were fitted into the published EM maps for the respective species (EMD codes: 2443, 5152) (Bui et al., 2013) using the Fitmap tool from *Chimera* (Pettersen et al., 2004). 1000 trials were run with an apparent resolution of 33 Å for human and 35 Å for the *S. cerevisiae* composite structure. For fitting the human composite model into EMD-2443 (Bui et al., 2013), the best solution had a correlation score of 0.8235 with 116/1000 hits (Fig. 8c). The next best solution (not shown) had a correlation score of 0.8219 with 8/1000 hits. For fitting the yeast composite model into EMD-5152 (Bui et al., 2013), the best solution had a correlation score of 0.6992 with 18/1000 hits (Fig. 8g). The next best solution (not shown) had a correlation score of 0.6374 with 19/1000 hits.

Fitting composite *H. sapiens* and *S. cerevisiae* Y-complex structures into 3-D reconstruction EM tomography map of the entire NPC

A procedure similar to the one outlined in Bui et al. (Bui et al., 2013) was followed to fit both human and yeast composite Y-complex structures into the published EM map of the human NPC (EMD code: 2444) (Bui et al., 2013) using *Chimera*. To highlight features that are invariant between the cytoplasmic and nuclear ring, a consensus map was generated by multiplication of a binarized copy rotated by 180 degrees with respect to the original EM map, as in Bui et al. (Bui et al., 2013). Subsequently, a membrane mask was constructed and subtracted, resulting in a membrane free consensus map, which was used for all fitting calculations. To avoid excluding alternate conformations resulting from flexibility of the long stalk of the Y-complex, Nup133 was omitted from the computational fitting procedure and was fitted manually. 30000 trials were run with an apparent resolution of 35 Å for both human and yeast composite Y-complex structures. For the human composite model, the best solution (conformation 1, Fig. 10b) had a correlation score of 0.7918 with 179/30000 hits. The second best solution (conformation 2, Fig. 10b) had a correlation score of 0.7661 with 252/30000 hits. The third best solution (inner ring, Fig. 11) had a correlation score of 0.7516 with 94/30000 hits. Other lower scoring solutions were not considered further based on steric considerations. Similar results were obtained when using the yeast composite Y-complex structure as a search model.

Multiple sequence alignment generation

Multiple sequence alignments were generated for all components of the hub structure using sequences representative of the eukaryotic tree of life (Ciccarelli et al., 2006; Fritz-Laylin et al., 2010). Sequences, except for Nup85 and Nup120, were obtained via *BLAST* searches and aligned using *Muscle* in *JalView* (Waterhouse et al., 2009). Nup85 and Nup120 multiple sequence alignments were obtained from Neumann et al. (Neumann et al., 2010) pruned, and realigned using *ClustalW*.

Yeast Plasmid Construction

The entire *NUP145*, *NUP85*, and *NUP120* genes with their endogenous promoters and terminators (~1000 nucleotides before the start codon and ~400 after the stop codon, respectively) were separately cloned into the multiple cloning site of the centromeric YCplac33 shuttle vector. Additionally, the entire *NUP145* and *NUP85* genes with their endogenous promoters and terminators were separately cloned into the multiple cloning site of the centromeric prs315 vector. All cloning was performed following the standard Gibson assembly method (Gibson et al., 2009). The constructed wild type vectors used in this study were named as follows: GKYp01 (*NUP120*, *URA3*, *CEN*), GKYp02 (*NUP85*, *URA3*, *CEN*), GKYp03 (*NUP145*, *URA3*, *CEN*), GKYp04 (*NUP85*, *LEU2*, *CEN*), and GKYp05 (*NUP145*, *LEU2*, *CEN*).

Vectors where the last C-terminal alpha helix of Nup120 (Δα30, Δ1016-1037), Nup85 (Δα30, Δ719-744), and Nup145C (Δα27, Δ689-712) has been deleted were also generated using GKYp01, GKYp04, and GKYp05, respectively. The deletion vectors were created by introducing an early stop codon using the Quikchange site-directed mutagenesis method. The constructed deletion vectors used in this study were named as follows: GKYp06 (*NUP120*Δα30, *URA3*, *CEN*), GKYp07 (*NUP85*Δα30, *LEU2*, *CEN*), and GKYp08 (*NUP145*Δα27, *LEU2*, *CEN*).

Yeast Strain Construction

Yeast strains used in this study are listed in Table 4. The Δ*NUP120* haploid strain was obtained from the *S. cerevisiae* deletion consortium (background BY4742, #14906), transformed with either the YCplac33 empty vector (negative control), GKYp01 (positive control), or GKYp06 (truncation), and selected on plates lacking uracil. The diploid Δ*NUP85* and Δ*NUP145* strains were also obtained from the *S. cerevisiae* deletion consortium (background BY4743, #26840 and #24459, respectively), transformed with either GKYp02 or GKYp03, respectively, and selected on uracil negative plates. The transformed diploids were then sporulated and subjected to tetrad dissection. The resulting Δ*NUP85* haploids containing GKYp02 were transformed with either prs315 empty vector (negative control), GKYp04 (positive control), or GKYp07 (truncation), whereas the Δ*NUP145* haploids containing GKYp03 were transformed with either prs315 empty vector (negative control), GKYp05 (positive control), or GKYp08 (truncation). The Δ*NUP85* and Δ*NUP145* haploid transformations were selected on plates lacking leucine.

Cell Growth Analysis

The Δ*NUP120* strains containing YCplac33, GKYp01, or GKYp06 were grown as a liquid culture in YPD overnight shaking at 30 °C. The cultures were then diluted in YPD to an OD_{600} of 0.1 and aliquoted into a 96-well plate (100 μl of culture / well). The plate was placed into a Synergy 2 multi-mode microplate reader (BioTek) and the OD_{600} of all the wells containing a culture were monitored every 15 min for 24 h. The plate was shook continuously and kept at 30 °C.

The Δ*NUP85* haploid strains containing GKYp02 and prs315, GKY04, or GKY07 and the Δ*NUP145* haploid strains containing GKYp03 and prs315, GKY05, or GKY08 were grown as a liquid culture in YPD overnight shaking at 30 °C. The cultures were then diluted in either SC medium (containing all 20 amino acids, uracil and 2% glucose) or SC medium containing 5-fluoroorotic acid (5-FOA) at 1 mg/ml. The plate was placed into a Synergy 2 multi-mode microplate reader and the OD_{600} of all the wells containing a culture were monitored every 15 min for 24 h. The plate was shook continuously and kept at 30 °C.

The growth of three separate clones of each strain was tested in quintuplicates. The data was graphed in Prism (GraphPad Software). All error bars represent the standard deviation of the mean.

PDB accession code. Coordinates and structure factors have been deposited in the Protein Data Bank under accession code PDB 4YCZ.

Acknowledgements

The X-ray crystallography work was conducted at the APS NE-CAT beamlines, which are supported by award GM103403 from the National Institute of General Medical Sciences, US National Institutes of Health (NIH). Use of the APS is supported by the US Department of Energy, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357. We thank K. Rajashankar (APS NE-CAT) for help in phasing the structure; E. Brignole (MIT) for help with generating the cryo-ET consensus map; L. Berchowitz and A. Amon (MIT) for help with the *in vivo* fitness analysis. Research was supported by the US NIH under grant number R01GM77537 (T.U.S.) and T32GM007287 (K.K. and K.E.K.) and the US National Science Foundation Graduate Research Fellowship under Grant No. 1122374 (K.E.K.).

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

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. (2007). The molecular architecture of the nuclear pore complex. Nature *450*, 695–701.

Algret, R., Fernandez-Martinez, J., Shi, Y., Kim, S.J., Pellarin, R., Cimermancic, P., Cochet, E., Sali, A., Chait, B.T., Rout, M.P., et al. (2014). Molecular architecture and function of the SEA complex, a modulator of the TORC1 pathway. Mol. Cell. Proteomics *13*, 2855–2870.

Andersen, K.R., Leksa, N.C., and Schwartz, T.U. (2013). Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins *81*, 1857–1861.

Bar-Peled, L., Chantranupong, L., Cherniack, A.D., Chen, W.W., Ottina, K.A., Grabiner, B.C., Spear, E.D., Carter, S.L., Meyerson, M., and Sabatini, D.M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science *340*, 1100–1106.

Belgareh, N., Rabut, G., Baï, S.W., van Overbeek, M., Beaudouin, J., Daigle, N., Zatsepina, O.V., Pasteau, F., Labas, V., Fromont-Racine, M., et al. (2001). An evolutionarily conserved NPC subcomplex, which redistributes in part to kinetochores in mammalian cells. J. Cell Biol. *154*, 1147–1160.

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.

Bilokapic, S., and Schwartz, T.U. (2012). Molecular basis for Nup37 and ELY5/ELYS recruitment to the nuclear pore complex. Proc. Natl. Acad. Sci. USA *109*, 15241–15246.

Boehmer, T., Jeudy, S., Berke, I.C., and Schwartz, T.U. (2008). Structural and functional studies of Nup107/Nup133 interaction and its implications for the architecture of the nuclear pore complex. Mol. Cell *30*, 721–731.

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

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.R., and Schwartz, T.U. (2009). The nuclear pore complex has entered the atomic age. Structure *17*, 1156–1168.

Bui, K.H., Appen, von, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M.-T., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., et al. (2013). Integrated structural analysis of the human nuclear pore complex scaffold. Cell *155*, 1233–1243.

Chen, V.B., Arendall, W.B.I., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Ciccarelli, F.D., Doerks, T., Mering, von, C., Creevey, C.J., Snel, B., and Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. Science *311*, 1283–1287.

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.

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.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Fernandez-Martinez, J., Phillips, J., Sekedat, M.D., Diaz-Avalos, R., Velazquez-Muriel, J., Franke, J.D., Williams, R., Stokes, D.L., Chait, B.T., Sali, A., et al. (2012). Structure-function mapping of a heptameric module in the nuclear pore complex. J. Cell Biol. *196*, 419–434.

Fritz-Laylin, L.K., Prochnik, S.E., Ginger, M.L., Dacks, J.B., Carpenter, M.L., Field, M.C., Kuo, A., Paredez, A., Chapman, J., Pham, J., et al. (2010). The genome of Naegleria gruberi illuminates early eukaryotic versatility. Cell *140*, 631–642.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

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.

Hetzer, M.W., and Wente, S.R. (2009). Border control at the nucleus: biogenesis and organization of the nuclear membrane and pore complexes. Dev. Cell *17*, 606–616.

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

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.

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. *16*, 782–788.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc *4*, 363–371.

Kim, S.J., Fernandez-Martinez, J., Sampathkumar, P., Martel, A., Matsui, T., Tsuruta, H., Weiss, T.M., Shi, Y., Markina-Inarrairaegui, A., Bonanno, J.B., et al. (2014). Integrative structure-function mapping of the nucleoporin Nup133 suggests a conserved mechanism for membrane anchoring of the nuclear pore complex. Mol. Cell. Proteomics *13*, 2911–2926.

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.

Liu, X., Liu, X., Mitchell, J.M., Wozniak, R.W., Blobel, G., Fan, J., and Fan, J. (2012). Structural evolution of the membrane-coating module of the nuclear pore complex. Proc. Natl. Acad. Sci. USA *109*, 16498–16503.

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.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., and Sliz, P. (2013). Collaboration gets the most out of software. eLife *2*, e01456.

Nagy, V., Hsia, K.-C., Debler, E.W., Kampmann, M., Davenport, A.M., Blobel, G., and Hoelz, A. (2009). Structure of a trimeric nucleoporin complex reveals alternate oligomerization states. Proc. Natl. Acad. Sci. USA *106*, 17693–17698.

Neumann, N., Lundin, D., and Poole, A.M. (2010). Comparative genomic evidence for a complete nuclear pore complex in the last eukaryotic common ancestor. PLoS ONE *5*, e13241.

Ori, A., Banterle, N., Iskar, M., Andrés-Pons, A., Escher, C., Khanh Bui, H., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., et al. (2013). Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. Mol. Syst. Biol. *9*, 648.

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

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605–1612.

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

Ries, J., Kaplan, C., Platonova, E., Eghlidi, H., and Ewers, H. (2012). A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat. Methods *9*, 582–584.

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.

Sampathkumar, P., Gheyi, T., Miller, S.A., Bain, K.T., Dickey, M., Bonanno, J.B., Kim, S.J., Phillips, J., Pieper, U., Fernandez-Martinez, J., et al. (2011). Structure of the C-terminal domain of Saccharomyces cerevisiae Nup133, a component of the nuclear pore complex. Proteins *79*, 1672–1677.

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

Seo, H.-S., Ma, Y., Debler, E.W., Wacker, D., Kutik, S., Blobel, G., and Hoelz, A. (2009). Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. Proc. Natl. Acad. Sci. USA *106*, 14281–14286.

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.

Strambio-de-Castillia, C., Niepel, M., and Rout, M.P. (2010). The nuclear pore complex: bridging nuclear transport and gene regulation. Nat. Rev. Mol. Cell. Biol. *11*, 490–501.

Stuwe, T., Correia, A.R., Lin, D.H., Paduch, M., Lu, V.T., Kossiakoff, A.A., and Hoelz, A. (2015). Nuclear pores. Architecture of the nuclear pore complex coat. Science *347*, 1148–1152.

Süel, K.E., Cansizoglu, A.E., and Chook, Y.M. (2006). Atomic resolution structures in nuclear transport. Methods *39*, 342–355.

Szymborska, A., de Marco, A., Daigle, N., Cordes, V.C., Briggs, J.A.G., and Ellenberg, J. (2013). Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. Science *341*, 655–658.

Terwilliger, T.C., DiMaio, F., Read, R.J., Baker, D., Bunkóczi, G., Adams, P.D., Grosse-Kunstleve, R.W., Afonine, P.V., and Echols, N. (2012). phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. J. Struct. Funct. Genomics *13*, 81–90.

Thierbach, K., Appen, von, A., Thoms, M., Beck, M., Flemming, D., and Hurt, E. (2013). Protein interfaces of the conserved Nup84 complex from Chaetomium thermophilum shown by crosslinking mass spectrometry and electron microscopy. Structure *21*, 1672–1682.

Vasu, S., Shah, S., Orjalo, A., Park, M., Fischer, W.H., and Forbes, D.J. (2001). Novel vertebrate nucleoporins Nup133 and Nup160 play a role in mRNA export. J. Cell Biol. *155*, 339–354.

Vollmer, B., and Antonin, W. (2014). The diverse roles of the Nup93/Nic96 complex proteins - structural scaffolds of the nuclear pore complex with additional cellular functions. Biol. Chem. *395*, 515–528.

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

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

Whittle, J.R.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.

Whittle, J.R.R., and Schwartz, T.U. (2010). Structure of the Sec13-Sec16 edge element, a template for assembly of the COPII vesicle coat. J. Cell Biol. *190*, 347–361.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

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.

