Structural and biochemical analysis of the Y-shaped Nup84 subcomplex of the NPC

by

Nina Carolina Leksa

B.S. Duke University
Durham, NC 2006

SUBMITTED TO THE DEPARTMENT OF BIOLOGY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTORATE OF PHILOSOPHY

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2012

© 2012 Massachusetts Institute of Technology
All rights reserved

Signature of Author: ...........................................

Nina C. Leksa
Department of Biology
January 6, 2012

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

Thomas U. Schwartz
Associate Professor of Biology
Thesis Supervisor

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

Robert T. Sauer
Luria Professor of Biology
Chair, Biology Graduate Committee
Structural and biochemical analysis of the Y-shaped Nup84 subcomplex of the NPC
by
Nina Carolina Leksa

Abstract

The eukaryotic cell is easily distinguishable from its prokaryotic counterpart through the presence of a complex endomembrane system. Most notable is the nucleus, which harbors and protects the genetic information of the cell. Such physical separation allows for extra layers of regulation for a variety of cellular processes, but at the same time, necessitates a mechanism for communication between the nucleus and cytoplasm. The nuclear envelope consists of an inner and outer nuclear membrane that fuse at distinct loci to form a gateway into the nucleus. Embedded into these circular openings are nuclear pore complexes (NPCs), which serve as the gatekeepers, mediating all exchange between the nucleus and the cytoplasm. At its core, the NPC consists of an 8-fold symmetric structural scaffold that serves as a docking site for some of the more dynamic components. Overall, the NPC is composed of ~30 proteins, or nucleoporins (nups), that form biochemically stable subcomplexes which are repeated in multiple copies to assemble the intact NPC. This macromolecular machine is not only essential to nucleo-cytoplasmic transport, but also plays a pivotal role in a myriad of other cellular processes. Thus, understanding its function and assembly in molecular detail is of great interest and importance.

Here, I investigate the structure and assembly of one of the major components of the NPC scaffold, the Nup84 complex. Also known as the Y complex, this 7-membered assembly adopts the shape of an extended Y in solution. We solved the structure of the heterodimeric Nup85 Seh1 complex, which forms one of the short arms of the Y. Nup85 was found to have an unexpected, yet conserved fold, termed ancestral coatomer element 1 (ACE1), which is also present in 3 additional scaffold nucleoporins, as well as Sec31, a major component of COPII vesicle coats. This discovery led to the first experimental evidence for a common ancestry of nucleoporins and vesicle coat proteins. Additionally, we solved the partial structure of Nup120, which also exhibited a unique and unexpected domain architecture. While initial secondary structure prediction methods classified all nups into the canonical α-solenoid and β-propeller fold types, an arsenal of recent nucleoporin structures now tells a different story.

To date, at least a partial structure of each of the components of the Nup84 complex has been solved. Using biochemical interaction data as a guide, we can now place each of these structures into an electron density map generated by electron microscopy (EM) to build an initial composite structure of a nearly complete Y-shaped Nup84 complex. Furthermore, we developed the lattice model for assembly of the NPC structural scaffold based on the similarities discovered between nucleoporins and COPII vesicle coats—two membrane-coating complexes whose proteins are evolved from a common ancestor. To develop the lattice model further, deciphering the complex interaction network linking each of the biochemically-defined subcomplexes will be paramount in arriving at a more detailed and accurate model that can provide mechanistic insight into NPC function.

Thesis Supervisor: Thomas U. Schwartz
Title: Associate Professor of Biology
Table of Contents

Abstract................................................................................................................ 2
Table of Contents.......................................................................................... 3
Figures and Tables ..................................................................................... 6
Acknowledgements ...................................................................................... 8

CHAPTER 1: Introduction............................................................................. 10

Introduction to the nuclear pore complex.................................................. 11
Regulated transport across the NE ......................................................... 13
Protein transport...................................................................................... 13
Transport of mRNA and ribosomal subunits ........................................... 14
Overall structure of the NPC.................................................................. 16
Protein composition and size.................................................................. 19
Modularity of the NPC.............................................................................. 21
Breaking down the NPC.......................................................................... 23
Protein folds within the NPC................................................................. 26
Assembly of the NPC................................................................................ 29
Similarity to COPII vesicle coats.............................................................. 32
Bridging the resolution gap........................................................................ 33
Solving the NPC structure from the inside out...................................... 34

Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats.............................................................. 36

Abstract...................................................................................................... 37
Introduction................................................................................................. 37
Results........................................................................................................ 39
Discussion................................................................................................... 58
Methods...................................................................................................... 67

Protein Production and Purification.......................................................... 67
Crystallization............................................................................................. 69
Structure Determination........................................................................... 70
CHAPTER 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture ......................................................... 76

Introduction ................................................................. 77

Results ............................................................................. 79

Structure Determination .................................................... 79
Crystal Structure of Nup120 ............................................... 80
The Main Crystal Contact is Formed by a Domain Swap ............. 82
Conservation of Nup120 and Comparison to the Human Ortholog Nup160 ................................................................. 83
The C terminus of Nup120 Directly Binds Nup145C and Nup85 ....... 85
Without its C-terminal domain Nup120 does not properly localize to the NPC ................................................................. 87
Nup120 is Topologically Different from other Scaffold Nucleoporins .................. 89

Discussion ........................................................................... 93

Nup120 in the Context of the NPC Scaffold .................................. 93

Experimental Procedures .................................................. 96

Protein Expression and Purification ....................................... 96
Protein Crystallization ....................................................... 97
Structure Determination .................................................... 98
Analytical Size Exclusion Chromatography ................................ 98
Yeast Strain Construction .................................................. 99
Fluorescence Microscopy ................................................... 99
ACCESSION NUMBER ....................................................... 100
ACKNOWLEDGEMENTS .................................................... 100

CHAPTER 4: Conclusion .......................................................... 101
Summary ............................................................................................................. 102
“Y” it needs more work ...................................................................................... 102
Building the Nic96 complex ............................................................................ 103
Models of NPC assembly .................................................................................. 104
Lattice model of the NPC and of COPII coats: similarities and differences .............................................................................................................. 106
β-Propellers and the Assembly of the NPC ..................................................... 110
Challenges .......................................................................................................... 112
Outlook ............................................................................................................. 114

References ........................................................................................................ 115
Figures and Tables

Figure 1. 1 - Protein import and export ........................................................... 14
Figure 1. 2 - mRNA export ............................................................................. 16
Figure 1. 3 - Overall structure of the nuclear pore complex ......................... 19
Figure 1. 4 - Inventory of nucleoporins .......................................................... 21
Figure 1. 5 - Schematic representation of modular NPC subcomplexes ......... 22
Figure 1. 6 - Assembly of the Y-shaped Nup84 subcomplex ....................... 24
Figure 1. 7 - Models of NPC assembly ............................................................ 31

Figure 2. 1 - Structure of the Nup85-Seh1 complex ...................................... 41
Figure 2. 2 - Arrangement of two Nup85-Seh1 heterodimers in the asymmetric unit ............................................................ 42
Figure 2. 3 - Nup85-Seh1 is a dimer in solution as determined by analytical ultracentrifugation ............................................................ 43
Figure 2. 4 - Nup85-Seh1 is a dimer in solution as determined by size exclusion chromatography ............................................................ 44
Figure 2. 5 - Comparison of Nup85-Seh1 and Nup145C-Sec13 and identification of the Nup84-Nup145C crown-crown binding interface .......... 46
Figure 2. 6 - Nup120 binds Nup145C and Nup85 via their tail modules ......... 47
Figure 2. 7 - A single point mutation in the predicted interaction helix of the Nup145C tail module disrupts Nup120 binding .............................. 48
Figure 2. 8 - Surface mutation of the Nup145C and Nup84 crowns does not negatively affect protein stability ............................................................ 49
Figure 2. 9 - Elimination of the Nup84 binding site on Nup145C results in nuclear pore assembly defects in vivo ............................................................ 51
Figure 2. 10 - Growth analysis of yeast strains .............................................. 52
Figure 2. 11 - Nup133-GFP and Nup84-GFP become more soluble in the Nup145-ELIEA strain ............................................................ 53
Figure 2. 12 - Architecture of ACE1 ............................................................... 55
Figure 2. 13 - Superposition of ACE1 modules ............................................. 56
Figure 2. 14 - Surface mutation of the Nup145C and Nup84 crowns ............. 57
Figure 2. 15 - Sequence alignments ............................................................... 64
Figure 2. 16 - Lattice model for the Nup84 complex and the structural scaffold of the nuclear pore complex ............................................................ 65
Acknowledgements

When I arrived at MIT over 5 years ago, I had no idea what was in store for me. I stepped foot on campus with little more than a strong curiosity and interest in science. What I lacked in experience I made up for with naïve optimism and excitement to simply be in the midst of such a stimulating and exciting scientific environment. It may be good that I did not quite realize what it would take to make it through graduate school at that time, but looking back, I realize it would have been impossible without the help of a large group of people.

First and foremost, I would like to thank my advisor Thomas Schwartz, for welcoming me into the lab and being willing to take a chance on me. I have been lucky to have such a driven advisor, whose enthusiasm is truly infectious. His door is always open and he always has time to offer advice and encouragement. Thomas has done a great job of building a stimulating, fun, friendly, exciting, efficient, helpful, and unique lab experience that I think will be hard to beat. I could not be happier or more grateful to have been a part of this lab and to have worked for Thomas over the past years.

Thank you also to the members of my thesis committee, Bob Sauer and Cathy Drennan. Both of you have been great over the years, providing constructive criticisms, advice, and suggestions that have helped me throughout my graduate career. Thank you also to Iain Cheeseman and Vlad Denic for joining my thesis committee. All of you have set an example of what it means to be a great scientist.

I would also like to thank all members of the Schwartz lab, past and present. In particular, I would like to thank Steve Brohawn, who started me off on the right foot when I joined the lab. Much of the work presented in this thesis was done in close collaboration with Steve, who not only mentored me and showed me how to be a successful scientist, but also how to balance work with fun. In addition, I would like to thank my baymates throughout the years—Silvija Bilokapic, James Partridge, and Kasper Andersen—for listening to me complain about anything and everything, from failed experiments to the cold Boston weather to staying up too late the night before for any number of reasons. Thanks also to Brian and Kevin for joining the lab even after having worked with me during their rotations, and to everyone else for helpful discussions over the years.