Tables

Table 1. Data collection and refinement statistics

	Nup85(SeMet)$_{257-1181}$, Nup120$_{952-1241}$, Sec13-Nup145C$_{233-791}$	Nup85(SeMet)$_{257-1181}$, Nup120(SeMet)$_{952-1241}$, Sec13-Nup145C(SeMet)$_{233-791}$
Data collection		
Space group	C2	P1
Cell dimensions		
a, *b*, *c* (Å)	104.98, 212.02, 170.64	118.96, 107.67, 163.09
α, β, γ (°)	90, 107.2, 90	74.3, 80.4, 63.2
Resolution (Å)	163 - 4.10 (4.25 - 4.10)[a]	157 - 4.00 (4.14 - 4.00)
R_{sym}	0.19 (1.00)	0.21 (0.97)
$I / \sigma I$	8.9 (0.9)	5.3 (0.8)
Completeness (%)	98.2 (93.5)	90.3 (81.3)
Redundancy	6.1 (4.7)	3.6 (2.9)
Refinement		
Resolution (Å)	163 - 4.10	
No. reflections	53648	
$R_{\text{work}} / R_{\text{free}}$	31.9/35.8	
No. atoms	10070	
Protein	10070	
B factors		
Protein	161.7	
r.m.s. deviations		
Bond lengths (Å)	0.002	
Bond angles (°)	0.64	

[a] Values in parentheses are for highest-resolution shell. One crystal was used for each dataset.

Protein	**Fully Resolved Residues[1]**	**Backbone Resolved Residues**	**Total Resolved Residues**	**Total Residues**
Sec13	239 (79%)	31 (10%)	270 (89%)	304
Nup145C$_{233-791}$	257 (46%)	188 (34%)	445 (80%)	559
Nup85$_{257-1181}$	424 (46%)	166 (18%)	590 (64%)	925
Nup120$_{952-1241}$	134 (46%)	63 (22%)	197 (68%)	290

Table 2. Proportion of residues and side chains resolved in the crystal structure.

[1] Percentage refers to the percent of total residues in the crystallization construct that are fully resolved (with side chains), backbone resolved (no side chains), or the sum of the two (total resolved).

	M. thermophila* to *S. cerevisiae		***M. thermophila* to *H. sapiens***		**S. cerevisiae to H. sapiens**	
Protein	**Identity**	**Similarity**	**Identity**	**Similarity**	**Identity**	**Similarity**
Nup85	12%	25%	14%	25%	13%	52%
Nup120	11%	46%	12%	29%	10%	31%
Nup145C	14%	43%	20%	40%	17%	40%
Nup84	17%	54%	17%	51%	19%	56%
Nup133	14%	41%	16%	44%	13%	46%
Sec13	60%	75%	50%	68%	50%	66%
Seh1	25%	54%	22%	52%	34%	76%

Table 3. Sequence Identity and similarity between *M. thermophila*, *S. cerevisiae*, and *H. sapiens* Y-complex nucleoporins.

Table 4. Yeast strains used in the study.

Plasmid	Description	Parent Plasmid
YCplac22	CEN plasmid with Ura- selection	----------
prs315	CEN plasmid with His- selection	----------
GKYp01	Full length Nup120 + promoter and terminator	YCplac33
GKYp02	Full length Nup85 + promoter and terminator	YCplac33
GKYp03	Full length Nup145 + promoter and terminator	YCplac33
GKYp04	Full length Nup85 + promoter and terminator	prs315
GKYp05	Full length Nup145 + promoter and terminator	prs315
GKYp06	Nup120 (1-1015; Δ α30) + promoter and terminator	GKYp01
GKYp07	Nup85 (1-718; Δ α30) + promoter and terminator	GKYp04
GKYp08	Nup145 (1-1293; Δ α27) + promoter and terminator	GKYp05

Strain	Description	Parent Strain
GKY112	nup120Δ::KanMX MATα his3Δ leu2Δ ura3Δ	----------
GKY113	NUP85/nup85Δ::KanMX MATα/MATA his3Δ/his3Δ leu2Δ/leu2Δ ura3Δ/ura3Δ	----------
GKY114	NUP145/nup145Δ::KanMX MATα/MATA his3Δ/his3Δ leu2Δ/leu2Δ ura3Δ/ura3Δ	----------
GKY115	nup120Δ::KanMX with YCplac22	GKY112
GKY116	nup120Δ::KanMX with GKYp01	GKY112
GKY117	nup120Δ::KanMX with GKYp06	GKY112
GKY121	NUP85/nup85Δ::KanMX with GKYp02	GKY113
GKY122	NUP145/nup145Δ::KanMX with GKYp03	GKY114
GKY123	nup85Δ::KanMX MATα with GKYp02	GKY113
GKY124	nup145Δ::KanMX MATα with GKYp04	GKY114
GKY125	nup85Δ::KanMX MATα with GKYp02 and prs315	GKY123
GKY125	nup85Δ::KanMX MATα with GKYp02 and GKYp04	GKY123
GKY125	nup85Δ::KanMX MATα with GKYp02 and GKYp07	GKY123
GKY126	nup145Δ::KanMX MATα with GKYp03 and prs315	GKY124
GKY127	nup145Δ::KanMX MATα with GKYp03 and GKYp05	GKY124
GKY128	nup145Δ::KanMX MATα with GKYp03 and GKYp08	GKY124

Figures

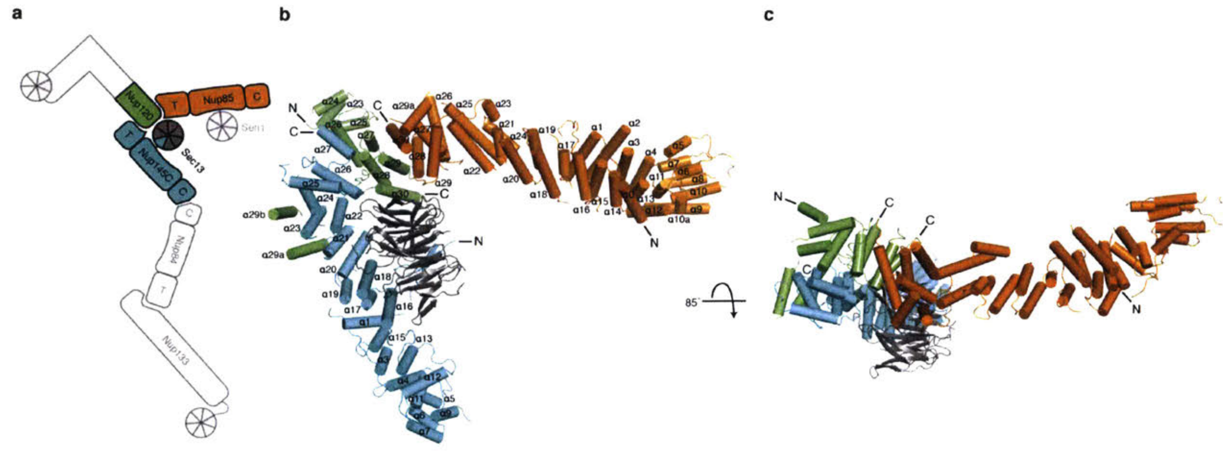


Figure 1. Structure of the *Myceliophthora thermophila* Y-complex hub at 4.1 Å resolution. (**a**) Schematic of the Y-complex. Regions included in the crystallization construct are colored, other Y-complex regions in grey. Elements of ACE1 fold proteins are indicated: T – tail and C – crown flank the central trunk element. (**b**) The hub structure is colored as follows: Nup85 (orange), Nup120 (green), Nup145C (cyan), and the Sec13 β-propeller (grey). N and C termini are indicated for the helical proteins. Helices are numbered according to previously solved *S. cerevisiae* fragments (Bilokapic and Schwartz, 2012; Brohawn and Schwartz, 2009; Brohawn et al., 2008). Numbers that include a letter modifier indicate helical elements not present in *S. cerevisiae*. (**c**) Top-down view of the hub. The N terminus of Nup145C is not indicated because it is obscured by the 85° rotation.

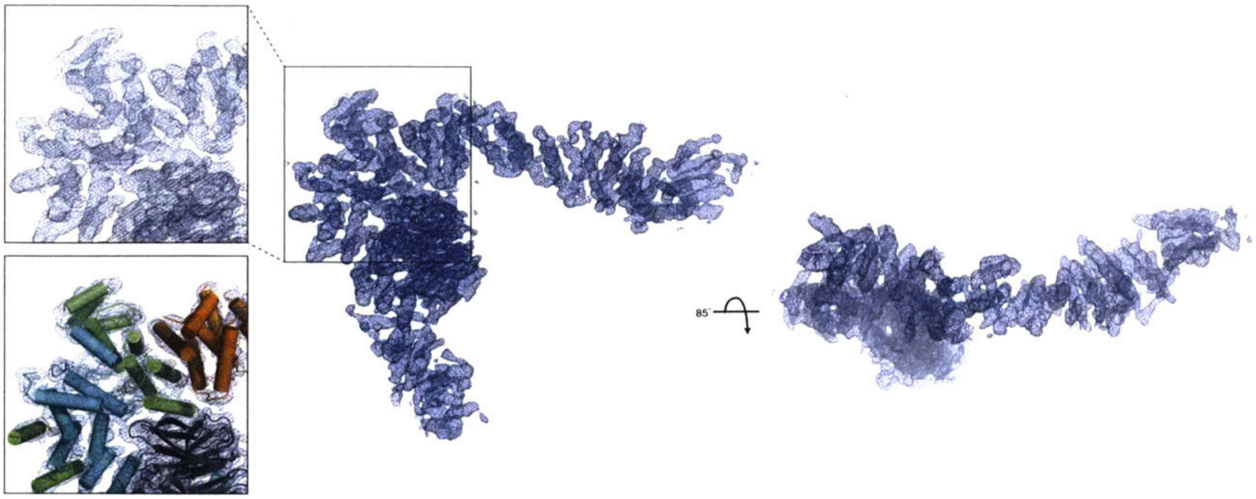


Figure 2. Electron density map for hub structure. Final 2Fo-Fc electron density map for the hub contoured at 1.5 σ, shown in the same view as Fig. 1 (Inset, top) zoom-in on the interaction region of the hub. (Inset, bottom) the hub structure is placed into the electron density.

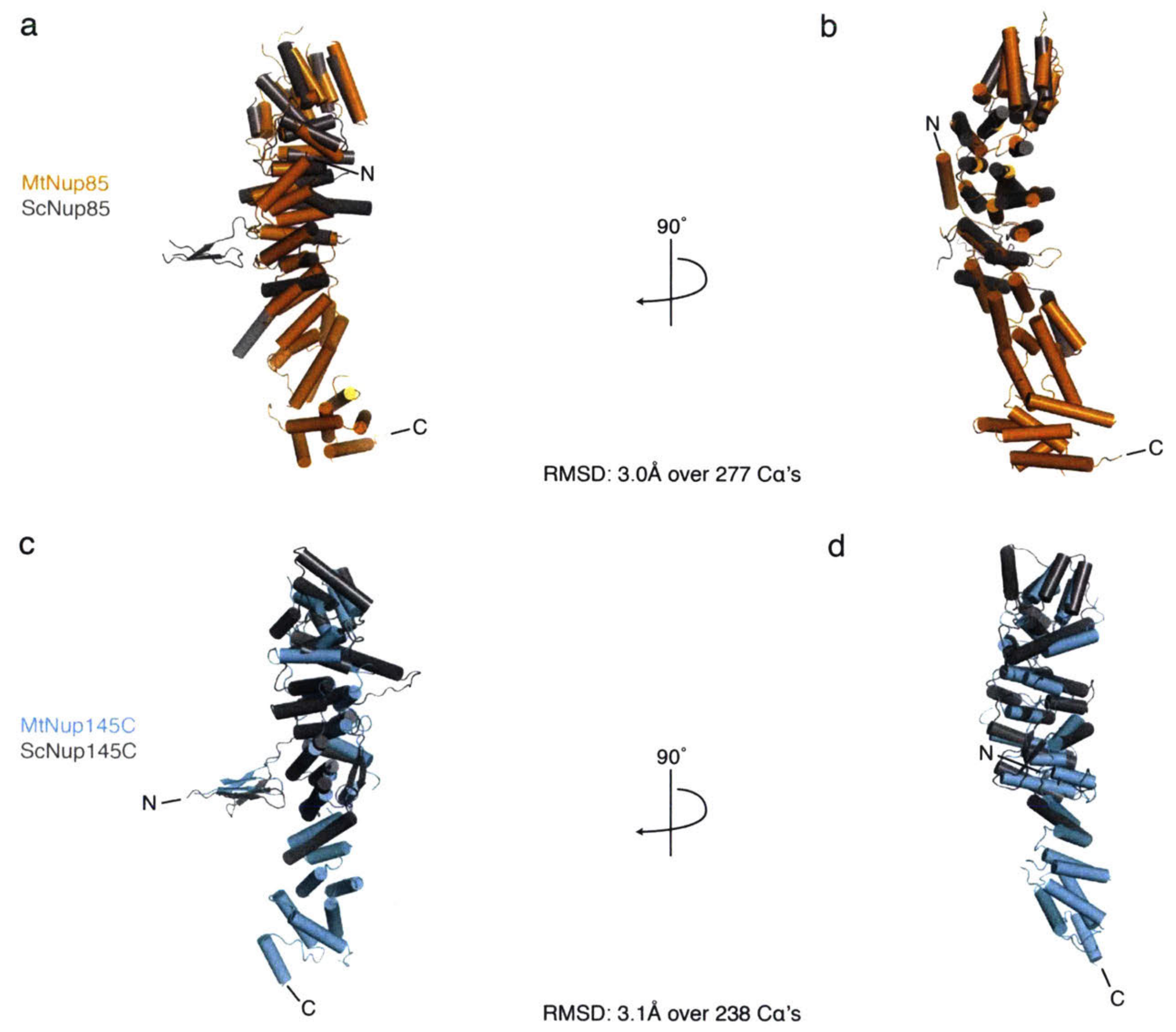
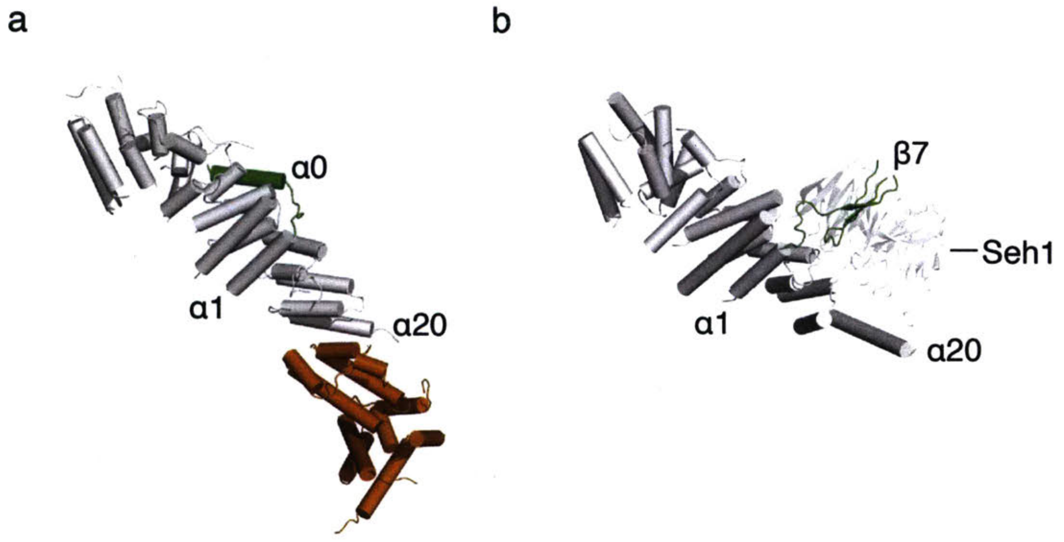


Figure 3. Superposition of *M. thermophila* and *S. cerevisiae* ACE1 proteins of the hub. (**a**) Overlay of mtNup85 (orange) and scNup85 (grey) reveals structural conservation, despite low (14%) sequence identity. N and C termini of the mtNups are labeled. (**b**) 90° rotation of the superposed structures. (**c**) Overlay of mtNup145C (cyan) and scNup145C (grey) (20% identity). (**d**) 90° rotation of the superposed structures.