To my classmates, thank you for being great. You are all extremely talented individuals and scientists with whom I hope to stay in touch throughout my scientific career. I have a learned a lot from you all, most importantly perhaps, I learned how to take a break and relax. Our class trips and get-togethers have been a great outlet for fun, and I am thankful that we shared those experiences. To Heather, thanks for helping me feed my urge for planning things. And for taking part in all of the awesome, yet slightly embarrassing, interests we share.

Lastly, I would like to thank my family. Thanks to my parents for giving me the opportunity to come to MIT, and to my sister for randomly long email chains that provide enticing distractions from work sometimes. Thank you also to Matt, who has had to listen to more about my successes and failures, both in life and in science, than just about anyone.

Thank you.
To Opi and Tille
CHAPTER 1: Introduction
Introduction to the nuclear pore complex

Prokaryotic and eukaryotic cells differ in a number of ways, but perhaps most notably by the presence of a complex endomembrane system in the latter. This endomembrane system allows for the intricate regulation and execution of a variety of cellular processes. The hallmark feature of this endomembrane system is the nucleus, which houses the genetic material of the cell and physically separates it from the cytoplasm. However, such compartmentalization necessitates a mechanism to shuttle mRNA, proteins, and ribosomal subunits across the double membrane boundary for proper maintenance and execution of cellular functions. This need is filled by the nuclear pore complex (NPC), which serves as the sole gatekeeper between the nucleus and cytoplasm of the eukaryotic cell and tightly regulates the exchange of macromolecules across the nuclear envelope (NE).

NPCs were first described in the mid-twentieth century when early electron micrographs showed 8-fold symmetric nuclear pores embedded in circular openings within the NE where the inner and outer nuclear membranes fuse (Callan et al., 1949; Watson, 1959; Gall, 1967). Since then, much has been learned about the physical structure of the intact pore as well as its role in regulating nuclear import and export. The entire mass of the NPC varies between species, and in the yeast Saccharomyces cerevisiae it forms a ~44 MDa nanomachine (Rout et al., 2000), constituting one of the largest multiprotein assemblies in the cell. The central channel of the NPC is lined with proteins containing long stretches of phenylalanine-glycine (FG) repeats, which prevent the passive diffusion of molecules larger than approximately 40 nm in diameter (Panté and Kann, 2002). However, nuclear import and export factors, known as transport receptors, bind to cargo molecules and form transport complexes that are able to traverse the FG-network, allowing for the exchange of molecules between the nucleus and cytoplasm (Cook et al., 2007). Amazingly, the NPC can accommodate cargoes of
vastly differing sizes, ranging from 40 kDa proteins to mRNPs that can be a staggering 100 MDa in size (Grünwald et al., 2011).

While the main function of the NPC lies with its ability to tightly and efficiently regulate transport across the NE, several proteins of the NPC have been implicated in a number of other processes. It has been shown that one of the subcomplexes of the NPC in higher eukaryotes, the Nup107 complex (Nup84 complex in yeast), is targeted to kinetochores during mitosis, although the exact function is unknown (Belgareh et al., 2001). Additionally, nucleoporins and associated transport factors have been associated with both gene activation and repression (Akhtar and Gasser, 2007; Heessen and Fornerod, 2007; Casolari et al., 2004). In particular, gene activation has been linked to the site of nuclear pores, which are also thought to recruit RNA-processing and export machinery (Akhtar and Gasser, 2007). In contrast, the associated RanGEF shows association to transcriptionally inactive genes, which often localize to non-pore sites at the nuclear periphery (Casolari et al., 2004). Lastly, there is increasing evidence of the involvement of several nucleoporins in a variety of human pathologies including Triple A syndrome, atrial fibrillation, a variety of leukemias, as well as a diverse set of developmental defects (Cronshaw and Matunis, 2003; Lupu et al., 2008; Nakamura et al., 1996; Saito et al., 2004; Capelson and Hetzer, 2009).

The NPC is a dynamic entity. Not only is it critical for maintaining cellular homeostasis, but the NPC and its components appear to have much more wide-reaching effects on a number of other processes, many of which are just now coming to light. Therefore, having an understanding of the structure of the NPC in molecular detail is not only interesting, but also essential in understanding its function.
**Regulated transport across the NE**

**Protein transport**

The NPC is critical in mediating the exchange of macromolecules across the nuclear envelope, though the way in which smaller molecules, such as proteins and small RNAs (i.e. tRNAs, microRNAs) are transported differs substantially from the transport of larger cargoes such as mRNAs and ribosomal subunits. For nuclear protein import, a nuclear localization signal (NLS) on the cargo molecule is recognized and bound by an import receptor, forming a cargo complex which can traverse the FG-network (Görlich et al., 1996; Moore and Blobel, 1993; Melchior et al., 1993) (Fig 1.1). Upon reaching the nucleoplasm, the cargo is released as RanGTP binds the import receptor. Similarly, for export, a cargo molecule carrying an NES (nuclear export signal) forms a ternary complex with an export receptor and RanGTP. Once in the cytoplasm, GTP is hydrolyzed, leading the dissolution of the complex and release of the cargo molecule. The directionality of transport is maintained by the asymmetric localization of the guanine exchange factor RanGEF in the nucleus and the GTP activating protein RanGAP in the cytoplasm (Görlich et al., 1996), resulting in high concentrations of nuclear RanGTP and cytoplasmic RanGDP. Subsequently, RanGDP is returned to the nucleus by the help of the transport receptor Ntf2, allowing Ran to be recycled for further rounds of transport (Nehrbass and Blobel, 1996; Ribbeck et al., 1998).
Protein import and export are mediated by transport receptors, termed importins and exportins, with the directionality of transport being governed by the asymmetric distribution of the small GTPase Ran, in two different nucleotide-bound forms. The guanine exchange factor RanGEF maintains a high concentration of RanGTP in the nucleus, while the GTP activating protein RanGAP maintains a high concentration of RanGDP on the cytoplasmic side. For import, a transport receptor binds the cargo molecule and together they traverse the FG-network of the NPC. Upon reaching the nucleus, RanGTP binds the import complex, leading to dissociation of the import cargo. For export, RanGTP, a transport receptor, and an export cargo molecule form a ternary complex and cross the NPC. Once in the cytoplasm, RanGTP is hydrolyzed by the help of RanGAP, allowing for the release of the export cargo. RanGDP is recycled to the nucleus by Ntf2.

**Transport of mRNA and ribosomal subunits**

While the overarching principle of transport described above holds true, large
molecules such as mRNAs and ribosomal subunits, require additional factors for successful transport. With export cargoes as large as 100 MDa, some of these transport substrates substantially exceed the diameter of the transport channel of the pore itself, and require remodeling in order to be transported across the NE (Grunwald et al., 2011). Additionally, while the directionality of protein transport is mediated by the asymmetric distribution of the small GTPase Ran, the directionality of mRNA export is mediated by the cytoplasmic localization of the DEAD-box helicase Dbp5. It is thought that ATP-bound Dbp5 binds to mRNA molecules that reach the cytoplasmic side of the pore and aids in the removal of the heterodimeric Mex67-Mtr2 mRNA transport factors (Fig 1.2). Next, inositol-hexakisphosphate (IP6)-bound Gle1 binds to Dbp5 and activates ATP hydrolysis and mRNA release, resulting in conformational change of Dbp5 (Montpetit et al., 2011). Subsequently, Nup159 binds to Dbp5 and blocks the mRNA from rebinding. Additional conformational change results in the release of ADP and prepares Dbp5 for another round of mRNA binding.
The heterodimeric Mex67-Mtr2 complex binds mRNA and transports it through the FG-network of the NPC. At the cytoplasmic face of the NPC, ATP-bound Dbp5 binds to the mRNA and causes the Mex67-Mtr2 dimer to dissociate. Subsequently, inositol hexakisphosphate (IP6) mediates the interaction of Gle1 with Dbp5 and ATP is hydrolyzed to ADP, leading to mRNA release and conformational change of Dbp5. This conformational change allows Nup159 to bind Dbp5, preventing the mRNA from re-binding and inducing further conformational change of Dbp5 that allows for the dissociation of ADP, and ultimately the entire complex.

**Overall structure of the NPC**

While the soluble components and transport processes of the NPC are quite well understood, knowledge of its structure and assembly is still emerging. Several studies
using electron microscopy (EM) and scanning electron microscopy (SEM) have provided insight into the overall features of the structure of the NPC (Fig 1.3A). These studies revealed an 8-fold symmetric ring embedded in the NE where the inner nuclear membrane (INM) and outer nuclear membrane (ONM) are fused (Callan et al., 1949; Watson, 1959; Gall, 1967). While the electron dense modules forming the 8 spokes of the structural scaffold around the central axis of the pore exhibit an additional 2-fold symmetry across the plane of the NE, the nucleo- and cytoplasmic faces differ in composition (Hinshaw et al., 1992; Akey and Radermacher, 1993). On the cytoplasmic side, fibrous extensions are attached to each of the spokes of the electron dense scaffold (Kiseleva et al., 2004). On the nuclear side, a ring structure suspended below the NE, termed the nuclear basket, is attached to the NPC via 8 filamentous extensions. Furthermore, the approximate diameter of the NPC was determined to be ~100nm (Yang et al., 1998; Stoffler et al., 2003; Beck et al., 2007), while the height of the NPC is a bit more variable, falling between 30-50nm (Yang et al., 1998; Alber et al., 2007; Elad et al., 2009) depending on the species. While some species-specific differences between NPC structures have been observed, the overall topologies and compositions are conserved from yeast to humans.

To date, the best picture of a complete NPC has been achieved by using cryo-electron tomography (cryo-ET), which extended the resolution to 6 nm (Fig 1.3B) (Beck et al., 2007). In these images, three separate rings can be observed which constitute the three major domains of the NPC—the cytoplasmic filaments, the structural scaffold, and the nuclear basket. Each of these domains is repeated around the central axis of the pore, forming 8 spokes. The domain constituting the cytoplasmic filaments appears to be the least electron dense, which is not unexpected since such long extensions are likely quite flexible, and therefore invisible in the final averaged structure. In addition, several of the nups in the cytoplasmic ring exhibit more transient and short-lived
interactions at the NPC in comparison to the scaffold nups, which exhibit long residence
times and very slow protein turnover (Rabut et al., 2004; Daigle et al., 2001). The
regions of the cytoplasmic ring that are visible in the 3D reconstruction seem to float
above the structural scaffold, indicating that the material connecting these two domains
is also not particularly electron-dense. Similarly, the central transport cavity of the
nuclear pore shows no structural features, consistent with the idea of an aqueous central
meshwork of extended and unfolded FG-repeats that act as a transport barrier to the
free diffusion of large molecules through the NPC.

Aside from the central transport channel, cryo-ET images also reveal the
possible presence of peripheral channels (Beck et al., 2007). The idea of peripheral
channels has been postulated by a number of groups. They are thought to transport not
only small molecules and ions, but perhaps also proteins destined for the inner and outer
nuclear membranes (Kramer et al., 2007). Since the ER, INM, and ONM are
contiguous, a protein may first be inserted into the ER membrane, and then shuttled
through the peripheral channels of the NPC to reach its final destination in the INM or
ONM (Powell and Burke, 1990; Zuleger et al., 2008). Gaps of ~ 9 nm between the
structural scaffold and the NE have been observed (Hinshaw et al., 1992; Stoffler et al.,
2003), which would be big enough to accommodate the passage of molecules ~40kDa
in size, which is consistent with the average size of the cytoplasmic domains of proteins
embedded in the INM or ONM. However, a mechanistic understanding of how such a
process might work is still lacking.

While cryo-EM and cryo-ET studies have greatly enhanced our understanding of
the overall structure of the NPC, a higher resolution picture is still necessary to
understand the molecular workings and assembly of individual components within the
NPC.
Figure 1. 3 - Overall structure of the nuclear pore complex

(A) Scanning electron micrographs of an intact yeast nucleus (Kiseleva et al., 2004). Nuclear pore complexes are false-colored in blue. The top right panel shows the cytoplasmic face of the NPC, while in the lower panel, the nuclear basket is clearly visible from the nuclear face of the NPC. The 8-fold rotational symmetry of the NPC is clearly visible. (B) Three distinct architectural units—the cytoplasmic ring, the structural scaffold, and the nuclear basket—are evident in this electron tomographic image of the NPC at ~6nm resolution (Beck et al., 2007). Four of the 8 spokes of the NPC are shown. While this represents the highest resolution image of an intact NPC to date, the resolution still precludes the docking of crystal structures, or even single particle reconstructions of NPC subcomplexes.

Protein composition and size

In total, the NPC is composed of about 30 different proteins (Fig 1.4) that exist in multiple copies to compose a roughly 44 MDa multiprotein assembly in yeast (Rout et
al., 2000). Much of the work that went into establishing this inventory of nups stems from two studies using *Saccharomyces cerevisiae* (Rout et al., 2000) or rat hepatocytes (Cronshaw et al., 2002) as starting material. The cell extracts were fractionated and enriched for NPCs, and subsequently subjected to mass spectrometry to identify each of the purified proteins.

Both studies determined a total of ~30 nucleoporins that can be classified into three categories based on distinct sequence or fold elements. Three of the identified proteins contain transmembrane domains, and likely anchor the NPC into the NE. Another set of about 15 proteins contains highly structured domains, including α-helical and β-propeller folds, indicating a role for these proteins in forming the structural scaffold of the NPC. The remaining subset of nups is characterized by long, disordered regions containing FG-repeats at the N- or C- termini. As discussed previously, many of these proteins emanate into the central channel of the NPC and form the aqueous meshwork comprising the transport barrier.

Calculations based on the stoichiometries determined in the above studies led to the approximation of an NPC with a mass of roughly 44 MDa for *S. cerevisiae*, or 60 MDa for rat. However, previous hydrodynamic and volumetric calculations led to size estimations ranging from 66 MDa in yeast (Rout and Blobel, 1993) to 125 MDa in vertebrates (Reichelt et al., 1990). These size discrepancies can be reconciled by a lattice like model of the NPC, in which the individual components assemble to form an extended lattice, or meshwork, rather than a dense, globular assembly of proteins.
Chapter 1: Introduction

Nucdeoponn Metazoan homolog

Nup192
Nup188
Nc95
Nup53
Nup157/170
Nup53/59
Nup188
Nup120
Nup160
Nup84
Nup85
Nup145C
Sec13
Ndc96
Nup157
Nup155
Nup133
Nup107
Nup35
Nup133
Nup188
Nup84
Nup85
Nup157
Sec13
Nup133
Nup107
Nup84
Nup85
Nup35
Nup133
Nup107
Nup84
Nup85
Nup35

Domain Architecture

Nup192
Nup188
Nc95
Nup53
Nup157/170
Nup53/59
Nup188
Nup120
Nup160
Nup84
Nup85
Nup145C
Sec13
Nup133
Nup107
Nup35
Nup133
Nup107
Nup84
Nup85
Nup35

Abundance  Mass per NPC  Fraction of total NPC

Figure 1.4 - Inventory of nucleoporins

Each of the known yeast nucleoporins and their metazoan homologs is listed above, along with domain architectures as determined by x-ray crystallography and/or structure prediction methods (Brohawn et al., 2009). The nups are further categorized into defined subcomplexes, with the Nic96 and Nup84 complexes (composing the structural scaffold) constituting roughly 55% of the total NPC mass. Nups specific to metazoans are shown in italics.

Modularity of the NPC

Solving the structure of the assembled NPC in molecular detail is currently an unattainable goal. However, by taking advantage of the modular nature of the NPC (Schwartz, 2005), structural biologists have started to chip away at this formidable
First, the NPC is arranged in spokes around an 8-fold symmetric axis, immediately reducing the complexity 8-fold. Secondly, the NPC is composed of a set of ~30 proteins with a limited set of domain architectures. Perhaps most helpful is the fact that each spoke of the NPC can be further broken down into stable subcomplexes that can be studied individually (Fig 1.5). By using such a ‘divide and conquer’ strategy, each of the individually solved subcomplexes can be pieced back together to forge a picture of the assembled NPC (Schwartz, 2005).

**Figure 1.5 - Schematic representation of modular NPC subcomplexes**

The nuclear pore complex can be broken down into a number of subcomplexes that, together, form one spoke that is repeated 8 times around the central axis of the pore. The core of the NPC is shown in blue. The Nup84 complex is located symmetrically on the nucleo- and cytoplasmic sides of the NPC, with the Nic96 complex being present in only one copy. The FG-containing Nsp1 complex lines the transport channel of the pore. Shown in green are the transmembrane nups, which tether the NPC to the NE. The more peripheral components of the NPC, including the cytoplasmic filaments, the nuclear basket, and more transient nups, are shown in gray.
Breaking down the NPC

The ~30 nups composing the architectural domains of the NPC—including the cytoplasmic ring, the structural scaffold, the central transport channel, and the nuclear ring/basket—can be categorized into a number of biochemically-defined subcomplexes. Many of these subcomplexes were identified as individual entities in mitotic cell extracts via coimmunoprecipitation as well as by yeast two hybrid interactions (Matsuoka et al., 1999; Belgareh et al., 1998; 2001). In higher eukaryotic cells, the NE breaks down during mitosis and NPCs disassemble into these stable subcomplexes. When the NE reforms, they are sequentially recruited to reassemble functional NPCs (Dultz et al., 2008).

Three of the main subcomplexes of the NPC are the Nup84 and Nic96 complexes, which together form the structural scaffold, as well as the Nsp1 complex, which forms the central channel of the pore. The heptameric yeast Nup84 complex is the best-understood subcomplex of the NPC to date (Fig 1.5 and 1.6). It is composed of seven proteins that are conserved across all species—Nup85, Seh1, Nup145C, Sec13, Nup120, Nup84, and Nup133. Deletion of Nup84 complex members leads to a number of architectural and functional defects, ranging from pore clustering to mRNA export defects (Siniossoglou et al., 1996; Aitchison et al., 1995; Boehmer et al., 2003; Harel et al., 2003). Furthermore, immunodepletion of the entire Nup84 complex from cellular extracts in in vitro nuclear reconstitution assays leads to nuclei devoid of NPCs altogether (Harel et al., 2003). When reassembled in vitro from recombinant proteins, this complex arranges into an extended Y-shaped structure as shown by EM (Fig 1.6A and 1.6B) (Lutzmann et al., 2002; Kampmann and Blobel, 2009). Additional biochemical experiments have led to a detailed map of the binary interactions that together form the Y-complex (Fig 1.6C) (Brohawn et al., 2008; Leksa et al., 2009). In many species, excluding S. cerevisiae, three additional proteins are stably associated with the Y-
shaped Nup84 complex—Nup37, Nup43, and ELYS (Fig 1.4) (Cronshaw et al., 2002; Franz et al., 2007; Rasala et al., 2006). While distinct architectural roles for these proteins are still unknown, Nup37 from S. pombe has recently been shown to bind tightly to Nup120 (Bilokapic and Schwartz, in publication).