M. thermophila
Nup85
This study

S. cerevisiae
Nup85/Seh1
PDB: 3EWE

Figure 4. mtNup85 lacks an Seh1 binding motif. (**a**) mtNup85, where the structural element N-terminal of α1 is an additional helix, α0 (green). The crown and trunk elements, which align to the solved scNup85 fragment, are shown in grey. The tail element, not solved in *S. cerevisiae*, is shown in orange. (**b**) scNup85, aligned to mtNup85, where the Seh1 insertion blade, β7 (green), is N-terminal to α1. Seh1 is labeled and scNup85 is shown in grey.

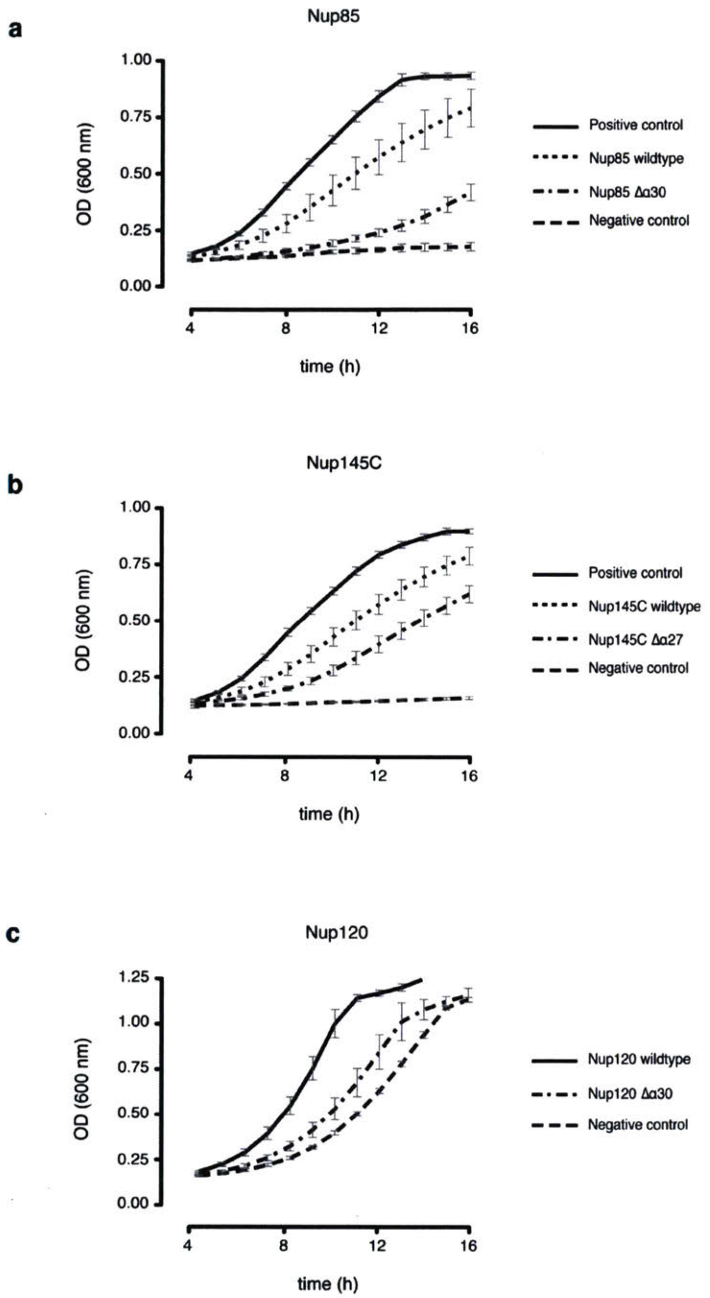


Figure 5. Fitness analysis of hub interactions. (**a**) Growth curves of *NUP85Δ* strains carrying *NUP85:URA3* and either empty pRS315 (negative control), Nup85 wildtype, or Nup85 Δα30 grown in the presence of 5-FOA. The positive control is the *NUP85Δ* strain carrying *NUP85/URA3* and empty pRS315 grown in the absence of 5-FOA. (**b**) Growth curves of *NUP145Δ* strains carrying *NUP145/URA3* and either empty pRS315 (negative control), Nup145C wildtype, or Nup145C Δα27 grown in the presence of 5-FOA. The positive control is the *NUP145Δ* strain carrying *NUP145/URA3* and empty pRS315 grown in the absence of 5-FOA. (**c**) The growth curves of *NUP120Δ* strains carrying YClac33 empty vector (negative control), Nup120 wildtype, or Nup120 Δα30 grown in YPD. Four technical replicates (*n*=4 OD measurements) for each of three biological

replicates (*n*=3), from separate colonies, were performed at 30 °C for all experiments. All error bars are standard deviation of the mean (s.e.m.).

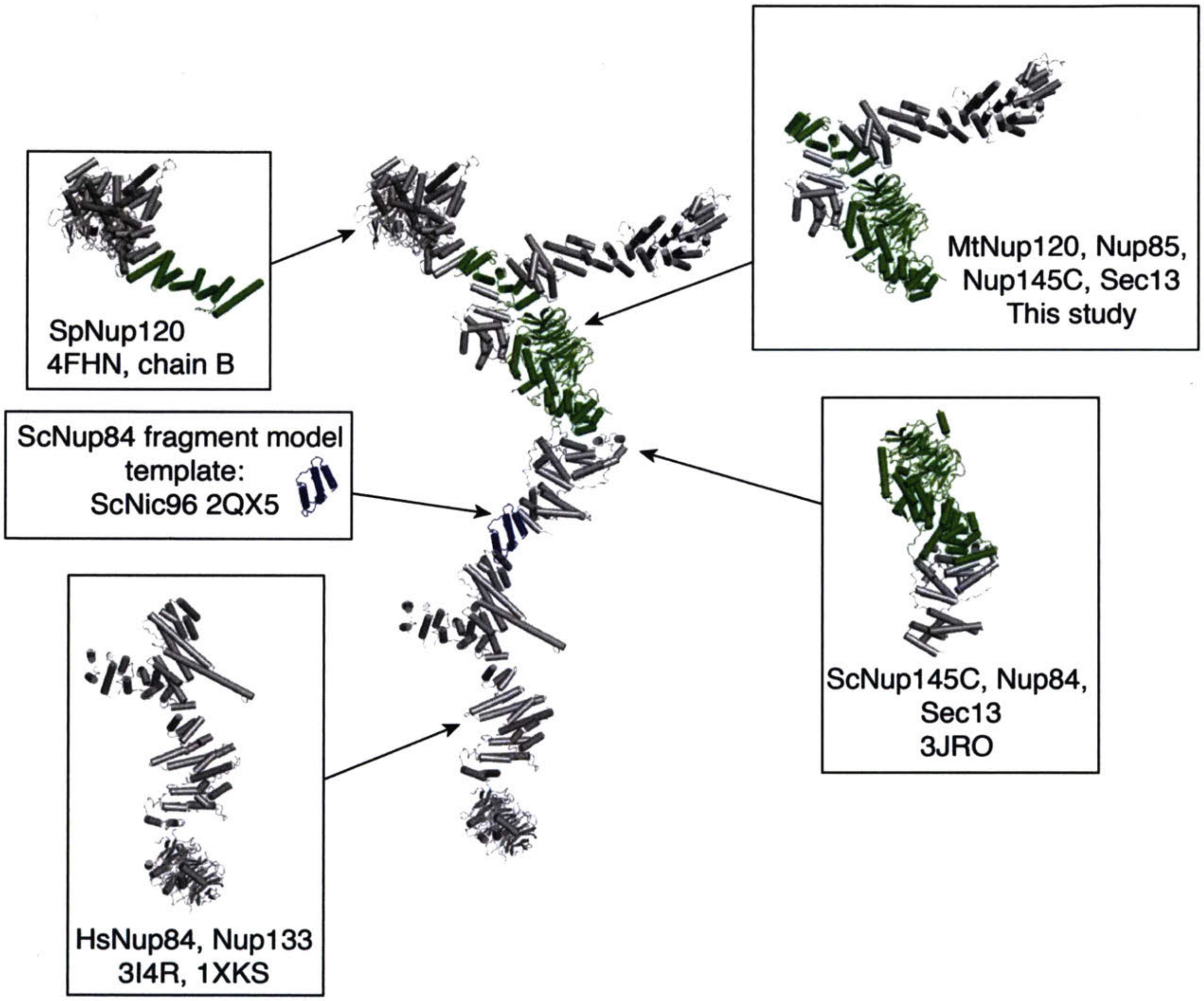


Figure 6. Composite Y-complex structure generated through overlapping crystal structure fragments. Overlapping elements between the hub and previously solved structures are shown in green. Grey elements are non-redundant crystal structures used to generate the composite. The only portion of the composite that is modeled is a four-helix bundle in Nup84 (blue).

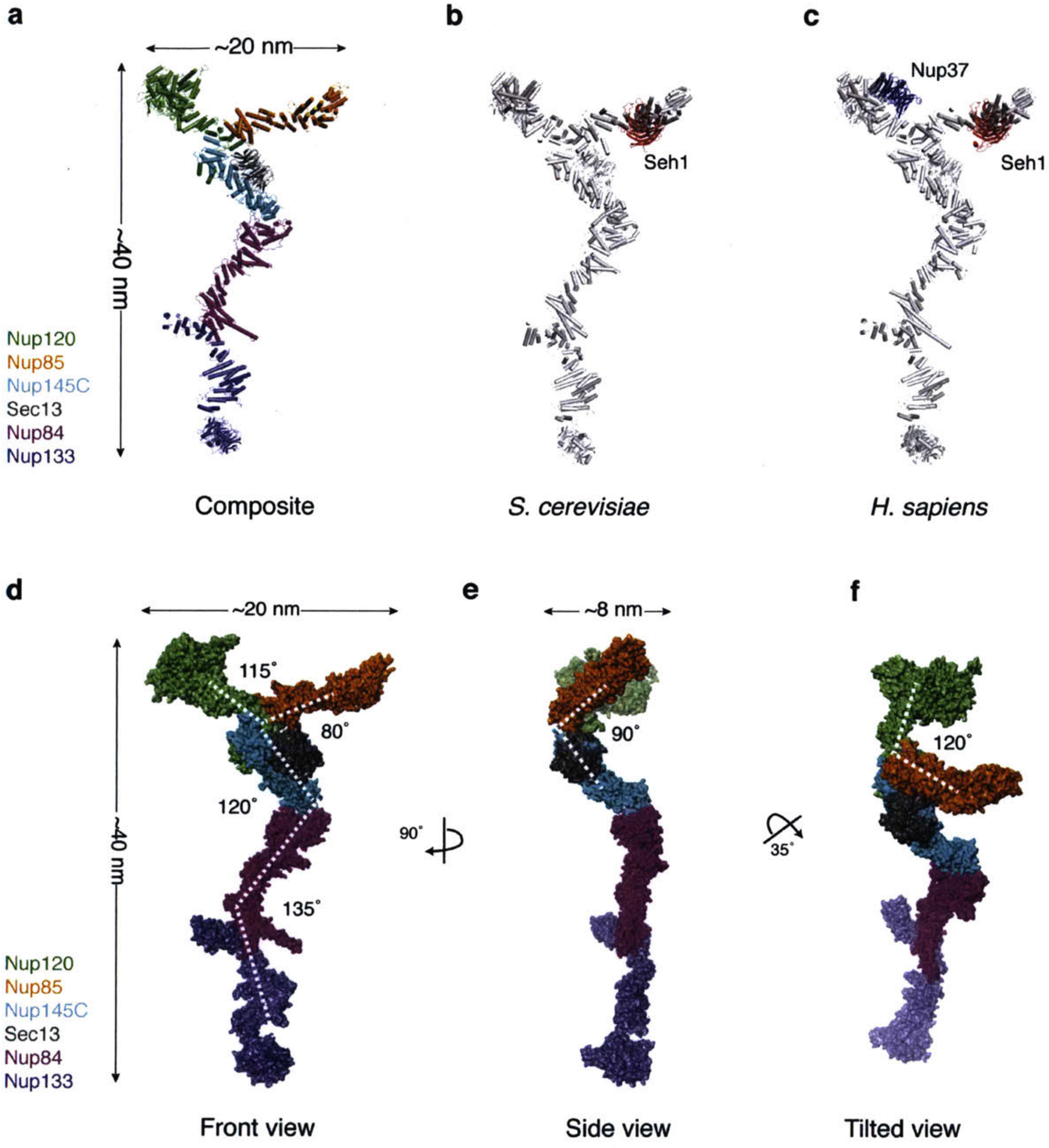
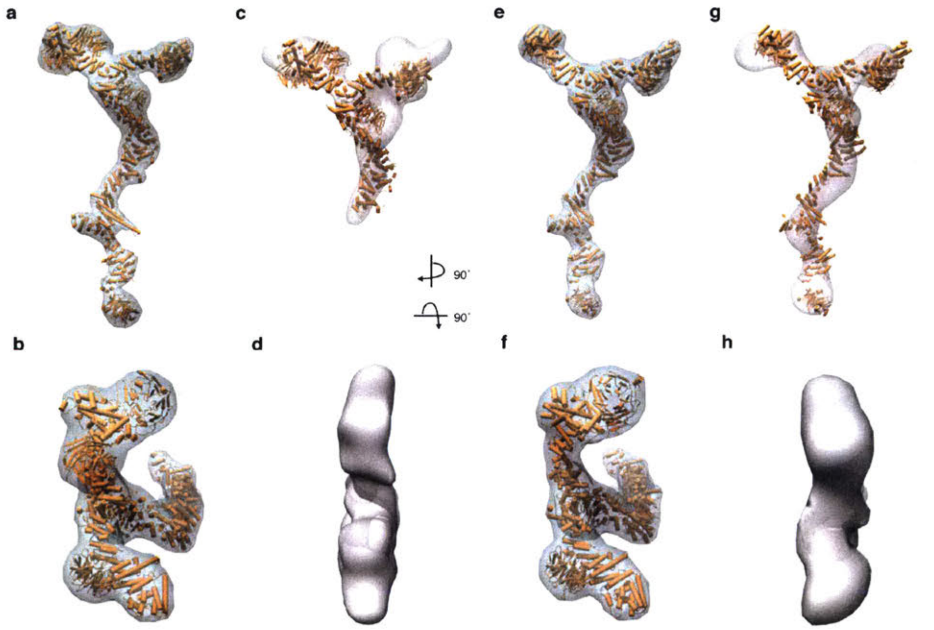


Figure 7. Composite high-resolution structure of the Y-complex. (**a**) Composite, hexameric Y-complex core constructed from the hub structure (Fig. 1) combined with previously published X-ray crystal structure fragments. Within Nup84, 4 helices were modeled computationally. (**b**) Composite *S. cerevisiae* Y-complex based on (a), with *S. cerevisiae* sequences threaded onto existing homologous structures. Compared to the universally conserved hexameric Y-complex core shown in (a), Seh1 (red) is an additional component found in many organisms, including yeast. (**c**) Composite *H. sapiens* Y-complex with *H. sapiens* sequences threaded onto existing homologous structures. Nup37 (blue) is another Y-complex component only found in a subset of

eukaryotes, including humans (Neumann et al., 2010). (**d**) Space filling surface view of the composite, hexameric Y-complex viewed from the front. (**e**) Side view. (**f**) Tilted view.



H. sapiens
35 Å
This study

H. sapiens
35 Å
Bui et al. 2013

S. cerevisiae
33 Å
This study

S. cerevisiae
33 Å
Kampmann et al. 2009

Figure 8. Comparison of the X-ray based, composite Y-complex structure with published 3-D EM reconstruction structures. (**a,b**) Electron density envelope around the composite *H. sapiens* Y-complex calculated for 35 Å resolution from front (a) or top (b) view. (**c,d**) 3-D EM reconstruction of the *H. sapiens* Y-complex with an overlay of the composite model, fitted computationally from front (c) or top (d) view. (**e,f**) Electron density envelope around the composite *S. cerevisiae* Y-complex calculated at 33 Å resolution from front (e) or top (f) view. (**g,h**) 3-D EM reconstruction of the *S. cerevisiae* Y-complex with an overlay of the composite model, fitted computationally from front (g) or top (h) view.

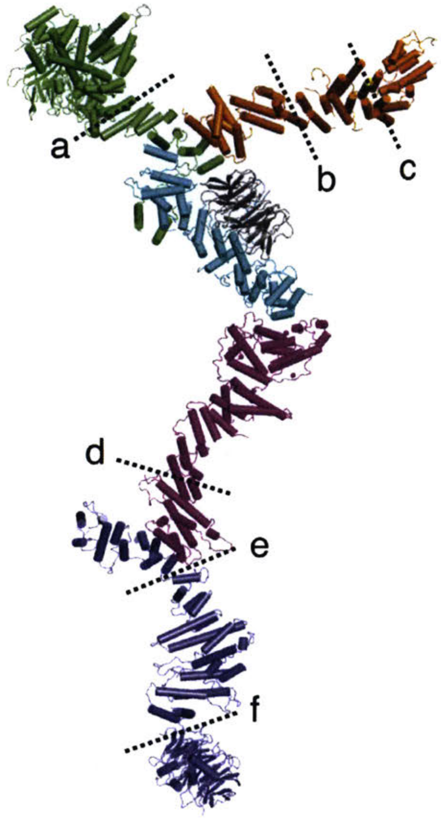


Figure 9. Flexibility of the Y-complex. Experimentally observed hinge regions of the Y-complex are denoted by dashed lines. **(a)** (Bilokapic and Schwartz, 2012). **(b–c)** (Brohawn et al., 2008). **(d–e), (**Bui et al., 2013) (Berke et al., 2004) (Boehmer et al., 2008) (Whittle and Schwartz, 2009) (Kampmann and Blobel, 2009) **(f)** (Berke et al., 2004).

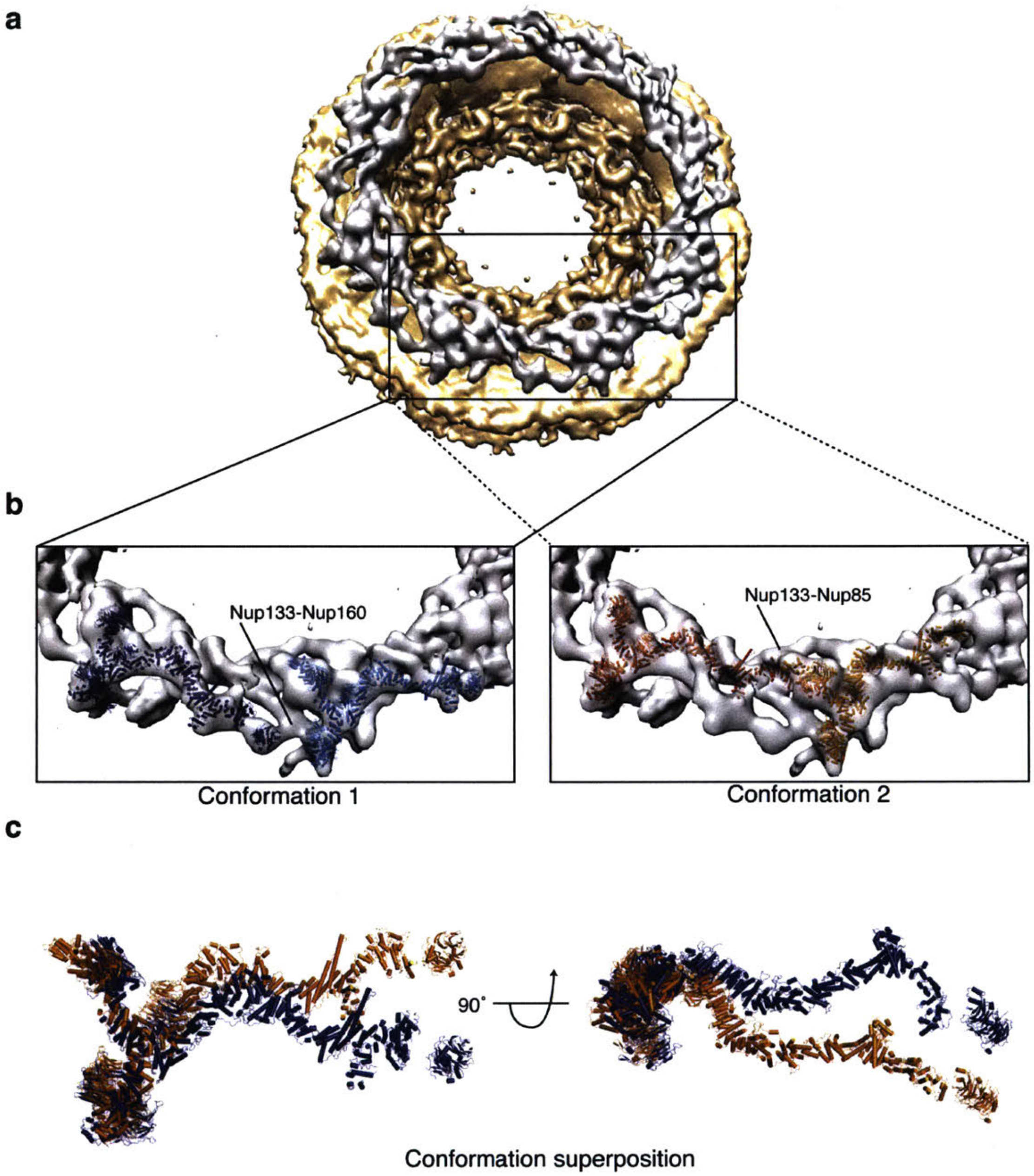
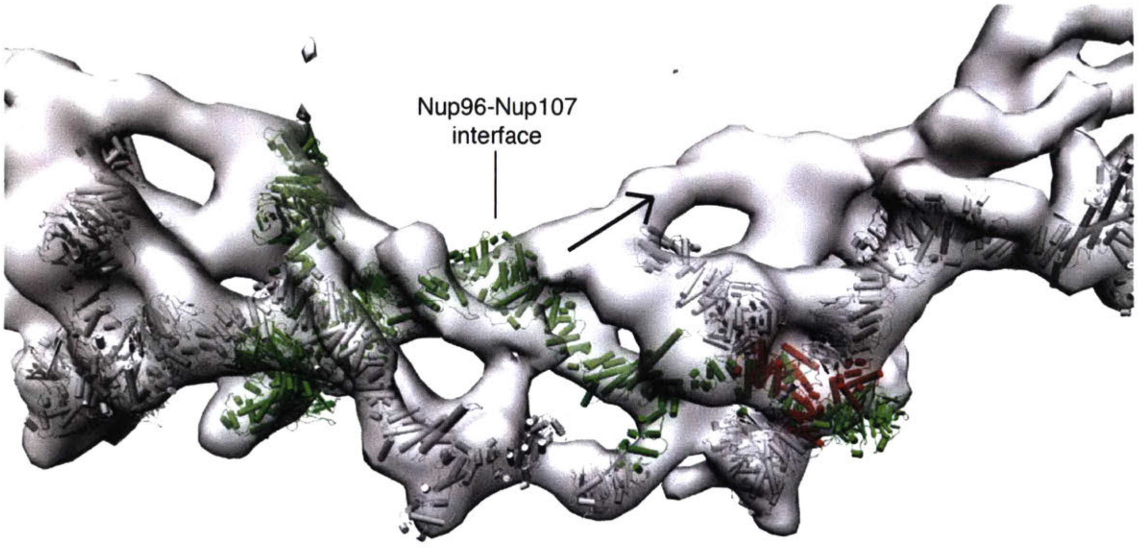


Figure 10. Fitting of the composite *H. sapiens* Y-complex into the cryo-ET map of the entire NPC. (**a**) Consensus map calculated from the cryo-ET map of the human NPC (EMD code: 2444) (Bui et al., 2013). The cytoplasmic ring density is highlighted in grey for clarity. (**b**) The top scoring fit of the composite *H. sapiens* Y-complex (conformation 1) and the second top scoring fit (conformation 2) are depicted. (**c**) A superposition of the two fits from (b) shows that they are related by a ~20° rotation about the Y-complex hub, and substantial bending of the long stack.



Bui et al. fit - Top view

Bui et al. fit - Side view

Figure 11. Fitting of the composite *H. sapiens* Y-complex into an inner ring position in the 3-D EM tomography map suffers from significant steric clashes with the outer ring fits. Inner ring fit (3rd top scoring solution) is colored green. One of the outer ring fits (1st solution) is colored grey and regions of steric clash with Nup133 of the inner ring fit are colored red. The arrow (top view) shows where the Y complex stem needs to move in order to match the Bui et al. inner ring placement and avoid steric clashes with the adjacent outer ring. Such stem placement would involve rotation at or around the Nup96-Nup107 interface, which is unlikely due to the energetic cost of disrupting the complex's hydrophobic core.

Chapter 3: Structural and Biochemical studies of the Nic96 complex

Introduction

The second large structural scaffold sub-complex of the NPC is the ~0.5 MDa Nic96 complex. Biochemically stable members of the Nic96 complex are Nic96, Nup157/170, Nup53/59, Nup188 and Nup192 (Amlacher et al., 2011; figure 1). The redundant scNup157/170 and scNup53/59 paralogs are a result of a genome duplication event during yeast evolution, but exist as a single copy in metazoa (Wolfe and Shields, 1997). The Nic96 complex is localized to the center of the transport axis and is thought to form a structural scaffold that fills the inner ring density observed in cryo-ET reconstructions of the NPC (Krull et al., 2004; Alber et al., 2007b; von Appen et al., 2015). Nic96 was first identified in *S. cerevisiae*, complexed to the FG-repeat barrier forming Nsp1 complex (Grandi et al., 1993). It is not considered to be a constituent of the Nsp1 complex, but anchors it to the NPC through its N-terminal α-helices (Schlaich et al., 1997). The entire complex is thought to be anchored to the membrane through interactions of the C-terminal amphipathic helix of Nup53 with transmembrane nucleoporin Ndc1 or inserting directly into the membrane (Marelli et al., 2001; Onischenko et al., 2009, figure 1). Thus, the Nic96 complex serves an important structural role that maintains the integrity of the NPC in addition to serving as a platform for barrier forming proteins.

A stable reconstitution of the Nic96 complex was first demonstrated by pull-down assays with thermophilic proteins from *C. thermophila.* GST-Nup53 immobilized on beads was able to recruit Nic96, Nup170 and either Nup192 or Nup188. Through competition studies, it was determined that Nup192 and Nup188 compete for binding to both Nup53 and Nic96, forming mutually exclusive tetrameric assemblies (Amlacher et al., 2011). In addition, the authors were able to coarsely map the interactions to regions of each member. To summarize their findings, they were mapped onto a cartoon representations of each member (Amlacher et al., 2011; figure 1). The 2nd of the two predicted N-terminal helices of Nic96 (R2 helix) binds to the C terminus of either Nup192 or Nup188. The N-terminal disordered region of Nup53 binds simultaneously to Nic96 and either Nup192 or Nup188. The C-terminal disordered region of Nup53 binds to Nup170 (Amlacher et al., 2011). In recent reports, the interaction mapping was further narrowed down. The R2 helix of Nic96 binds to the very C terminus of either Nup192 or Nup188 (Stuwe et al., 2014). The R1 helix nests itself into the coiled coil fold-back architecture of the Nsp1 complex (Stuwe et al., 2015). A co-crystal structure of Nup53 and Nup170 revealed that a short C-terminal stretch of Nup53 forms an α-helix that binds to the N-terminal β-propeller of Nup170 (Lin et al., 2016). Likewise, a co-crystal structure of Nup53 and Nic96 showed that a short N-terminal stretch of Nup53 forms an α-helix that binds to the crown module of Nic96 (Lin et al., 2016).

Although most folded regions of each Nic96 constituent have been structurally characterized by X-ray crystallography, assembly structures of the entire complex are still missing (Andersen et al., 2013; Stuwe et al, 2014; Jeudy and Schwartz, 2007; Schrader et al., 2008; Whittle and Schwartz, 2009; Seo et al., 2013; Handa et al., 2006). On the one hand, both Nic96 and Nup157/170 are related to Y complex proteins Nup84/ Nup145C/Nup85, and Nup133, respectively, which do form the distinct Y shape. However, contrary to the Y complex, it seems that many of the interactions within the Nic96 complex involve small interfaces separated by long flexible linkers (Lin et al., 2016; figure 1). Therefore it is not clear wether or not the assembly adopts a distinct shape.

To gain insight into the assembly structure of the Nic96 complex, we focused on the Nic96/Nup192 binding interface. Encouraged by our success in characterizing the Y complex hub, all experiments were performed with proteins from *M. thermophila.* Single particle EM analysis of negatively stained samples and limited proteolysis experiments suggests that the Nic96/Nup192 interface involves more than just the R2 helix of Nic96. The conformational heterogeneity observed in 2-D class averages suggests that the interface is flexible, but semi-ordered. Due to this inherent conformational heterogeneity, single particle cryo-EM was not pursued. Instead, to gain high resolution structural details of the interface, we attempted to capture a single state by coaxing conformational homogeneity through crystallization. Various assemblies were designed, purified, and subject to crystallization trials. Since the Nic96/Nsp1 complex binding interface is relatively close to the Nic96/Nup192 binding interface, the coiled coil region of the Nsp1 complex was included later, with the hope that it would stabilize the complex and aid in crystallization. Finally, in the same spirit as with including the Nsp1 complex, nanobodies were generated to serve as crystallization chaperones. Extensive screening yielded promising quasi crystals and phase separation, but no diffracting crystals.

Results

Purification of the dimeric Nic96/Nup192 assembly.

Full length mtNup192 was heterologously expressed in *E. coli* with an N-terminal His-tag and purified by nickel affinity (Ni). Multiple sized species were observed in the elution fraction when visualized by Coomassie stained SDS-PAGE gel, potentially due to early termination. This problem is mitigated by switching the His-tag to the C terminus. mtNic96 can be heterologously expressed in *E. coli* with an N-terminal His-tag, yielding a single prominent band in the Ni elution fraction. The first 210 residues of mtNic96 are predicted to be unstructured and were therefore not included in any of the constructs. A large proportion of both mtNup192 and mtNic96 are sequestered in the insoluble fraction. In addition, mtNic96 alone has very poor solubility. Improved yields of both proteins was achieved by co-expression of mtNup192 with a C terminal His-tag and mtNic96, suggesting mutual stabilization of the dimeric assembly. Since mtNup192 is limiting, we expect a stoichiometric dimeric complex after Ni purification. Indeed, Coomassie stained SDS-PAGE gel band intensities of the Ni elution fraction were proportional to their molecular weights, consistent with a 1:1 stoichiometry (figure 2, A). Human rhinovirus 3C protease was added to the elution fraction and incubated at 4 °C for ~3 hours to cleave solubility tags. The cleaved fraction was concentrated to 2 mg/ml (6.5 μM) and subject to size exclusion chromatography, yielding a single mono-disperse peak (figure 2, B, C).