![Diagram of Nup84 complex](image)

**Figure 1.6 - Assembly of the Y-shaped Nup84 subcomplex**

(A) Early electron microscopic studies of the Nup84 subcomplex show that its 7 members can be reassembled in vitro to form a Y-shaped structure (Lutzmann et al., 2002). (B) A recent single particle reconstruction of the Nup84 complex is shown at ~3.5nm resolution, revealing several hinge regions within the Y-shaped complex and allowing for the general placement of solved crystal structures into the EM density (Kampmann and Blobel, 2009). While largely correct, the resolution still allows for misplacement of nups. In particular, the Nup85-Seh1 heterodimer is placed with the C-terminus of Nup85 pointing out from the Y, when biochemical data clearly place the C-terminus at the hub of the Y, where the three arms meet. (C) The ACE1 fold is found in three of the proteins of the Y-shaped complex—Nup85, Nup145C, and Nup84. This fold is characterized by three distinct modules. The N-terminus begins in the trunk region and leads into the crown module where the polypeptide chain makes a 180° turn and meanders back down the other side of the trunk, finally leading into a C-terminal tail domain. The placement of the three ACE1 proteins within the Y-shaped complex is indicated in the schematic on the right.
The Nic96 complex is composed of Nic96, Nup53/59, Nup157/Nup170, Nup188, and Nup192 (Marelli et al., 1998; Hawryluk-Gara et al., 2005; Amlacher et al., 2011), and is currently less well understood than the Nup84 complex. It is possible that the interactions within the Nic96 complex are weaker than those found in the Nup84 complex, making it more difficult to recapitulate and observe these interactions in vitro. In addition, 4 of the proteins in this complex are quite large and substantially more difficult to express recombinantly and purify for in vitro experiments as has been done with the Nup84 complex (unpublished data). Nonetheless, Nic96 has been shown to interact with Nup53 (Fahrenkrog et al., 2000), and also with Nup188 and Nup192 (Nehrbass et al., 1996; Kosova et al., 1999; Zabel et al., 1996). However, a detailed interaction map of the proteins within this complex is only beginning to emerge. Recent strides have been made to this end by using a thermophilic eukaryote, Chaetomium thermophilum, whose proteins are more stable and biochemically more tractable (Amlacher et al., 2011). In this study, Nup53 was shown to bind to Nic96, Nup170, Nup188 and Nup192 individually. In addition, Nic96 could bind directly to both Nup188 and Nup192 via an N-terminal helix, though never simultaneously, indicating competitive binding between Nup188 and Nup192 (Amlacher et al., 2011; Theerthagiri et al., 2010).

Lastly, the Nsp1 complex is formed through interaction of the C-terminal coiled-coil domains of Nsp1, Nup49, and Nup57 (Grandi et al., 1993). The N-terminal domains of these proteins contain natively unfolded FG-repeat domains that form the central transport barrier of the NPC. Recently, the structures of minimal binding complexes of Nsp1-Nup57 and Nup57-Nup49 were solved (Solmaz et al., 2011). The authors developed a model of the assembled transport channel based on crystal-packing interactions. However, since the coiled-coil domains of each of the three constituent proteins are cleaved into small, non-native fragments, some of the packing interactions seen in the crystal may be artifacts. Often, when proteins are truncated, hydrophobic
surfaces that are normally buried become exposed. In the context of the highly concentrated crystallization environment, these exposed hydrophobic surfaces provide optimal sites for crystal packing contacts (Hsia et al., 2007; Brohawn et al., 2008). Thus, models based on crystal packing contacts must be interpreted cautiously, especially when sufficient experimental evidence to support such a model is lacking.

Also, a second pool of Nsp1 forms a complex with Nup82 and Nup159 that is positioned on the cytoplasmic side of the pore (Belgareh et al., 1998), while the Mlp1/2 proteins together with Nup60 form the nuclear basket structure (Feuerbach et al., 2002; Strambio-de-Castillia et al., 1999). In addition, three transmembrane nups—NDC1, Pom34, and Pom152—tether the NPC to the NE (Onischenko and Weis, 2011).

Protein folds within the NPC

Aside from the subcomplexes discussed above, the limited fold composition of the NPC offers another level of modularity (Fig 1.4). Before crystal structures of nups were available, structure prediction methods led to the classification of all nups into four distinct fold categories—FG-repeat, coiled-coil, β-propeller, and α-helical (Devos et al., 2004; 2006; Schwartz, 2005).

In total, ~10 FG-repeat-containing proteins form the cytoplasmic filaments and the aqueous meshwork lining the central channel of the NPC. Among these, there is still substantial variety, as different sets of nups contain sets of either FG, GLFG, or FXFG repeats separated by spacer sequences of unique amino acid composition (Terry and Wente, 2009). Several studies have attempted to address the question of specificity of each of these FG-repeat types. While in vivo experiments show that there is some redundancy between FG-repeats and deletion of a subset of them can still maintain the integrity of the transport, in vitro experiments have also shown that certain transport receptors exhibit specific preferences for certain FG-repeat types (Terry and Wente,
While coiled-coil domains can often be promiscuous, the interactions between the coiled coils of the Nsp1 complex in the NPC appear to be very specific. Various in vitro experiments, as well as a recent structure of the complex, show that while Nup57 binds Nsp1 and Nup49 and, Nsp1 and Nup49 do not bind to one another directly (Schlaich et al., 1997; Solmaz et al., 2011). The coiled-coil domains of the nuclear basket proteins Mlp1/2, however, may be less specific and provide a platform for recruiting multiple proteins and accessory factors to the NPC (Strambio-de-Castillia et al., 1999).

β-propellers compose a significant portion of the NPC scaffold and are, in general, an abundant fold class across all eukaryotes. β-propellers have a variety of functions and often serve as binding platforms, mediating multiple interactions (Chaudhuri et al., 2008). Within the NPC, Sec13 and Seh1 were the first proteins to be identified as β-propellers due to the presence of canonical WD-40 motifs characteristic of β-propellers (Pryer et al., 1993). More nups have since been identified as β-propellers, despite the lack of canonical sequence motifs (Devos et al., 2004). Experimentally, the first β-propeller to have its structure solved was the N-terminal domain of Nup133, which is flexibly linked to a C-terminal helical domain (Berke et al., 2004). The structures of Sec13 and Seh1 soon followed, and both displayed a unique, open 6-bladed propeller structure (Hsia et al., 2007; Brohawn et al., 2008). In each case, the open propeller is completed by a 7th blade, which is contributed in trans by its binding partner. The propeller of Nup120 also proved to be unique, featuring a 4-helix bundle inserted between blades 6 and 7. (Leksa et al., 2009; Seo et al., 2009). In addition to the β-propellers of the structural scaffold, the structures of Nup159 and Nup82 from yeast, as well as Nup214 and Rae1 from human, have been solved as well (Weirich et al., 2004; Yoshida et al., 2011; Napetschnig et al., 2007; Ren et al., 2010). Judging from the
Chapter 1: Introduction

typical role of the β-propeller in mediating interactions, it is thought that the β-propellers of the NPC will likely play a pivotal role in linking multiple subcomplexes to assemble the NPC (Leksa and Schwartz, 2010). However, these interactions may be difficult to pin down, as they may be weaker.

The α-helical fold represents over half of the total mass of the NPC. Initial secondary structure predictions suggested that many of the architectural nups should fold into α-solenoids (Devos et al., 2006). α-solenoids are typically constructed of sets of 2-3 helices that are stacked repeatedly, resulting in an elongated molecule that runs from N- to C- terminus in one direction, often with a super-helical twist (Kobe and Kajava, 2000). In contrast to predictions, the structures of a number of helical nups revealed unique α-helical fold classes.

When the 65 kDa α-helical portion of Nic96 was solved, it was immediately evident that it did not represent a canonical α-solenoid, but instead, a new and unique class of helical fold (Jeudy and Schwartz, 2007; Schrader et al., 2008). Overall, the molecule is oblong and forms a J-shape. The N-terminus of the helical domain begins at the middle of the length of the protein and leads to one end of the ~140Å long molecule. Here, the polypeptide chain makes a 180 degree turn back on itself and meanders down the other side, past the N-terminus, extending into a tail domain with an additional 9 helices. In addition to Nic96, three other scaffold nups adopt this unique fold as well—Nup145C, Nup85, and Nup84 (Hsia et al., 2007; Brohawn et al., 2008; Debler et al., 2008; Brohawn and Schwartz, 2009b; Nagy et al., 2009). Despite extremely low sequence conservation, these proteins are structurally conserved, pointing to a common ancestor and leading to a new fold class termed ACE1 (ancestral coatomer element) (Brohawn et al., 2008). The ACE1 fold is characterized by a trunk, crown, and tail domain (Fig 1.6C). So far, Nic96 is the only ACE1 with all three domains solved. A second type of unique α-helical fold was observed when the structures of the α-helical
domains of Nup133 and Nup170 were solved (Boehmer et al., 2008; Whittle and Schwartz, 2009). These proteins are distantly related and share an extended and stretched α-helical fold. Lastly, another unprecedented domain architecture was discovered in the partial structure of Nup120, comprising the N-terminal β-propeller domain and the first ~350 residues of the α-helical domain (Leksa et al., 2009; Seo et al., 2009). The α-helical portion of the structure is built around a central stalk formed by 2 long helices that are encircled by an additional 9 helices. Furthermore, this α-helical domain forms an interface with the previously mentioned 4-helix bundle, inserted between the 6th and 7th blades of the propeller domain. Recently, the full-length structure of Nup120 from *S. pombe* was solved and revealed that while the first half of the Nup120 helical domain has a novel domain architecture, the C-terminal half of the helical domain is the most regular of all architectural nups, resembling a HEAT-repeat fold (Bilokapic and Schwartz, 2011, in publication). Overall, it is evident that the vastly differing α-helical topologies of the scaffold nups provide a significant challenge for structure prediction methods.

**Assembly of the NPC**

While breaking the NPC down into subcomplexes has provided an avenue for studying this huge macromolecular structure at high resolution (Fig 1.5) (Schwartz, 2005), a full understanding of the assembled NPC can only be achieved once each of these smaller subcomplexes are pieced back together. While some general information on the number and location of each of the subcomplexes of the NPC has been gathered based on the stoichiometric studies discussed above (Rout et al., 2000; Cronshaw et al., 2002), as well pull-downs and immuno-gold labeling experiments indicating relative locations of subcomplexes within the NPC (Rout et al., 2000; Fahrenkrog et al., 2002; Strambio-de-Castillia et al., 1999), the model for NPC assembly continues to be a topic
of debate within the field. Different approaches have led to varying models (Fig 1.7).

Alber et al. used a computational approach, taking into account vast amounts of biochemical and proteomic data to come up with general distance restraints and ultimately generate a 3-dimensional model of the NPC (Alber et al., 2007). In this model, two sets of 8 copies of the Nup84 subcomplex form separated rings, one on the cytoplasmic and one on the nucleoplasmic periphery, with Nic96 complexes forming an inner ring that is sandwiched in between (Fig 1.7). While the overall arrangement and relative positions of nups are plausible, it is clear that this approach has several limitations. Most notably, the Nup84 complex, which has been shown to form an extended Y-shape (Lutzmann et al., 2002; Kampmann and Blobel, 2009), is predicted to form a compact, oblong structure.