Purification of the pentameric Nic96/Nup192/Nsp1/Nup49/Nup57 assembly.

The pentameric Nic96/Nup192/Nsp1/Nup49/Nup57 assembly was co-expressed heterologously in *E. coli* off of two plasmids and purified by Ni affinity (figure 3). Only the coiled coil regions of the Nsp1 complex were included. After Ni purification, human rhinovirus 3C protease was added to the elution fraction and incubated at 4 °C for ~3 hours to cleave solubility tags (figure 3, A). The cleaved fraction was concentrated to ~2 mg/ml (~5 μM) and subject to size exclusion chromatography, yielding a single peak corresponding to the pentameric assembly (figure 3, B, C). The first time this construct was purified, we noticed that the Nsp1 complex has a natural affinity towards Ni resin and was in excess of the Nic96 complex in the Ni elution fraction (figure 4, A, B). This was mostly mitigated by including 10% w/v glycerol in the binding buffer (figure 3, A). Interestingly, the reconstitution of the pentameric assembly from purified Nsp1 complex and Nic96/Nup192 is no longer stable on size exclusion chromatography (figure 4, C, D).

Single particle EM analysis of the dimeric Nic96/Nup192 assembly.

To structurally characterize the binding interface, single particle EM analysis was performed on negatively stained samples of purified mtNup192/mtNic96. Upon visual inspection of the raw images, significant chemical and conformational heterogeneity was observed among the particles. Despite this challenge, the large and distinctive S-shape of Nup192 was clearly visible even in the raw micrographs. Upon closer inspection, ~10% of the Nup192 particles had extra density at the very C terminus. To boost the signal of the raw micrographs, 2-D unsupervised classification was performed on ~10,000 manually picked particles. Most class averages show a well defined S-shape corresponding to Nup192 in various conformational states, but without any extra density (figure 5, A). However, in some of the class averages, there was clear extra density at the very C terminus of Nup192. Within these classes were examples where the extra density had a length consistent with the known length of the Nic96 ACE1 domain (figure 5, B). We therefore interpret this extra density as belonging to Nic96.

The R2 helix of Nic96 is connected to the ACE1 domain by a ~80 residue long flexible linker (figure 1). If the binding interface between Nic96 and Nup192 is limited only to the R2 helix, the ACE1 domain will be distant and randomly oriented with respect to Nup192 and therefore would be averaged out during classification. The observation of 2-D class averages with Nic96 right adjacent to Nup192 suggests that the binding interface involves at least one more contact site between either the crown or tail module of the ACE1 domain and the C terminus of Nup192 (figure 5, C).

The inhomogeneity observed in the EM micrographs complicates the 2-D class average analysis since only ~10% of the particles are intact Nic96/Nup192 assemblies. Dissociated particles are not very easy to distinguish from intact particles due to the low signal to noise ratio of the micrographs. Therefore, they cannot be avoided at the particle picking step. This diminishes the signal coming from intact particles, in effect blurring the Nic96 density in the class averages and exacerbating the problem with flexibility. The dissociation of Nic96 from Nup192 is potentially due to the low concentrations (~5 nM) needed to avoid overcrowding. Another possibility is that the harsh condition of the negative stain (2% w/v uranyl acetate, pH ~4) causes dissociation. Other stains with a neutral pH such as 1% w/v sodium phosphotungstate were tried, but on average, they displayed poor staining when compared to uranyl acetate, and therefore were not pursued further. Since the N terminus of Nic96 is near its Nup192 binding helix and the C terminus of Nup192, we reasoned that they could be covalently linked without disrupting the binding interface. First, the C terminus of Nup192 was genetically fused with the N terminus of Nic96. A few variants with increasing linker lengths between the two proteins were tried, but none expressed heterologously in *E. coli.* An alternative strategy is to fuse the two proteins after expression and purification. Sortase mediated fusion (Theile et al., 2013), which is heavily dependent on the concentration of the two proteins to be linked, surprisingly did not go to completion, despite the high local concentration of the two termini. Although the amount of intact assemblies increased, there was still a significant amount of dissociated particles. The SpyCatcher mediated fusion(Li et al., 2014), similarly did not go completion, in addition to producing undesirable off-target species as judged by SDS-PAGE gel. Optimization of the Sortase and SpyCatcher reactions should overcome these shortcomings and increase intact particles. This will lead to better resolved class averages.

Limited proteolysis analysis.

Limited proteolysis is a technique that can potentially identify flexible regions and uncover mutual stabilization in protein complexes by comparing proteolytic products of the individual proteins with that of the complex (Koth et al., 2003). We wanted to complement the single particle EM analysis of the Nic96/Nup192 dimer with limited proteolysis. The same molar amount of Nic96, Nup192, and the Nsp1 complex was incubated with increasing amounts of trypsin to analyze the proteolytic products of the individual proteins (figure 6, A, B). The reaction was quenched by boiling with SDS-PAGE loading buffer. The samples were run on SDS-PAGE gel and visualized by Coomassie stain. The same procedure was repeated with the dimeric Nic96/Nup192 and the pentameric Nic96/Nup192/Nsp1/Nup49/Nup57 assemblies (figure 6, C, D). The addition of Nic96 to Nup192 yields four products that are different from either Nic96 or Nup192 alone (figure 6, A, C). This is consistent with our findings from the 2-D EM class averages, but in the absence of more dramatic differences, does not add further evidence of a potentially rigid Nic96/Nup192 structure. The addition of the Nsp1 complex to Nic96/Nup192 did not result in any difference in proteolytic products when compared to the individual proteins (figure 6, D). Comparable results were obtained with V8 protease (data not shown).

Generation of crystal constructs

To gain detailed structural information of the Nic96/Nup192 interface and its connection to the Nsp1 complex, we attempted to crystallize various assemblies with and without the Nsp1 complex and the N terminus of Nup53. All crystal constructs are summarized in (figure 7). The purification procedures were identical to those presented in the previous sections.

In addition to full length Nup192 and Nic96, rational construct design was implemented to mirror previously crystallized fragments and to be consistent with biochemical data. The flexible tail module of the Nic96 ACE1 domain was removed as well as its unstructured N terminus. Nup192 was truncated to the C-terminal residues 1379-1734, consistent with the previously solved Nup188 C terminus (Andersen et al., 2013). Point mutations were made in the N-terminal helix to substitute two hydrophobic residues for serines to prevent aggregation (Stuwe et al., 2014). Nsp1, Nup49, and Nup57 were limited to the structured C-terminal coiled coil domains (Stuwe et al., 2015). Nup53 was limited to its C terminus.

The largest assembly attempted was Nic96/Nup192/Nsp1/Nup49/Nup57. Extensive crystallization trials yielded phase separation but no crystals. Nic96(325-1193)/Nup192 (full length) did yield quasi crystals, but they were too fragile to harvest and therefore could not be exposed to X-rays. Attempts were made to improve the stability of the crystals by fine screening around the original crystallization condition and micro seeding, but similar pathologies remained. It was not clear if any further modifications to the constructs would be beneficial, so instead, we looked to modify the properties of the assemblies with nanobodies.

Generation of nanobodies

Nanobodies are small single domain antibodies obtained from the variable region of Camelid immunoglobulins (Fridy, 2014). Despite their small size, nanobodies can have high affinity for its target protein and can be produced in large quantities (Fridy, 2014). Therefore, they are an attractive alternative to full immunoglobulins in assisting with crystallization. Nanobodies have been successfully applied as crystallization chaperones in our lab and others (Demircioglu, 2016; Pardon, 2014).

To raise antibodies for crystallization trials, mtNup192 (full length) and mtNic96 (210-1193) were injected into alpaca. A plasmid library containing nanobody sequences was generated and enriched for antigen binders by a phage display protocol. The antigens that were challenged against the nanobody library were Nup192 (1379-1734), Nic96 (210-1193), and Nup192 (1379-1734)/Nic96 (210-1193). The C terminus of Nup192 was challenged to generate nanobodies that only bind to its C terminus. Candidate nanobodies were cloned into expression vectors and expressed heterologously in *E. coli*. 5, 6, and 9 sequence unique nanobodies were generated for Nup192 (1379-1734), Nic96 (210-1193), and Nup192 (1379-1734)/Nic96 (210-1193), respectively. To verify binding nanobodies, binding was first tested with the pentameric Nic96(210-1193)/Nup192(1379-1734)/Nsp1(501-718)/Nup49(242-470)/Nup57(70-329) assembly by size exclusion chromatography. An example of a successfully binding nanobody is shown in figure 8. Nanobodies that fail to bind to the pentameric assembly were excluded from further analysis. During this analysis, it was observed that of the 5 nanobodies that were generated for Nup192(1379-1734), 2 did not bind the pentameric assembly. The other 3 displayed pathological behavior. Incubation with these nanobodies dissociates Nup192 from the rest of the assembly, likely by disrupting interactions between Nup192 and Nic96 (figure 9). Nanobodies panned against Nup192 (1379-1734)/Nic96 (210-1193) were incubated with either Nup192 or Nic96 and analyzed by size exclusion chromatography to narrow down which of the two they bound to. Furthermore, non-redundantly binding nanobodies were identified by competitive binding assays. In summary, 8 sequence unique and non-redundantly binding nanobodies specific to Nic96 were generated for use in crystallization trials.

Discussion

To gain insight into the assembly structure of the Nic96 complex, we focused on the Nic96/Nup192 binding interface. Single particle EM analysis of negatively stained samples and limited proteolysis experiments suggests that the Nic96/Nup192 interface involves at least one more contact between the C terminus of Nup192 and the ACE1 domain of Nic96 in addition to the known interaction between Nup192 and the R2 helix of Nic96. Although the conformational heterogeneity observed in 2-D class averages suggests that the interaction is flexible, the ACE1 domain of Nic96 is adjacent to the C terminus of Nup192 despite the ~80 residues that separate it from the R2 helix, contrary to previously proposed models (compare figure 1 to figure 5C; Amlacher et al., 2011; Lin et al., 2016).

In order to gain high resolution structural details of the interface, we subject various assemblies to crystallization trials. In addition to full length constructs, truncation constructs were designed based on existing crystal structures of individual members. The Nsp1 complex was included since it is proximal to the Nic96/Nup192 binding interface. Nup53 was added as a potentially stabilizing agent since it binds to both Nic96 and Nup192. Our initial crystallization trials yielded promising leads, but no diffracting crystals. To modify the properties of the assemblies and promote crystallization, nanobodies were raised against Nic96 and Nup192. 8 unique nanobodies specific for Nic96 were generated. Moving forward, another attempt should be made at generating Nup192 nanobodies in order to get better coverage, as well as raising nanobodies to the Nsp1 complex.

It is exciting to see if and how Nup170 integrates with Nic96 and Nup192. Nup53 binds to all three proteins, but the binding sites are separated by a ~250 residue long flexible linker (figure 5; Amlacher et al., 2011; Lin et al., 2016). A recent report using nucleoporins from *C. thermophila* demonstrated that another flexible protein, Nup145N, binds to both Nup192 and Nup170 with a much shorter flexible linker of ~50 residues, placing Nup170 relatively close to Nup192 (Fischer et al., 2015). EM analysis of a pentameric Nic96/Nup192/Nup170/Nup53/Nup145N assembly however indicated that Nup170 was completely dissociated from Nic96/Nup192. The low concentrations necessary to avoid overcrowding is likely the issue. Alternative strategies to promote Nup170 occupancy with specific or non-specific cross linkers may elucidate the location of Nup170 within the Nic96 complex, and by extension, the assembly structure of the entire Nic96 complex.

References

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). Determining the architectures of macromolecular assemblies. Nature 450, 683–694.

Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D.P., Arumugam, M., Bork, P., and Hurt, E. (2011). Insight into Structure and Assembly of the Nuclear Pore Complex by Utilizing the Genome of a Eukaryotic Thermophile. Cell 146, 277.

Andersen, K.R., Onischenko, E., Tang, J.H., and Kumar, P. (2013). Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. Elife 2, e00745.

Demircioglu, F.E., Sosa, B.A., Ingram, J., Ploegh, H.L., Schwartz, T.U. (2016) Structures of TorsinA and its disease-mutant complexed with an activator reveal the molecular basis for primary dystonia. Elife 5, e17983.

Fischer, J., Teimer, R., Amlacher, S., Kunze, R., Hurt, E. (2015) Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. NSMB 22, 774–781.

Fridy, P.C., Li, Y., Keegan, S., Thompson, M.K., Nudelman, I., Scheid, J.F., Oeffinger, M., Nussenzweig, M.C., Fenyö, D., Chait, B.T., Rout, M.P. (2014). A robust pipeline for rapid production of versatile nanobody repertoires. Nature Methods 11, 1253–1260.

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.

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, 114-24.

Jeudy, S., and Schwartz, T.U. (2007). Crystal structure of nucleoporin Nic96 reveals a novel, intricate helical domain architecture. Journal Of Biological Chemistry 282, 34904-12-34912.

Koth, C.M., Orlicky, S.M., Larson, S.M., Edwards, A.M. (2003) Use of Limited Proteolysis to Identify Protein Domains Suitable for Structural Analysis. Methods in Enzymology 368, 77-84.

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

Li, L., Fierer, J.O., Rapoport, T.A., Howarth, M. (2014) Structural analysis and optimization of the covalent association between SpyCatcher and a peptide Tag. J. Mol. Biol. 425, 309-217.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., Davenport, A.M., Jeon, Y.E., Hoelz, A. (2016) Architecture of the symmetric core of the nuclear pore. Science 353, 308.

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

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. The Journal Of Cell Biology 185, 475-91-491.

Pardon, E., Laeremans, T., Triest, S., Rasmussen, S.G.F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W.G.J., Kobilka, B.K., Steyaert, J. (2014). A general protocol for the generation of Nanobodies for structural biology. Nature Protocols 9, 674–693.

Schlaich, N.L., Haner, M., Lustig, A., Aebi, U., and Hurt, E.C. (1997). In vitro reconstitution of a heterotrimeric nucleoporin complex consisting of recombinant Nsp1p, Nup49p, and Nup57p. Mol. Biol. Cell *8*, 33–46.

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

Seo, H., Blus, B.J., Jankovic, N.Z., and Blobel, G. (2013). Structure and nucleic acid binding activity of the nucleoporin Nup157. Proceedings Of The National Academy Of Sciences 110, 16450-5-16455.

Seo, H., Ma, Y., Debler, E.W., Wacker, D., Kutik, S., Blobel, G., and Hoelz, A. (2009). Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. Proceedings Of The National Academy Of Sciences 106, 14281-6-14286.

Stuwe, T., Lin, D.H., Collins, L.N., Hurt, E., and Hoelz, A. (2014). Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. Proceedings Of The National Academy Of Sciences 111, 2530.

Stuwe, T., Bley, C.J. , Thierbach, K., Petrovic, S., Schilbach, S., Mayo, D.J., Perriches, T., Rundlet, E.J., Jeon, Y.J., Collins, L.N., Huber, F.M., Lin, D.H., Paduch, M., Koide, A., Lu, V., Fischer, J., Hurt, E., Koide, S., Kossiakoff, A., Hoelz, A. (2015). Architecture of the fungal nuclear pore inner ring complex. Science, 350, 56-64

Theile, C.S., Witte, M.D., Blom, A.E.M., Kundrat, L., Ploegh, H.L., and Guimaraes, C.P. (2013). Site-specific N-terminal labeling of proteins using sortase-mediated reactions. Nature Protocols 8, 1800-1807.

von Appen, A., Kosinski, J., Sparks L., Ori, A., DiGuilio, A. L., Vollmer, B., Mackmull, M.T., Banterle, N., Parc, L., Kastritis, P., Buczak, K., Mosalaganti, S., Hagen, W., Andres-Pons, A., Lemke, E. A., Bork, P., Antonin, W., Glavy, J. S., Bui, K. H., and Beck, M. (2015). In *situ* structural analysis of the human nuclear pore complex. Nature 526, 140-143.

Whittle, J.R.R., and Schwartz, T.U. (2009). Architectural Nucleoporins Nup157/170 and Nup133 Are Structurally Related and Descend from a Second Ancestral Element. Journal Of Biological Chemistry 284, 28442.

Wolfe, K.H., and Shields, D.C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387, 708-13.

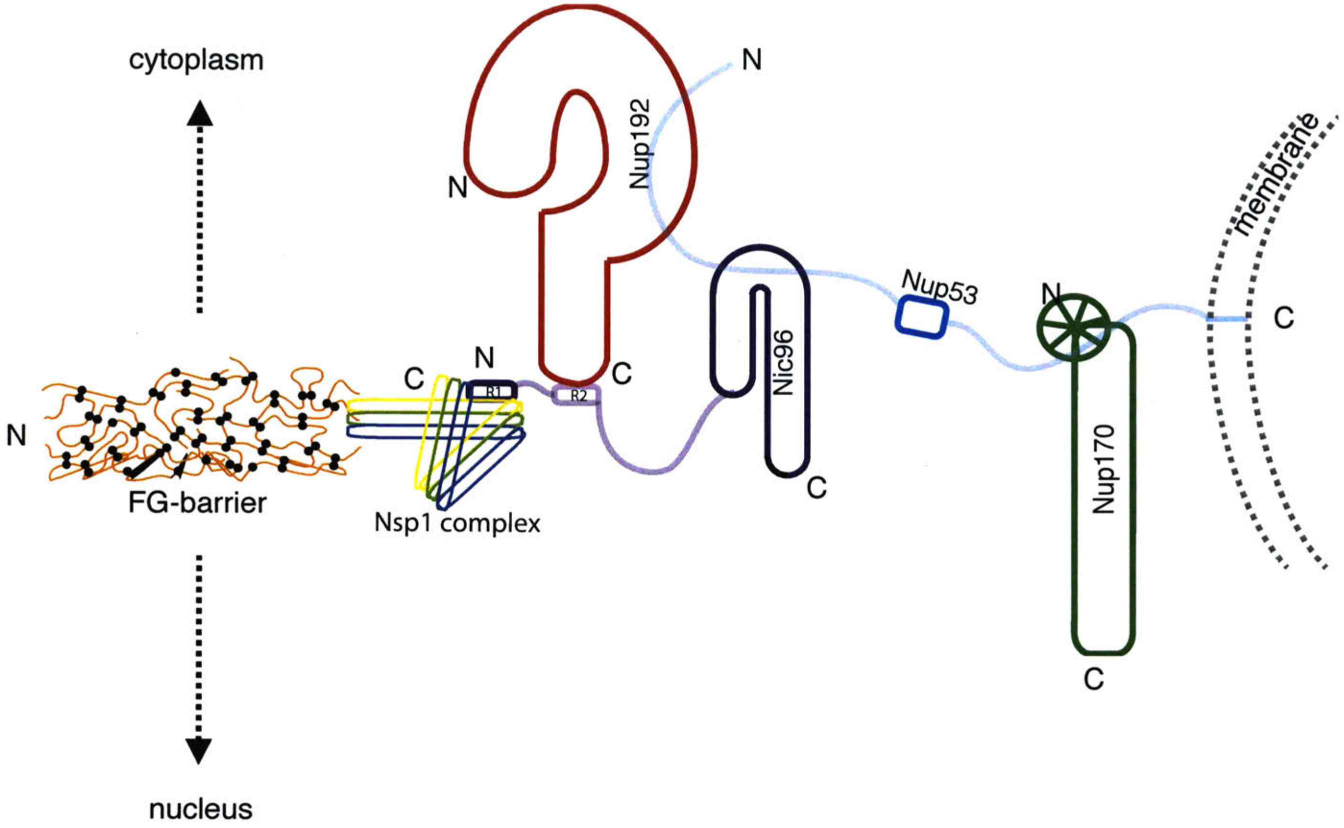


Figure 1 Cartoon representation of the Nic96 complex. Regions for which crystal structures are available are colored. Yet unsolved regions are transparent. Nup192/ Nup188 has a characteristic S-shape. ACE1 protein Nic96 binds to the C terminus of Nup192/Nup188 through its N-terminal R2 helix and to the Nsp1 complex through its R1 helix. Nup53 binds both Nup192/Nup188 and Nic96 through its N-terminal unstructured region and to Nup170 through its C-terminal unstructured region. Nup53 has a C-terminal amphipathic helix that inserts into the membrane. Nup170 has a N-terminal β-propeller followed by an α-helical stack which shares structural homology with Y-complex member Nup133. The Nsp1 complex docks onto Nic96 and forms the FG-repeat rich selective barrier at the center of the NPC.

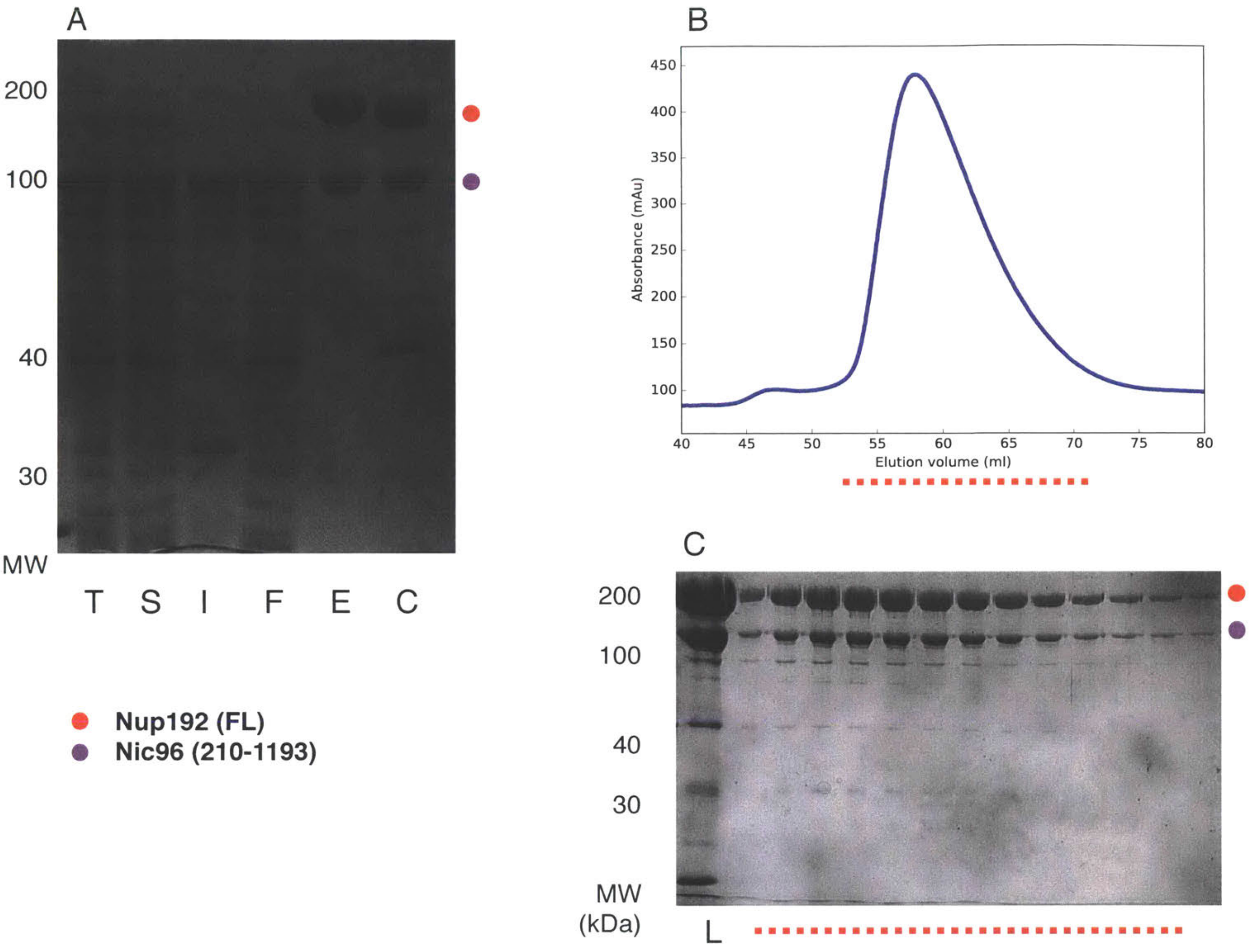


Figure 2. Purification strategy for the dimeric Nic96/Nup192 assembly. (A) Ni affinity purification visualized by Coomassie stained SDS-PAGE gel. T: total lysate, S: soluble, I: insoluble, F: flow through, E: elution, C: elution + protease. **(B)** Size exclusion chromatography 280 nm trace. **(C)** L: load. Size exclusion chromatography fractions underlined in dashed red line, visualized by Coomassie stained SDS-PAGE gel.

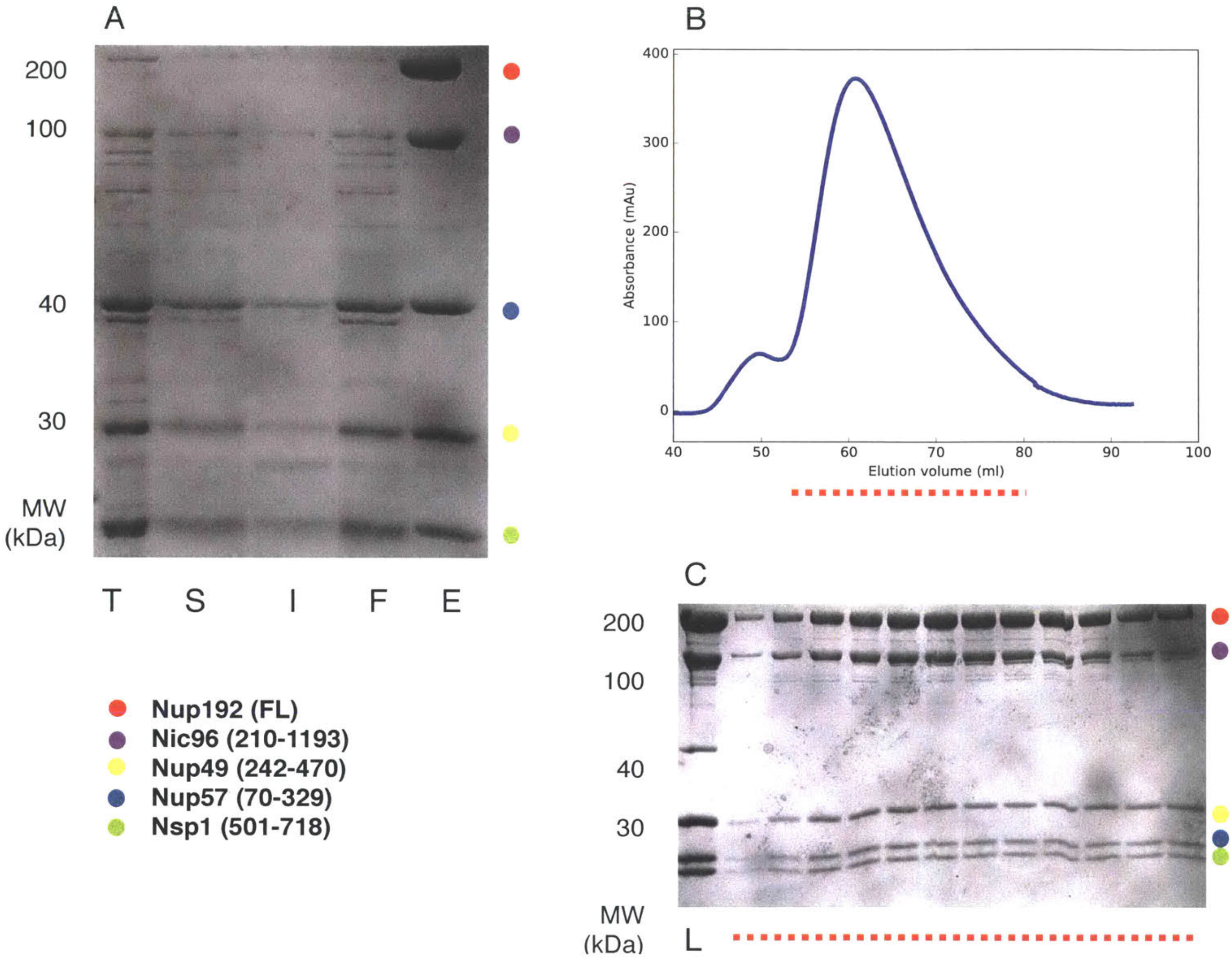


Figure 3. Purification of the pentameric Nic96/Nup192/Nsp1/Nup49/Nup57 assembly. (A) Ni affinity purification visualized by Coomassie stained SDS-PAGE gel. T: total lysate, S: soluble, I: insoluble, F: flow through, E: elution, C: elution + protease. **(B)** Size exclusion chromatography 280 nm trace. **(C)** L: load. Size exclusion chromatography fractions underlined in dashed red line, visualized by Coomassie stained SDS-PAGE gel.