The Blobel group, on the other hand, has used mostly crystal-packing interactions to formulate a model for NPC assembly in which the Y-shaped Nup84 complex is arranged horizontally around the central axis of the pore (Hoelz et al., 2011). In this model, four copies each of the Nup145C·Sec13 and Nup85·Seh1 dimers form hetero-octomeric poles spanning the height of the NPC, leading this model to be termed the fence-pole model. More recently the model has been renamed the head-to-tail model, based on an interaction that was detected between Nup120 (head) of one Y-shaped complex and Nup133 (tail) from another. While there is some experimental evidence seemingly in support of this arrangement (Seo et al., 2009; Kampmann et al., 2011), the results are weak, leaving room for alternate interpretation and calling into question the validity of this model.

In our lab, we have come up with a lattice like model for NPC assembly that is based on its similarity to another membrane coating complex—COPII vesicle coats. The overall placement of NPC subcomplexes most closely resembles that of the computational model, though we predict a more porous and lattice like framework that is
in line with the previously discussed studies on stoichiometries and total protein composition of the NPC.

Figure 1.7 - Models of NPC assembly

Three differing models of NPC assembly are illustrated (Brohawn et al., 2009). Each model proposes the arrangement of proteins within the Y-complex, how the Y-complexes are organized relative to one another, as well as how they are arranged with respect to other subcomplexes of the NPC. In the first model, the Y complexes are arranged horizontally around the central axis of the pore in a head to tail fashion, with the Nup145C-Sec13 and Nup85-Seh1 heterodimers forming vertical poles perpendicular to the Y-shaped complex. In this model, interaction between Nup120 and Nup133 stabilizes this head to tail arrangement. A computational model was proposed based on distance restraints derived from biochemical and proteomic data. While the general placement of the Nup84 subcomplexes is generally accepted, the organization of the Y-shaped complex itself is not in line with biochemical data and known EM structures of the complex. Lastly, the lattice model was proposed based on the similarity between the ACE1 proteins of the Y-shaped complex and the COPII vesicle coat proteins. In this model, the Y-shaped complex is oriented so that the Nup145C-Nup84 edge element lines the positive curvature of the NE. However, it remains to be seen whether multiple Y-complexes form direct contacts with each other and also how they interact with other subcomplexes of the NPC.
Chapter 1: Introduction

Similarity to COPII vesicle coats

As discussed previously, the crystal structures of four nups—Nic96, Nup145C, Nup85, and Nup84—showed a novel tripartite α-helical fold termed ACE1 (Fig 1.6C) (Brohawn et al., 2008). Comparison to other solved structures revealed that one of the two members of the COPII vesicle coats, Sec31, shared a similar domain topology despite extremely low sequence conservation (6-10%). The other component of the COPII coat is Sec13. While an evolutionary relationship between nucleoporins and vesicle coat proteins had been suggested based on predicted domain architectures (the “protocoatomer hypothesis”) (Devos et al., 2004), we were able to provide the first experimental evidence for a common ancestry. This structural similarity was initially obscured due to the fact that the ACE1 fold contains 3 distinct modules that, between different proteins, exhibit various orientations relative to one another (Brohawn et al., 2008). When the domains are compared separately, it was evident that the α-helical domain of Sec31 also falls into the ACE1 fold category. In the crystal structure, two Sec31-Sec13 dimers interact via the crown domains of Sec31 (Fath et al., 2007). We were able to show that the crowns of Nup145C and Nup84 interact in a homologous fashion (Brohawn et al., 2008; Brohawn and Schwartz, 2009b).

When fitting the Sec31-Sec13 structure into EM images of the assembled COPII coat, the extended α-helical domains of two Sec31 molecules form the edge elements, while the β-propeller domain of Sec31 and the Sec13 β-propeller form the vertex elements (Fath et al., 2007; Stagg et al., 2006). Based on binding assays and EM reconstructions of the Y-shaped complex, it is clear that the helical elements are also extended, with the Sec13 and Seh1 β-propellers in close proximity to one another near the hub where the three arms of the Y meet (Kampmann and Blobel, 2009). The fact that the same protein, Sec13, exists in both the COPII vesicle coat as well as the NPC supports the notion that the assembly of these two membrane-coating complexes is
similar. We hypothesize that the Y-shaped complex forms an assembly unit of the NPC lattice in a way similar to which two Sec31·Sec13 dimers form an assembly unit of the COPII coat (Brohawn et al., 2008). In this model, the curved Nup84-Nup145C edge element of the Y lines the curvature of the nuclear membrane, similar to the way in which the Sec31 homodimer stabilizes the curved vesicle membrane (Brohawn and Schwartz, 2009b). Since Nic96 shares the ACE1 fold as well, it is possible that the Nic96 complex will also form an extended structure and aid in the formation of a lattice-like coat, stabilizing the highly curved NE where the inner and outer nuclear membranes meet. However, whether the Nup84 and Nic96 complexes form direct interactions is currently unknown.

Bridging the resolution gap

Often, structures of large, multimeric protein complexes can be solved by combining EM and crystallographic data to arrive at a composite structure for the entire assembly (Chiu et al., 2006). The assembly of the COPII coat was solved by placing the crystal structure of the two Sec31·Sec13 heterodimers into the 30Å EM map of the assembled COPII cage (Fath et al., 2007; Stagg et al., 2006). While EM maps of ~15 Å are typically required for the confident fitting of crystal structures, the COPII cage forms a very branched lattice, with limited placement options for the Sec31·Sec13 assembly unit. Currently, the resolution of the best EM structure for the assembled NPC is ~60Å (Beck et al., 2007), making it impossible to fit the x-ray structures of nups into the context of the entire NPC. Furthermore, while a 3D single-particle reconstruction of the Y-shaped Nup84 complex has been solved to 35Å (Fig 1.6B) (Kampmann and Blobel, 2009), this resolution still precludes the confident fitting of individual components, as evidenced by the incorrect placement of the Nup85·Seh1 dimer. Currently, the C-terminus of Nup85 points out from the center of the Y while biochemical binding data
clearly places it at the hub of the Y where the three arms meet (Fig 1.6C) (Brohawn et al., 2008; Brohawn and Schwartz, 2009b).

The resolution gap between available EM structures of the NPC and crystal structures of individual components or complexes is currently still too great to answer many of the questions concerning overall assembly. While single particle reconstructions provide a promising means for moving forward, it is also important to expand our knowledge of inter-subcomplex interactions. Building a detailed interactome of proteins within the NPC will surely enhance our understanding of how nucleoporin subcomplexes come together to assemble the intact NPC. To date, many pull down experiments have been performed in order to answer some of these questions (Alber et al., 2007; Amlacher et al., 2011). Unfortunately, they have succeeded only in identifying interactions within subcomplexes, suggesting that the interactions linking multiple subcomplexes may be weaker or more transient in nature. Additionally, while some yeast two-hybrid (Y2H) experiments have been performed, they are not exhaustive and cover only a portion of the full-length nups. Since a single nup may have multiple domains and mediate multiple interactions at once or in turn, it will be important not only to distinguish between specific interaction regions of each nup, but also to use high throughput methods that allow for the thorough testing of all nups and their individual domains to arrive at a high resolution interactome of the NPC.

Solving the NPC structure from the inside out

Coming in at just under 50 MDa and a total of ~500 proteins in yeast (Rout et al., 2000), the NPC’s shear size makes solving its structure an arduous task for structural biologists. However, the NPC has a number of intrinsic features, including its modularity and symmetry elements, that are key in forging a path toward a high resolution structure of the assembled pore (Schwartz, 2005). It is clear that the end goal can only be
reached using a multi-disciplinary approach. By starting with crystal structures of individual proteins or small complexes and combining them with ever-improving single particle reconstructions of subcomplexes, EM maps of the assembled NPC, and detailed biochemical interaction networks of inter-subcomplex contacts, we will eventually reach a full picture of the NPC. If the progress from the last few years is any indication of what is to come, it won’t be long before we solve the structure of the nuclear pore complex from the inside out.
Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats

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

*These authors contributed equally to this work.

S.G.B. and N.C.L. designed and conducted all biochemical, biophysical, and crystallographic experiments and wrote the manuscript. E.D.S. designed and conducted *S. cerevisiae* experiments and K.R.R. helped with crystallographic data collection and analysis. T.U.S. advised the project and wrote the manuscript.
Abstract

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

Introduction

Exchange of macromolecules across the nuclear envelope is exclusively mediated by NPC (Tran and Wente, 2006; Weis, 2003; D'ap Angelo and Hetzer, 2008). Whereas much progress has been made understanding the soluble factors mediating nucleocytoplasmic transport, the structure of the ~40-60 MDa NPC itself is still largely enigmatic. Cryo-electron tomography (cryo-ET) and cryo-electron microscopy (cryo-EM) have established the NPC structure at low resolution (Beck et al., 2007; Drummond et al., 2006; Stoffler et al., 2003). Crystal structures of scaffold NPC components are emerging (Berke et al., 2004; Boehmer et al., 2008; Hsia et al., 2007; Jeudy and Schwartz, 2007), but the resolution gap still precludes fitting into the cryo-ET structure. Overall, the NPC has eight-fold rotational symmetry with an outer diameter of ~100 nm and a core scaffold ring ~30 nm wide. The central FG-repeat containing transport channel measures ~40 nm in diameter, defining the maximum size of
The modularity of the NPC assembly suggests a path toward a high-resolution structure (Schwartz, 2005). Of the ~30 bona fide nucleoporins (nups) that comprise the NPC, only a core subset is stably attached (Rabut et al., 2004). In S. cerevisiae, this core includes two essential complexes; the heptameric Nup84 complex and the heteromeric Nic96-containing complex (hereafter called the Nic96 complex; unless noted all proteins are from S. cerevisiae). The Nup84 complex is composed of one copy each of Nup84, Nup85, Nup120, Nup133, Nup145C, Sec13 and Seh1. It self-assembles from recombinant proteins in vitro and forms a branched Y-shaped structure (Lutzmann et al., 2002). Deletion or depletion of individual components of the Nup84 complex leads to severe assembly defects in many organisms (Fabre and Hurt, 1997; Galy et al., 2003; Harel et al., 2003). The Nic96 complex is less well characterized, but appears to contain the architectural nucleoporins Nup157/170, Nup188, Nup192, Nup53, and Nup59 (Lusk et al., 2002; Marelli et al., 1998; Zabel et al., 1996). β-propellers and stacked α-helical domains form the building blocks of the constituents of the Nup84 and Nic96 complexes (Schwartz, 2005; Devos et al., 2006). Because vesicle coats (including COPI, COPII, and clathrin) share similar elements, a common ancestry has been hypothesized despite very low sequence homology and the absence of experimental structural evidence (Devos et al., 2004).

A recent computer-generated model of the NPC, based on a plethora of primary data from different sources, places the Nup84 complex at the NPC periphery, sandwiching the Nic96 complex in the center (Alber et al., 2007). In contrast, a model solely based on the structure of the Nup145C-Sec13 heterodimer and crystal packing interactions has been proposed and is diametrically opposed to the computer model (Hsia et al., 2007).
Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats

Results

We solved the structure of a complex of Nup85 residues 1-564 (of 744) and intact Seh1 (referred to as Nup85•Seh1) at 3.5 A (Table 2.1). Seh1 and Nup85 form distinct units in a tightly associated complex (Fig 2.1). Seh1 folds into an open six-bladed β-propeller structure. The blades fan out consecutively around a central axis, typical for canonical β-propeller structures (Chaudhuri et al., 2008). Between blades 1 and 6, the N-terminus of Nup85 is inserted and forms a three-stranded blade that completes the Seh1 propeller in trans. Following its N-terminal insertion blade, Nup85 forms a compact cuboid structure composed of 20 helices, with two distinct modules, referred to as ‘crown’ and ‘trunk’. Helices α1-α3 (residues 100-200) meander along one side of the trunk; the other side is formed by helices α12-α19 (residues 362-509) running in the opposite direction in an anti-parallel zig-zag to the C-terminus. The trunk elements are separated by an intervening crown composed of helices α4-α11 (residues 201-361) that form a distinct bundle that caps one end of the trunk. Helices α5-α10 in the crown module are almost perpendicular to the helices in the trunk.

In the asymmetric unit of the crystal, two heterodimers are aligned along a non-crystallographic dyad generating patches of contacts (Fig 2.2). This interaction is unlikely to be functionally meaningful, as the contact residues are poorly conserved in orthologs. Moreover, analysis of Nup85•Seh1 by analytical ultracentrifugation (AUC) showed a single species of ~104 kDa with a hydrodynamic radius of 4.4 nm (Fig 2.3). This hydrodynamic radius is close to the theoretical value calculated from the atomic coordinates using HYDROPRO (García De La Torre et al., 2000) and reflects the elongated shape of the 103 kDa Nup85•Seh1 complex (a spherical protein of 220 kDa would have the same radius). Gel filtration also showed that Nup85•Seh1 is a single 103 kDa heterodimer at concentrations up to 20 mg/ml (Fig 2.4). Hence, we restrict our analysis to this heterodimer.
### Table 2.1 - Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>Nup851-364-Seh1 Native</th>
<th>Nup85-Seh1 Selenomethionine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P4,2,2</td>
<td>P4,2,2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9792</td>
<td>0.9792</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a=b, c$ (Å)</td>
<td>112.6, 350.5</td>
<td>112.5, 351.2</td>
</tr>
<tr>
<td>$\alpha=\beta=\gamma$ (*)</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>40-3.4</td>
<td>50-3.7</td>
</tr>
<tr>
<td>$R_{sym}$ (%)</td>
<td>10.7</td>
<td>12.7</td>
</tr>
<tr>
<td>$I/\sigma$</td>
<td>13.7 (1.8)</td>
<td>21.9 (1.8)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.9 (98.8)</td>
<td>99.1 (99.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.1 (3.3)</td>
<td>6.3 (6.1)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>30-3.5</td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>28300</td>
<td></td>
</tr>
<tr>
<td>$R_{work} / R_{free}$</td>
<td>32.6 / 36.9</td>
<td></td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>9638</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>R.m.s deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Bond angles (*)</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Coordinate error (Å)</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats

Figure 2.1 - Structure of the Nup85-Sehl complex

The structure of the heterodimeric Nup85-Sehl complex is shown in two views (A, B), related by a 90° rotation around the horizontal axis. Nup85 has a trunk (orange, helices α1-α3 and α12-α20) and a crown (blue, helices α4-α11) module. The β-strands at the extreme N-terminus of Nup85 form an insertion blade, which complete the Sehl (green) β-propeller. (C) 2Fo-Fc omit map (contoured at 1.2σ) with a Ca-trace of the Nup85-Sehl complex.
Figure 2.2 - Arrangement of two Nup85-Seh1 heterodimers in the asymmetric unit

(A) Association and orientation of the two Nup85-Seh1 heterodimers in the asymmetric unit. The heterodimers form an interface of ~1800 Å² and associate lengthwise along a two-fold axis. (B) Surface conservation of Nup85-Seh1 in a view 90° rotated from (A) with outlines corresponding to contact regions involved in forming the interface. Red and blue outlines indicate contacts made with Nup85 and Seh1, respectively, and conservation is shaded from white (not conserved) to orange (conserved). (C) The electrostatic surface potential of the Nup85-Seh1 heterodimer (colored from red (-8 kT/e) to blue (+8 kT/e)) with outlines as in (B). Yellow and blue outlines correspond to contacts made with Nup85 and Seh1, respectively. The view is the same as in (B).
Figure 2.3 - Nup85·Seh1 is a dimer in solution as determined by analytical ultracentrifugation

(A) C(s) distribution analysis of Nup85·Seh1 sedimentation velocity data. Sedimentation data from three concentrations of Nup85·Seh1 were analyzed globally in Sedphat (Schuck, 2000) with a hybrid local continuous distribution and global discrete species model. Data was fitted from 2 to 10 s⁻¹ with fixed partial specific volume. Residual plots for each concentration are shown below. The sedimentation coefficient and corresponding rmse values for the samples in order of decreasing concentration were 5.56 s⁻¹, 5.58 s⁻¹, and 5.60 s⁻¹ and 0.0079, 0.0056, and 0.0045, respectively, which corresponds to a single species with a molecular weight of ~104 kDa and frictional ratio (f/f₀) of 1.43. The calculated molecular weight for Nup85·Seh1 is 103 kDa. (B) Sedimentation equilibrium analysis of Nup85·Seh1. Sedimentation data from 6 concentrations at
three speeds were analyzed globally in Ultrascan 9.0. The data was best fit by an ideal single species model. The top panel shows data points as yellow triangles with fit curves overlaid (13.5 k rpm – dark blue, 17.5 k rpm – medium blue, 22.8 k rpm – light blue). The lower panel shows residuals of the fitted curves. The molecular weight was determined to be 99 kDa with a standard deviation of 0.4 kDa, closely matching the results obtained via sedimentation velocity.

![Graph showing elution profiles](image)

**Figure 2.4 - Nup85-Seh1 is a dimer in solution as determined by size exclusion chromatography**

Elution profiles of the Nup85-Seh1 complex at 1, 5, 10, and 20 mg/ml show a single peak eluting at 12.2 ml on a Superdex 200 10/300 (GE Healthcare) column indicating a hydrodynamic radius of 4.4 nm. The hydrodynamic radius was independently determined by sedimentation velocity and is consistent with the value calculated from the experimental crystal structure using HYDROPRO (Garcia De La Torre et al., 2000). The elution profile of Nup145C-Sec13 is shown for comparison (dashed black line).

The connectivity and topology of secondary structure elements and the three-dimensional folds of Nup85-Seh1 and Nup145C-Sec13 (Hsia et al., 2007) are remarkably similar (Fig. 2.5A), despite very low sequence identity between Nup85 and Nup145C (10%) and moderate identity between Seh1 and Sec13 (32%). Like Nup85, Nup145C has an N-terminal three-stranded β-sheet that provides a seventh blade to close open β-propeller of Sec13. The trunk and crown modules of Nup145C are also
similar to those in Nup85, although their relative orientation is modestly different in the two proteins.

The most conserved regions of Nup85 are involved in the interaction with Seh1. The corresponding interface between Nup145C and Sec13 is also well conserved, but Nup145C has an additional highly conserved surface on the crown module, around helix \( \alpha 8 \), that is not observed in Nup85 (Fig 2.5B). This region is reasonably polar and poorly conserved in Nup85 but highly conserved and distinctly hydrophobic in Nup145C, suggesting a protein-protein interaction site. Nup84 and Nup120 bind to roughly opposite sides of Nup145C•Sec13 in the Y-shaped complex (Lutzmann et al., 2002), and the C-terminal helical region of Nup145C is necessary for binding Nup120 (Fig 2.6). Thus, we hypothesized that the \( \alpha 8 \) crown surface of Nup145C is the binding site for Nup84.