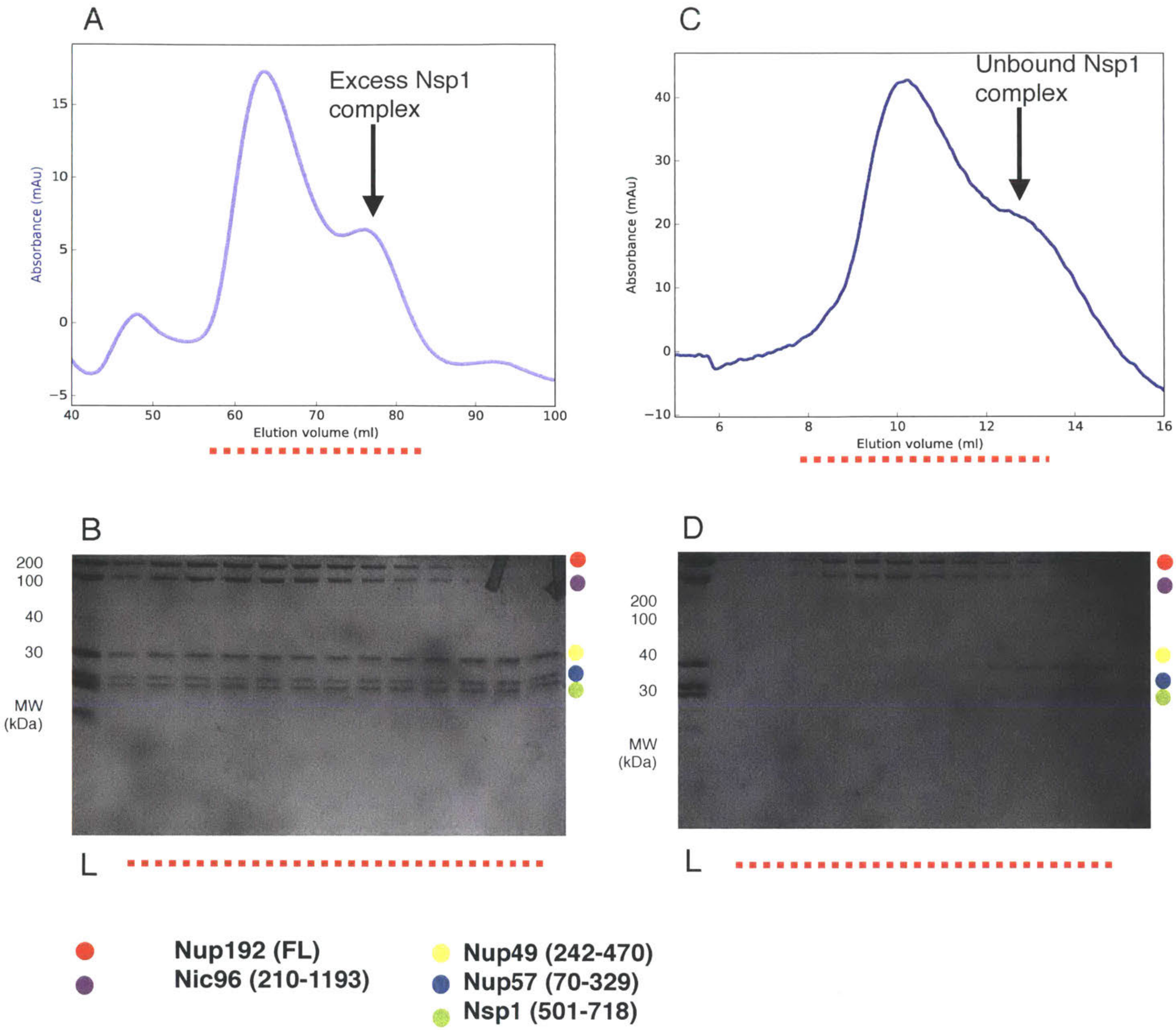


Figure 4. Behavior of the Nsp1 complex during Ni affinity purification and in reconstitution assays. (A) Size exclusion chromatography 280 nm trace of Nic96/Nup192/Nsp1/Nup49/Nup57 purified without 10% glycerol. An excess of Nsp1 complex can be seen in a peak adjacent to the pentameric assembly peak. **(B)** L: load. Size exclusion chromatography fractions of **(A)** underlined in dashed red line, visualized by Coomassie stained SDS-PAGE gel. **(C)** Size exclusion chromatography 280 nm trace of Nic96/Nup192 pre-incubated with Nsp1/Nup49/Nup57. The Nsp1 complex does not co-elute with Nic96/Nup192, indicating a failure to form a pentameric complex. **(D)** L: load. Size exclusion chromatography fractions of **(C)** underlined in dashed red line, visualized by Coomassie stained SDS-PAGE gel.

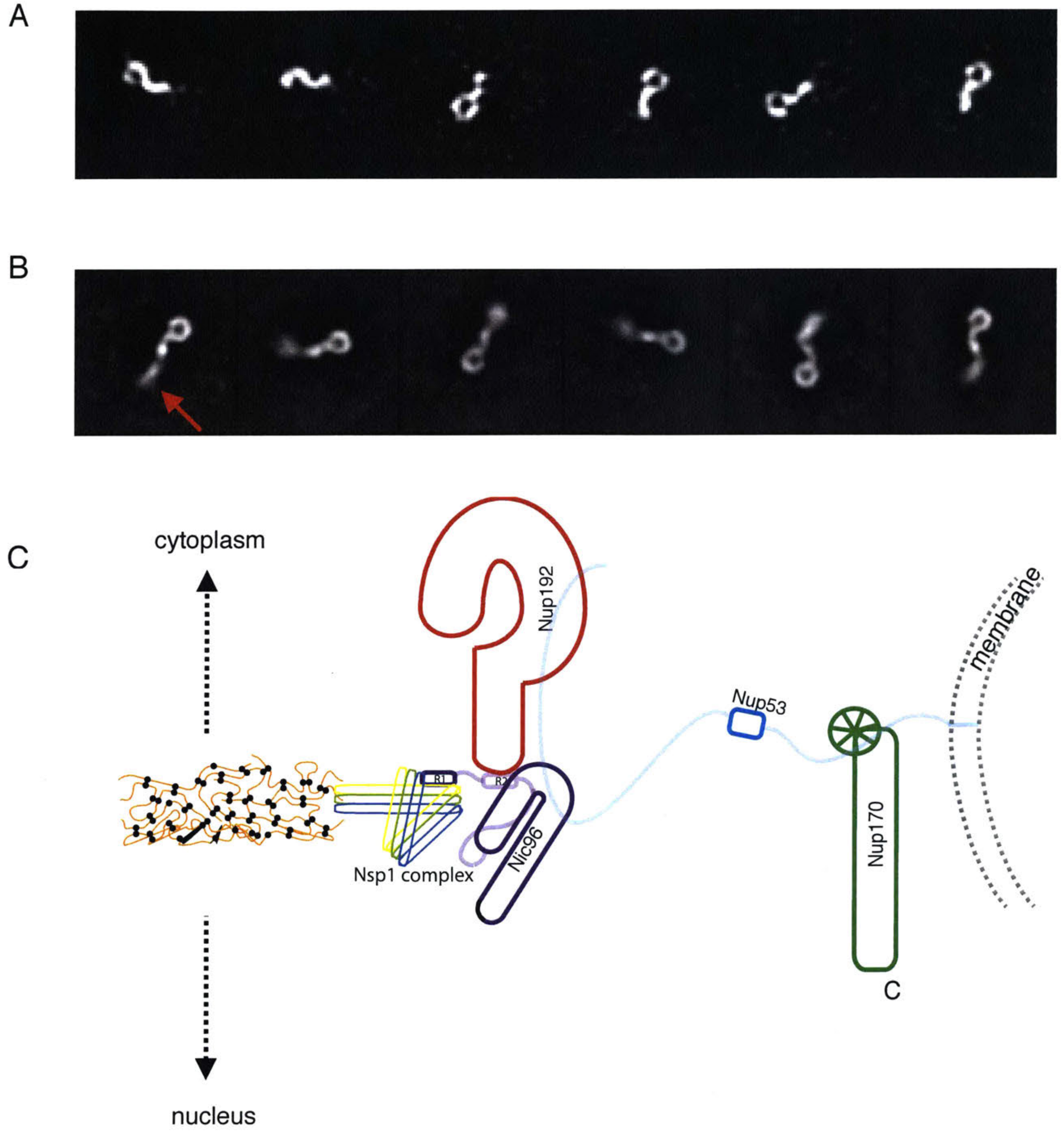


Figure 5. Single particle EM analysis of the dimeric Nic96/Nup192 assembly. 2-D class averages of negatively stained Nic96/Nup192. (A) 2-D class averages of Nup192 alone. **(B)** 2-D class averages of intact Nic96/Nup192. Extra density is observed at the C terminus of Nup192 that is consistent with the dimensions of Nic96 (red arrow). **(C)** Cartoon representation of the Nic96 complex similar to figure 1, but with Nic96 positioned at the C terminus of Nup192 to reflect its position in the class averages displayed in **(B)**.

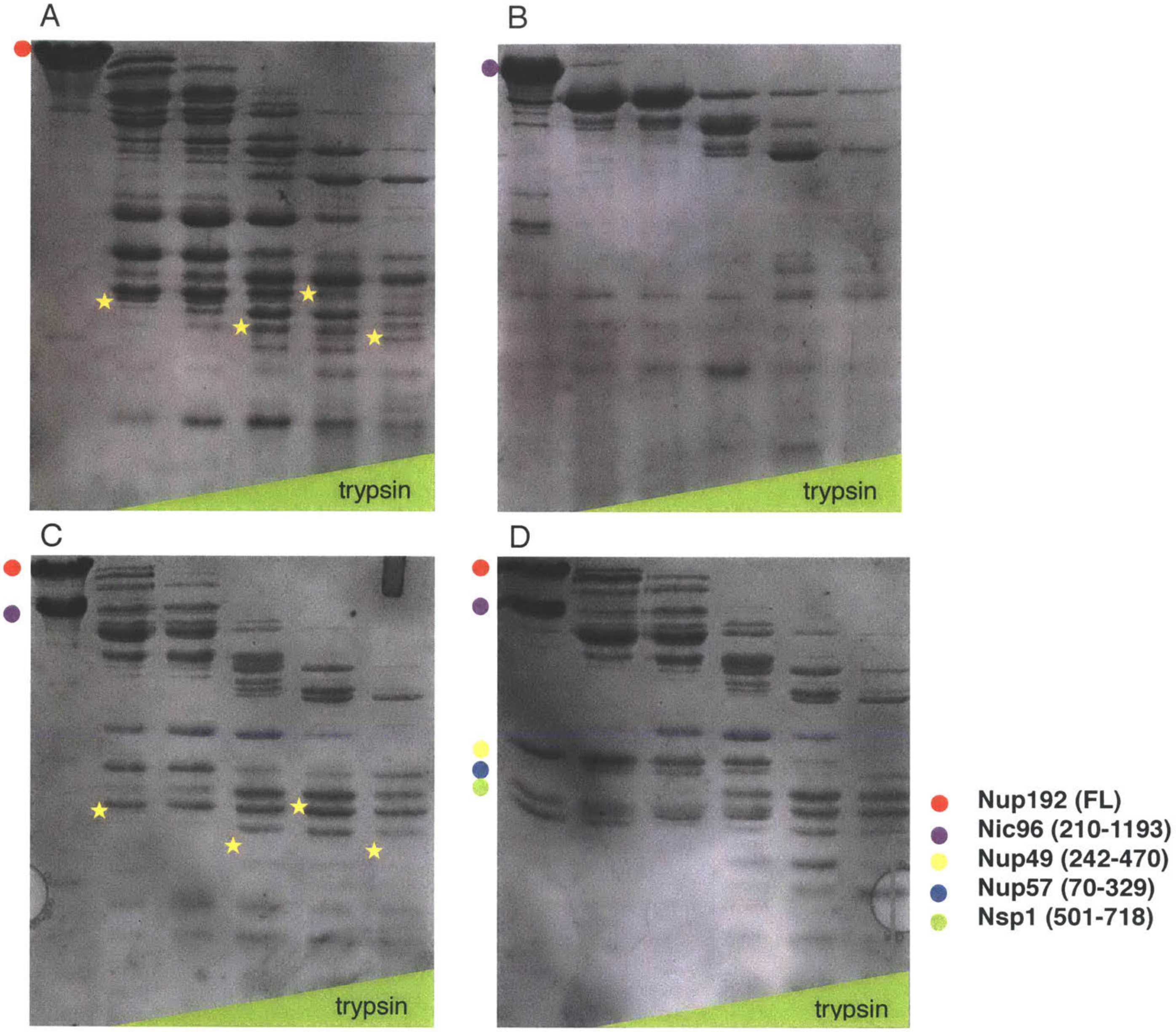


Figure 6. Limited proteolysis. Each sample was incubated with an increasing amount of trypsin and visualized by Coomassie stained SDS-PAGE gel. Stars denote bands that are different between the dimeric Nic96/Nup192 and Nup192 alone that are not attributable to Nic96 alone, indicating mild proteolytic protection of Nup192 by Nic96. **(A)** Nup192. **(B)** Nic96. **(C)** Nup192/Nic96. **(D)** Nup192/Nic96/Nsp1/Nup49/Nup57.

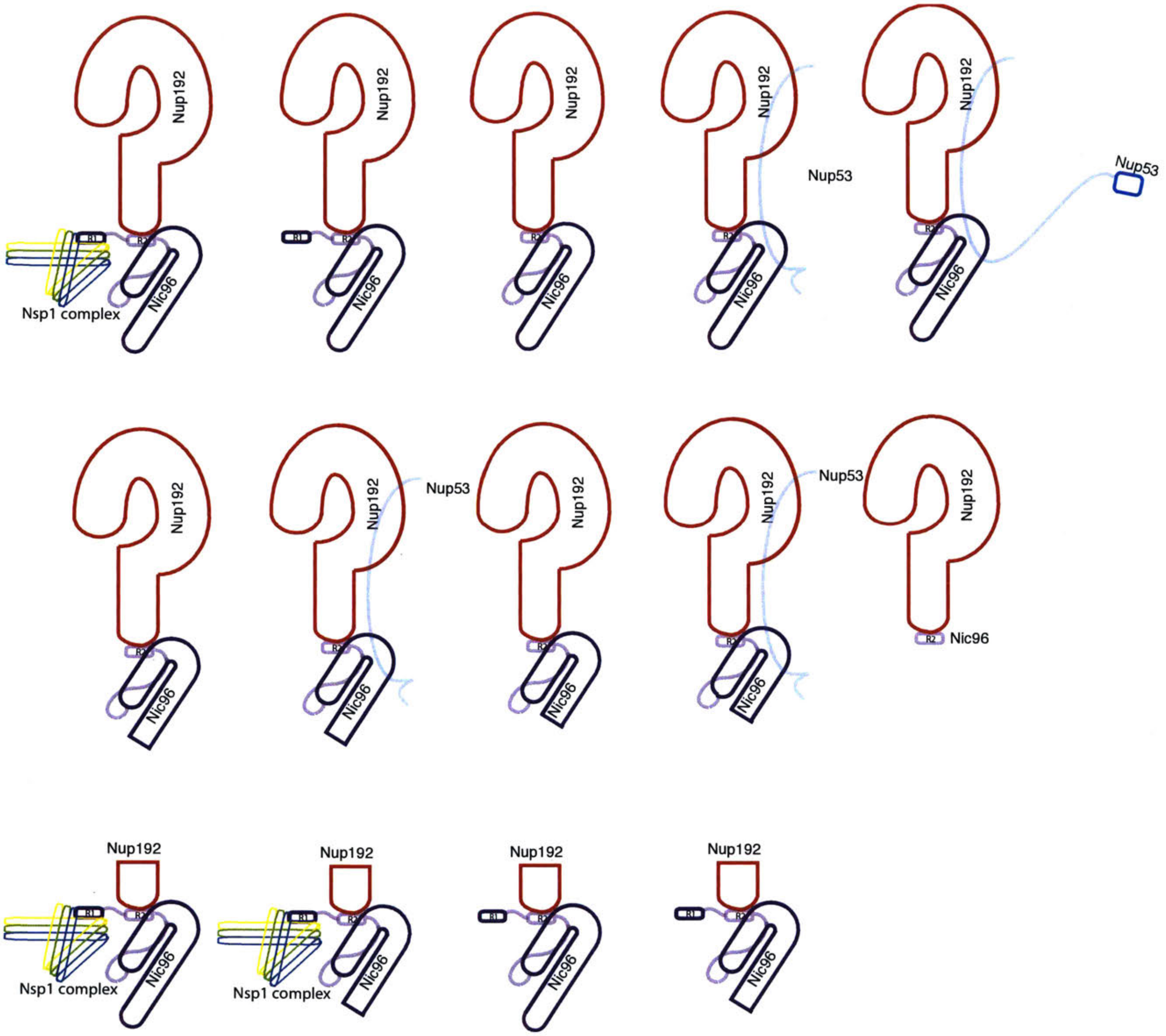


Figure 7. Cartoon representation of crystal constructs tried.