To test this hypothesis, we mutated the Nup145C sequence "VLISY" in \( \alpha 8 \) to "ELIEA," introducing two negative charges and eliminating a conserved aromatic side chain on the crown surface (Fig 2.5B). The overall structure of Nup145C did not appear to be perturbed by this modification: (i) Nup145C-ELIEA•Sec13 bound to Nup120 to form a 1:1 complex indistinguishable from Nup145C•Sec13 in gel-filtration experiments (Fig 2.7); (ii) Nup145C•Sec13 and Nup145C-ELIEA•Sec13 complexes had comparable thermostability (Fig 2.8) and (iii) showed identical behavior in gel filtration (Fig 2.5C). The ELIEA mutation completely eliminated Nup84 binding. In isothermal-titration-calorimetry (ITC) experiments, Nup84 bound wild-type Nup145C•Sec13 tightly (\( K_D = 3 \pm 2 \) nM; 1:1 stoichiometry) but not Nup145C-ELIEA•Sec13 (Fig 2.5D). Similarly, Nup84 formed a stable complex with Nup145C•Sec13 but not with Nup145C-ELIEA•Sec13 in gel-filtration (Fig 2.5C). We conclude that the Nup84 binding site on Nup145C includes the exposed surface of helix \( \alpha 8 \).
Figure 2.5 - Comparison of Nup85-Seh1 and Nup145C-Sec13 and identification of the Nup84-Nup145C crown-crown binding interface

(A) The topologies of the Nup85-Seh1 (left) and Nup145C-Sec13 (right, PDB code 3BG1 (Hsia et al., 2007)) complexes are shown, illustrating an overall similarity with three shared structural elements—trunk, crown, and β-propeller. Colors are assigned as in Figure 2.1. (B) Surface representations of the crowns of Nup85 and Nup145C are shown colored according to electrostatic surface potentials (top) and sequence conservation (bottom) in a view rotated 90° from (A). Sequence conservation is based on the phylogenetic tree of budding yeasts (Dujon, 2006) and is colored from white (not conserved) to orange (conserved). A partial sequence alignment of helix α8 (indicated by arrows in (A)) is also shown with surface exposed residues indicated by green dots, residues buried in the hydrophobic core by blue dots, and residues not modeled in the structure by dashes. Mutations made in this helix in Nup145C are shown above the sequence alignment and the corresponding residues are outlined in the surface representations of Nup145C. (C) In the upper panels gel filtration data of Nup84 alone,
Nup145C•Sec13 (wild type or -ELIEA mutant) alone, and Nup84 plus Nup145C•Sec13 (wild type or -ELIEA mutant) are shown. The shift in the Nup84 plus wild type Nup145C•Sec13 chromatogram indicates complex formation and is absent in the case of the -ELIEA mutant. In the lower panels gel filtration data of Nup145C•Sec13 alone, Nup84 alone (wild type or -DSICD mutant) alone, and Nup145C•Sec13 plus Nup84 (wild type or -DSICD mutant) are shown. The shift in the Nup145C•Sec13 plus wild type Nup84 chromatogram indicates complex formation and is absent in the case of the -DSICD mutant. (D) Isothermal titration calorimetry data illustrating high-affinity binding for wild-type Nup145C•Sec13 and Nup84 (black). Experimental values for N, K_D, ΔH, and TΔS are shown. In contrast, binding is lost for both crown-surface mutants Nup84-DSICD (grey) and Nup145C-ELIEA (red).

Figure 2.6 - Nup120 binds Nup145C and Nup85 via their tail modules
(A) The heterodimeric Nup85(fl)•Seh1 and Nup145C•Sec13 complexes were analyzed on a Superdex200 HR26/60 column. A mixture of both complexes (orange) does not result in a higher molecular weight species indicating that the complexes do not directly interact. (B) The
heterodimeric Nup85-Sehl complex (without the Nup85 tail module) and Nup120 were analyzed on a Superose 6 HR10/300 column. Again, a mixture of both samples (orange) does not result in a higher molecular weight species. (C) Nup85(fl) including the tail module binds Nup120 (blue), and adding Nup145C•Sec13 results in a pentameric complex (orange) (Superdex S200 HR10/300). This complex is not formed when the tail module is removed from Nup145C (dashed). Taken together, this series of experiments demonstrates that the tail modules of both Nup145C and Nup85 are responsible for Nup120 binding.

Figure 2.7 - A single point mutation in the predicted interaction helix of the Nup145C tail module disrupts Nup120 binding

Formation of a pentameric Nup120•Nup85(fl)•Sehl•Nup145C•ELIEA•Sec13 complex (orange) is disrupted by the Q691G mutation in the tail module of Nup145C (dashed) (Superdex S200 HR25/60). Formation of the Nup120•Nup85•Sehl complex is unaffected by this mutation. Fractions were analyzed by SDS-PAGE.
Figure 2.8 - Surface mutation of the Nup145C and Nup84 crowns does not negatively affect protein stability

Formation of a pentameric Nup120-Nup85-Fl-Nup145C-ELIEA-Elie13 complex (orange) is disrupted by the Q691G mutation in the tail module of Nup145C (dashed) (Superdex S200 HR26/60). Formation of the Nup120-Nup85-Elie1 complex is unaffected by this mutation. Fractions were analyzed by SDS-PAGE.

To determine the consequences of abolishing the Nup84 binding site on Nup145C in vivo, we introduced the Nup145C-ELIEA mutation into the NUP145 gene in yeast. Strains carrying NUP145-ELIEA in a ΔNUP145/NUP84-GFP or
ΔNUP145/NUP133-GFP background displayed a marked defect in incorporating Nup84-GFP and Nup133-GFP into the NPC (Fig 2.9 and 2.10). Compared to wild type, a significantly larger fraction of GFP-tagged proteins were found in the cytoplasm, indicating that the Nup84 binding interface on Nup145C is crucial in recruiting both Nup133 and Nup84 to the NPC (Fig 2.11). In addition, nuclear pores were clustered into discrete foci on the nuclear envelope of the strains expressing Nup145C-ELIEA, indicative of severe NPC assembly defects and similar to Nup84 and Nup133 null strains (Doye et al., 1994; Siniossoglou et al., 1996). Cells expressing wild-type Nup145C demonstrated the expected punctate nuclear rim staining in both Nup84-GFP and Nup133-GFP strains. Thus, disruption of the Nup84 binding site on Nup145C affects NPC assembly and function and causes loss of Nup84 and Nup133 from pores. The loss of Nup133 can be rationalized because it is attached to the Y-shaped Nup84 complex through a binary interaction with Nup84 (Boehmer et al., 2008; Lutzmann et al., 2002). Some Nup84 and Nup133 proteins remain associated with nuclear pores in the Nup145C-ELIEA expressing strains, arguing for the existence of additional weaker attachment sites for both proteins in the NPC. It has been shown that an ALPS membrane-binding motif is present in Nup133 (Drin et al., 2007). Because Nup133 and Nup84 are tightly associated (Boehmer et al., 2008), the ALPS motif might be weakly functional in recruiting Nup133-Nup84 to the NPC even when the Nup84-Nup145C interaction is compromised.

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

Figure 2.9 - Elimination of the Nup84 binding site on Nup145C results in nuclear pore assembly defects in vivo

(A) NUP145/NUP133-GFP and NUP145-ELIEA/NUP133-GFP or (B) NUP145/NUP84-GFP and NUP145-ELIEA/NUP84-GFP were grown at 24 °C and visualized by fluorescence microscopy. DIC, GFP-fluorescence, DNA (visualized with Hoechst dye), and false-colored overlay (GFP fluorescence – green, DNA – blue) images of the same field are shown in columns from left to right.
Figure 2.10 - Growth analysis of yeast strains

(A) The nup145C-ELIEA mutant supports viability. nup145Δ/NUP133-GFP and a nup145Δ/NUP84-GFP strain carrying SBYp115 (NUP145/CEN/URA3) and empty vector, or vector encoding NUP145, or vector encoding nup145C-ELIEA were grown in SMM-leu overnight, serially diluted and grown on SMM-leu or SMM-leu + 5-FOA plates at 30°C for two days. (B) nup145C-ELIEA is lethal at elevated temperatures in rich media. A nup145Δ strain carrying plasmid-borne NUP145 (SBYp116) or nup145C-ELIEA (SBYp117) were grown in minimal medium at 24°C. Serial-diluted cells were plated onto YPD plates and grown for 2 days at 24°C, 30°C or 37°C.
Figure 2.11 - Nup133-GFP and Nup84-GFP become more soluble in the Nup145-ELIEA strain

(A) NUP145 NUP133-GFP (YS221) and nup145C-ELIEA NUP133-GFP (YS222) or (B) NUP145 NUP84-GFP (YS223) and nup145C-ELIEA NUP84-GFP (YS224) strains were spheroplasted, and total lysates (T) were separated into 16,000×g soluble (S) and pellet (P) fractions. Equal cell equivalents from each fraction were analyzed by immunoblotting using rabbit anti-GFP or monoclonal anti-Pgk1 antibodies.

The structural similarity between Nup85 and Nup145C extends to at least three other proteins (Figs 2.12 and 2.13). First, the architectural nucleoporin Nic96 (Jeudy and Schwartz, 2007) shares a common structural core (Fig 2.13) but has a distinct N-terminus (Fig 2.12). The shared cores mutually superimpose with an rmsd of 3.0-3.5 Å. Nic96 has a trunk module (α1-α3 and α12-α19), a crown module (α4-α11), and an N-terminal coiled-coil extension (instead of the insertion blade of Nup145C and Nup85).
that tethers it to the FG-containing Nsp1 complex (Grandi et al., 1995). Apart from the N-
terminal differences the three proteins differ mainly in the relative orientation of the
crown and trunk modules. Although a previous comparison of Nup145C to the COPII
coop component Sec31 did not reveal a strong similarity (Hsia et al., 2007), comparison
with Nup85, Nup145C and Nic96 shows that Sec31 has corresponding trunk (α1-α3 and
α12-α18) and crown (α4-α11) modules. Sec31 homodimerizes to create an “edge
element” in the COPII coat by an internal domain-swap between two crown modules
(Fath et al., 2007). This domain swap results in a mixed crown module that is identical in
topology to the unmixed crowns in Nup85, Nup145C, and Nic96 (Fath et al., 2007).
Structural prediction using Phyre (Bennett-Lovsey et al., 2008) also places Nup84 in the
Nup85/Nup145C/Nic96/Sec31 group. Similarity extends beyond the trunk and crown
modules to a ‘tail’ module that has been structurally characterized in the C-terminal
domain of human Nup107 (homolog of yNup84) and in Nic96 (Boehmer et al., 2008;
Jeudy and Schwartz, 2007) (Fig 2.13C). The last three helices in the tail module of
Nup107 form the interaction site with Nup133 (Boehmer et al., 2008). In Nic96 this
region is predicted to be protein binding site as well (Jeudy and Schwartz, 2007).
Because we find this characteristic tripartite structural element of crown, trunk and tail in
architectural proteins of the NPC and the COPII coat, we term it Ancestral Coatomer
Element 1 (ACE1).
Figure 2.12 - Architecture of ACE1

(A) ACE1 containing proteins are shown as cylinders and sheets. Crowns are shown in blue, trunks in orange, tails in green, and other domains in gray. Modules with predicted structures are shown half-transparent. (PDB codes: 2QX5, Nic96; 3BG1, Nup145C; 3CQC, Nup107 (Nup84 homolog); 2PM6, Sec31) (B) Cartoon illustrating the similarity and modular nature of the ACE1 element. The N-terminal elaborations are for Nic96 a coiled-coil domain that interacts with the Nsp1 complex, for Nup85 the Seh1-interacting insertion blade, for Nup145C the Sec13-interacting insertion blade preceded by an autocatalytic cleavage domain and Nup145N, and for Sec31 the Sec13-interacting insertion blade is preceded by its own N-terminal 7-bladed β-propeller. Sec31 has a unique proline rich insertion C-terminal to its trunk module followed by a conserved region predicted to be α-helical.
Figure 2.13 - Superposition of ACE1 modules

Superposition of (A) crowns, (B) trunks, and (C) tails of ACE1 with known structures (PDB accession codes 3bg1, Nup145C; 2qx5, Nic96; 3cqc, Nup107; 2pm6, Sec31). Helices are labeled according to Nup85 in (A) and (B) and Nic96 in (C). In (A) Helix α4 of Sec31 and the short beta sheet between helices α5 and α6 in Nic96 are omitted for clarity. The rmsd between modules is 2.9-3.2Å in (A), 3.4-3.5Å in (B) and 2.7Å in (C).