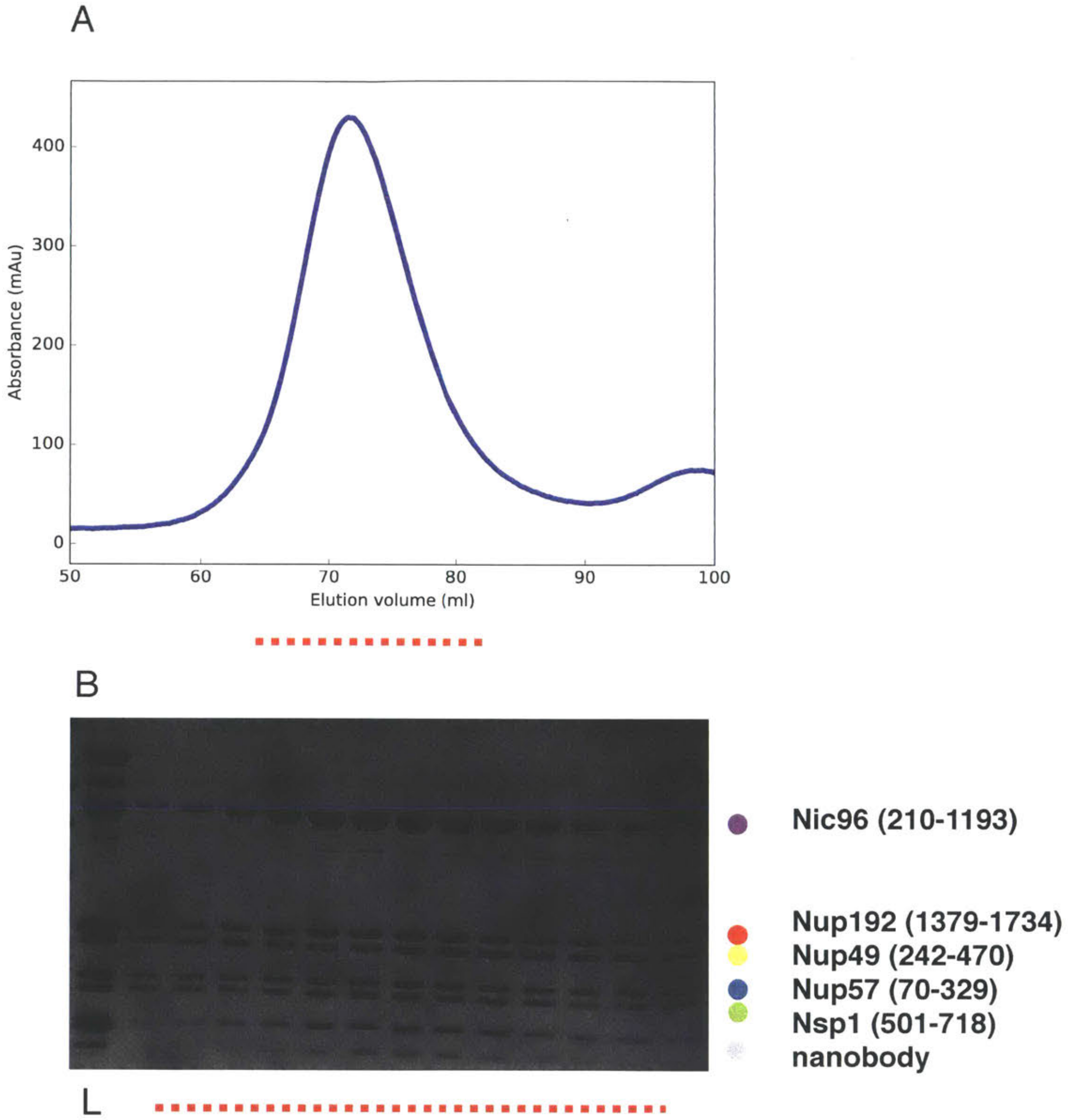


Figure 8. An example of a nanobody that binds to the pentameric Nic96/Nup192/Nsp1/Nup49/Nup57 assembly. Nic96(210-1193)/Nup192(1379-1734)/Nsp1(501-718)/Nup49(242-470) was incubated with nanobody and subject to size exclusion chromatography. **(A)** Size exclusion chromatography 280 nm trace. **(B)** L: load. Size exclusion chromatography fractions underlined in dashed red line, visualized by Coomassie stained SDS-PAGE gel. This particular nanobody clearly binds and co-elutes with the pentameric assembly.

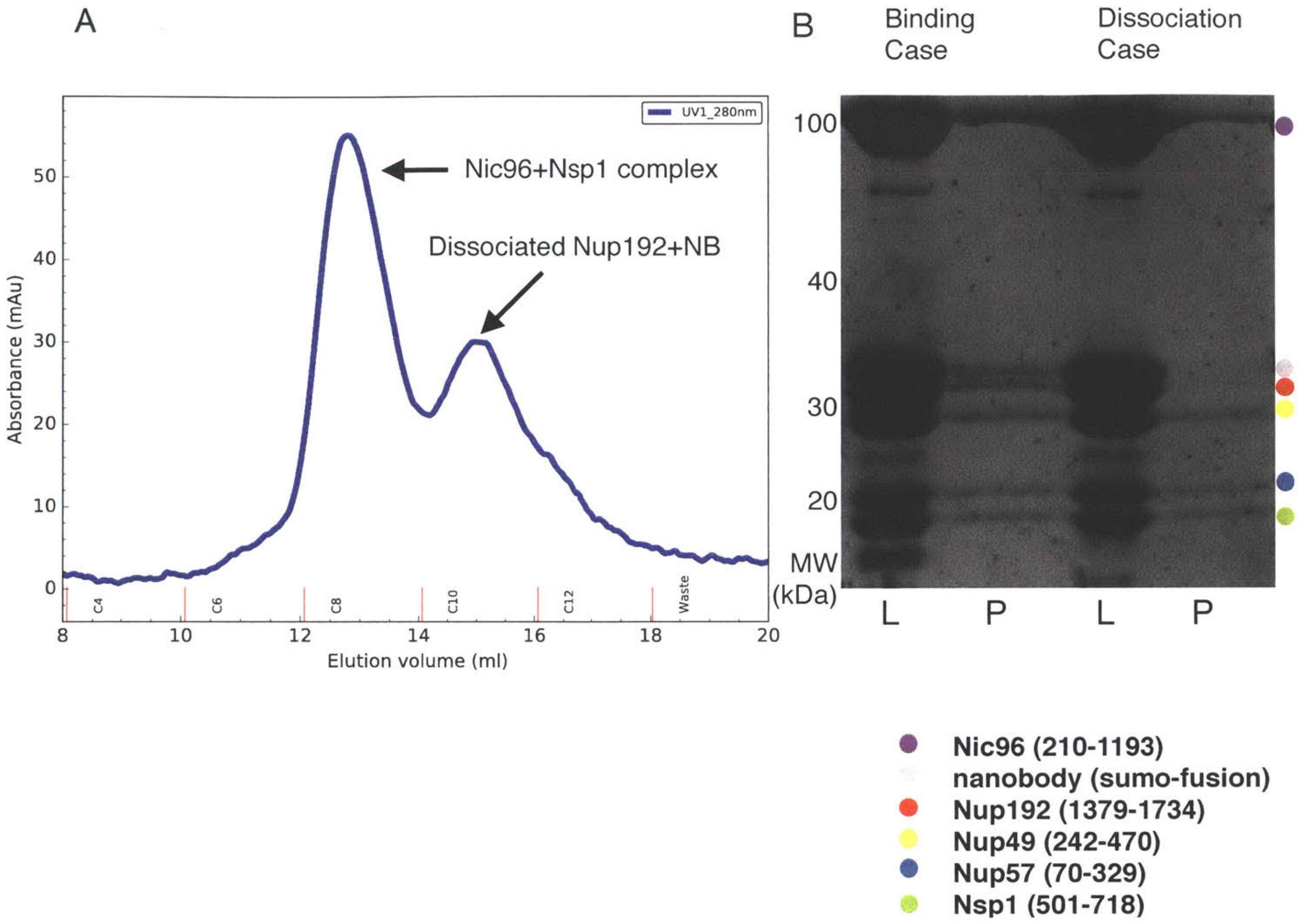


Figure 9. An example of a nanobody that dissociates Nup192. Nic96(210-1193)/ Nup192(1379-1734)/Nsp1(501-718)/Nup49(242-470) was incubated with a nanobody and subjected to size exclusion chromatography. **(A)** Size exclusion chromatography 280 nm trace. Nanobody binds to Nup192 competitively, dissociating it from the rest of the complex. **(B)** L: load. P: peak. Left, size exclusion chromatography peak fraction of a successfully binding nanobody. Right, size exclusion chromatography peak fraction of a Nup192 dissociating nanobody from **(A)**, visualized by Coomassie stained SDS-PAGE gel.

Chapter 4: Conclusions

Summary

The work presented here has let to a 4.1 Å structure of the Y complex hub from *Myceliophthora thermopila.* The structure revealed that the junction between the three arms of the Y complex forms a highly curved, pyramidal shape, causing it to deviate from the previously assumed flat architecture. We believe this discrepancy comes from the artificial flattening of the hub region in previously reported EM reconstructions of the Y complex. The structure also revealed that Seh1 is not a required member of the Y complex, leading to the conclusion that the minimal Y complex is a hexamer instead of a heptamer. We combined our hub structure with previously solved X-ray structures to construct a complete, composite high resolution structure of the Y complex. We found that the Y complex structure is similar across eukaryotes; built around a conserved hexametric core, with species specific additions that decorate, but do not significantly alter its overall shape. Docking the composite model into previously solved random conical tilt(RCT) and tomographic reconstructions of negatively stained samples of the Y complex showed overall consistency between the three methods, yet we highlight important structural differences, especially at the hub region, that constrains where the Y complex is placed within the NPC. By docking the composite model into the 32 Å cryo-ET reconstruction of the entire NPC (Bui et al., 2013) with consideration of both steric requirements and possible flexibility, we propose an arrangement of a single head-to-tail Y complex ring on both nuclear and cytoplasmic sides. Even with the significant improvements that our composite model provides over previously solved low resolution structures of the Y complex, we note that many reasonable solutions result from the docking procedure. Therefore, the inclusion of additional experimental evidence is necessary in order to definitively rule out or confirm a particular model.

This work has also lead to preliminary results on the assembly structure of the Nic96 complex. Single particle EM analysis of negatively stained samples suggests that the Nic96/Nup192 interface involves at least one more contact between the C terminus of Nup192 and the ACE1 domain of Nic96, in addition to the known interaction between Nup192 and the R2 helix of Nic96. Although the conformational heterogeneity observed in the 2-D class averages suggests that the interaction is flexible, the ACE1 domain of Nic96 is adjacent to the C terminus of Nup192, despite the ~80 residues that separate it from the R2 helix. To obtain high resolution information of the interface, we subject various assemblies to crystallization trials. Our initial crystallization trials yielded promising leads and serves as a basis for future experiments. Finally, to promote crystallization, nanobodies were raised against Nic96 and Nup192.

Future directions

Y complex

The composite structure of the Y complex determined in this work is a significant advance over its low resolution predecessors. However, important structural questions remain for how exactly species specific additions integrate into the minimal hexametric assembly. Although the X-ray crystal structure of the *S. pombe* Nup37/Nup120 assembly has been solved (Bilokapic and Schwartz, 2012), we are still missing structures for the Nup43/Nup85 and ELYS/Nup120 interface. These structures will result in more accurate species specific composite structures that may further constrain possible higher order assemblies of the Y complex.

In order to arrive at a definitive higher order Y complex assembly structure, we should better map inter sub-complex interactions and if possible, obtain structural information of their interfaces. This will also narrow down possible solutions from docking experiments into the cryo-ET maps of the entire NPC. For example, in a recent report it was show that the Nup82 complex interacts with the Y complex through contacts with Nup85 (Fernandez-Martinez et al., 2017). A co-crystal structure would reveal the accurate orientation of the Nup82 complex relative to the Y complex and is expected to significantly alter the shape of the Y complex core structure since its structured region is relatively large and elongated. The only reported physical connection between the Y complex and the Nic96 complex is between Nup120/Nup170 or mediated by Nup145N (Lutzmann, et al., 2005; Lin et al., 2016). The Nup120/Nup170 interaction was demonstrated by pull down assays. Nup145N is a mostly unstructured FG-repeat rich protein that links Nup145C of the Y complex to Nup192, Nup170, and the Nup82 complex (Fischer et al., 2015). Although the interactions are assumed to be separated by unstructured linkers, Nup145N is potentially an important connector that has a distinct structural role. It would be exciting to see how the inclusion of Nup145N affects the binding interfaces between the Y, Nic96, and Nup82 complexes. It is also possible that other inter sub-complex contacts have evaded detection so far due to dynamic or transient interactions. Methods that measure both on and off rates of binary binding events such as surface plasmon resonance or bio-layer interferometry would would be useful in such cases.

Nic96 complex

Better characterization of the Nic96 complex assembly structure should involve improving the resolution of the Nic96/Nup192 interface structure as well as defining whether or not it assumes a particular orientation relative to Nup170. Towards the first goal, another round of panning should performed to raise nanobodies that are specific to the C terminus of Nup192 and the Nsp1 complex. We expect these nanobodies, in conjunction with the ones already raised towards Nic96, to improve the odds of crystallization of both the dimeric Nic96/Nup192 and pentameric Nic96/Nup192/Nsp1/ Nup49/Nup57 assemblies. Defining the relative position of Nup170 within the Nic96 complex is challenging due to the instability of the Nup53/Nup170 interface. Consequently, Nup170 tends to dissociate from the rest of the Nic96 complex at concentrations appropriate for EM studies. This can potentially be mitigated by mild cross-linking agents that target either surface lysines or cysteins. This chemical stabilization should at the very least allow for intact assemblies that can be analyzed by EM. 2-D classification of cross-linked particles with and without Nup170 may reveal the position of Nup170, or alternately, may reveal that Nup170 does not adopt a defined orientation relative to the rest of the Nic96 complex.

References

Bilokapic, S., and Schwartz, T.U. (2012). Molecular basis for Nup37 and ELY5/ELYS recruitment to the nuclear pore complex. Proceedings Of The National Academy Of Sciences 109, 15241.

Bui, K.H., Appen, von, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., et al. (2013). Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. Cell 155, 1233.

Fernandez-Martinez, Z., Kim, S.J., Shi. Y., Upla, P., Pellarin, R., Gagnon, M., Chemmama, I.E., Wang, J., Nudelman, I., Zhang, W., Williams, R., Rice, W.J., Stokes, D.L., Zenklusen, D., Chait, B.T., Sali, A., Rout, M.P. (2016) Structure and Function of the Nuclear Pore Complex Cytoplasmic mRNA Export Platform. Cell 167, 1159-1160.

Fischer, J., Teimer, R., Amlacher, S., Kunze, R., Hurt, E. (2015) Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. NSMB 22, 774–781.

Lutzmann, M., Kunze, R., Stangl, K., Stelter, P., Tóth, K.F., Böttcher, B., and Hurt, E. (2005). Reconstitution of Nup157 and Nup145N into the Nup84 complex. Journal of Biological Chemistry 280, 18442-51.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., Davenport, A.M., Jeon, Y.E., Hoelz, A. (2016) Architecture of the symmetric core of the nuclear pore. Science 353, 308.