Can we predict ACE1 functional sites from established interactions? Analogy to
Sec31 monomers in the COPII edge element suggests that Nup145C and Nup84 might interact crown to crown. Based on the Phyre-model and structural alignment we constructed a surface point mutant replacing two conserved hydrophobic residues on helix a8 of the Nup84 crown with aspartate (Nup84-ISICM to Nup84-DSICD) (Fig 2.8 and 2.14). Nup84-DSICD disrupts Nup145C binding in a manner analogous to Nup145C-ELIEA severing Nup84 binding, as shown by gel filtration and ITC (Fig 2.5C,D). Thus, the Nup84•Nup145C interface is a crown-crown interaction involving a8 helices as in Sec31 homodimerization. Additionally, we found that the tail modules of Nup145C and Nup85 are necessary for interaction with Nup120 in a manner analogous to the human Nup107 interaction site for human Nup133 (Boehmer et al., 2008) (Fig 2.6 and 2.7).

Figure 2.14 - Surface mutation of the Nup145C and Nup84 crowns

The crown helix a8 is highlighted in (A) Sec31 (PDB ID 2pm6) (B) Nup145C (PDB ID 3bg1) and (C) Nup84. In (A), one Sec31 monomer is shown half transparent and the three hydrophobic residues buried in the interaction between crown helix a8 in each monomer are shown as sticks. In (B), the three exposed a8 residues in Nup145C mutated to disrupt binding to Nup84 are shown (VLISY to ELIEA). In (C), the two exposed hydrophobic a8 residues in the Phyre-predicted Nup84 structure mutated to disrupt Nup145C binding are shown (ISICM to DSICD).

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

**Discussion**

ACE1 is different from regular α-helical repeat structures, including HEAT-repeats and TPR-repeats (as discussed in (Jeudy and Schwartz, 2007)). The α-helical modules that compose ACE1 are distinctly irregular, most notably with elements that fold back onto themselves forming a U-turn within the crown module. The trunk is composed of two zig-zaging helical units running in opposite directions. We propose that this architecture confers rigidity to the trunk and thus distinguishes it from regular helical repeat structures that are often inherently flexible (Conti et al., 2006). As a consequence of the specific arrangement of the helices in ACE1, several helices in trunk and crown are encased by neighboring helices and thus have a characteristic hydrophobic character (typically helices α6 and α10) (Fig 2.15). This pattern of hydrophobic helices may help to find additional ACE1 proteins. Several sequence elements, notably in the crown and at the predicted hinge regions, distinguish ACE1 from other α-helical domains (Jeudy and Schwartz, 2007). Nevertheless, these characteristics are subtle enough to remain undetected in typical primary sequence (i.e. BLAST) searches and candidate proteins need to be examined using all available tools, including phylogeny and secondary and tertiary structure analysis.
Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats

A

**Nup85**

- **S. cerevisiae**: MTID---DSNKLMMQDQDFDLDDCTAQLSLNKTDEEQLYKRPDSVGSAILVPM 51
- **C. glabrata**: DLVMDYDFQDPQPSMETKAA---NEJTRPIGQSVVIPV-- 41
- **K. lactis**: MANDEFA---DTQQLMMQDQDFDLDDCTAQLSLNKTDEEQLYKRPDSVGSAILVPM 52
- **A. gossypii**: ---hmsham---ELMLDDQCMDFVDEDMECAGE---LQFSTDPPSNAPMVMSF 44
- **D. hansenii**: MCDQ---TSLSKFDIEMLEPDDD---CTESFS 34
- **C. albicans**: ---MPLFY---NEFQESDFMLDELPDEDDD---C---ENSIDNISSETECSIS 41
- **Y. lipolytica**: MFSPAPATINGADYNISPDLTAREFGEAMETLIDQDD---ELPEITLPCAPI 52

**B**

- **S. cerevisiae**: TVNODQI-EKNCDKMPL---FKLCPFLSYPYQNMAT---ITAKDKYLYPLYPR 97
- **C. glabrata**: DAKFPKESHNL---VFLKGCAGANINM---VYQSKDVKYNPVLPVL 85
- **K. lactis**: ---PTEQPS-ALKELRKYMKNVSSESRAF---DNSTKDRKYLNVPFLL 92
- **A. gossypii**: PSTQQP---DALKRFLFPVLSSRSAF---NGCSRRTCYLCYQVYPY 85
- **D. hansenii**: DSAS---EPQLDKPSQDDYLRLQDWKKS---DDDI1LDQDSDQGKRTLD 79
- **C. albicans**: DSDDSSSSSYQDWSVPI-FKSKSDEVILYDSLQVWINT---KCELKPEFQXKRYRKGIE 91
- **Y. lipolytica**: PEGQVEEWQKS---RNMGFVMDPVIAARCIAWFADKPCPGDQGTHAKDRELTTPTRV 105

**C**

- **S. cerevisiae**: L-D-TKKEPSFASYVSSL-EIYRDLCDDDVF---NYPTIGVV---NSFAX 138
- **C. glabrata**: L-D-FSEIFINXENLFLDIYQULCNSIF---SRPTIGVI---SSYDEL 126
- **K. lactis**: L-D-EKSEPAFYQVSLFLYELQCLCENK---DVPTIGLI---KQTRSL 133
- **A. gossypii**: YQMDKFPQAYQAKLTLTVPYSLGKADQ---AVPTIGLI---HNTSTL 126
- **D. hansenii**: N-D-VSKSTSYTILSNK-IIGCFDNTIDRILDLCDDDPIGLTSASKFCSARKA 133
- **C. albicans**: SRP-FDEKEYLYYYNNMKIQFQKLDTEDDTRFDLEDESSPIGLVMDMS---MGSSAKQV 147
- **Y. lipolytica**: C-DEWMDNKKGFQVQEAATTLLC-PQ--- 128

**D**

- **S. cerevisiae**: EHENATVNLAMCPLNLLENFLVFTVRKWDQQ---DCRVRNFVFYELLESLTWNLCTR 168
- **C. glabrata**: EHTRVNYMLSDMIELLLEELIASYDKLE---NANAIRILELQELCQLLGLCRR 177
- **K. lactis**: EHFISINLABHVTLETEFIKESIKY---TNKLRQICDLEECLELSNLGKTL 162
- **A. gossypii**: EHNTQTVNALLAEAVSSELEFSKLEY---SCRQKQVLDLLECLESNLKTV 175
- **D. hansenii**: QRMHKIDFETLTKYQYSHDESIG---VDDTEEVOQFYLYSLDELKAN 163
- **C. albicans**: RKKFKXIALARELIDYQLQDNSVPIKEE---QESYSQSGCQLHLVLCEFEAN 167
- **Y. lipolytica**: EAELTDQRFADFESIKHYREEIKRDVQDVEAELPLETALASY 176

**E**

- **S. cerevisiae**: YFDLQOADO---VEENRSEFESLLENNWIRSDGEDEPEYEQIEVSVKDSTACKK 259
- **C. glabrata**: QFDQ---TDINDFROSLLMKHTOGCPKREVYQGIQGQ---DQGVQ 230
- **K. lactis**: YFTLQDSEP---Y-KGEDDPLESINMHRSDGSCPSVIELKQFERDTTLL-TRRKV 230
- **A. gossypii**: QTFTLDSEE---CRSRAKIDSLINWYRTQDESEAYIAKRLCDG---KQTP 220
- **D. hansenii**: YETCDATMR---PESIAKIRNIAKTDYDPDDLKLVLQMV---NSKP 222
- **C. albicans**: HYMNDVQQ---KXNHTVTOPDIFDCLP---ELRNNTRQP 258
- **Y. lipolytica**: YHNGCAAGQSFGFSQSTYERVQPLYEVEVSQVESVLEKVEMSSTQ---P--- 176

**F**

- **S. cerevisiae**: FEQTYQFWK--LNLVQLRGLCSAIACIGCRSSOLPPLYLSDTCAVSTDASVSDIIHELQK 264
- **C. glabrata**: FFFETEFWK-EVLYQCPDDCNLAIKSMELKYLEXKCTETTTFITPLEVLKRN 276
- **K. lactis**: YEQQSDPDOL-LAQTGDRCAQVAAGAEKSSLAYLDSKASCATHTMQMVNLQDCGQ 226
- **A. gossypii**: LENPYFWR--LQCQRLIQDLLQDALVAQAAEEKSSLYADKSCAATHTMQMVNLQDC 217
- **D. hansenii**: YTHQPCFNWYQSLRCLRCIIAASIAKQY-EELENNPAFELSSVQIQSMILLEN 277
- **C. albicans**: YHPFIFWNCTSGQELTEKLIGYFTQKEFPELYAVGMDTLIL 291
- **Y. lipolytica**: VQICDFSTCGAGQSHCAGCA-CEVDFSCQGTECQTILDIYLYRPQ 280

**G**

- **S. cerevisiae**: YPXDS-STFREKSNLVKLQSLQAGQFSATGDSATIGSE---LDYDFELFVLCNGQ 344
- **C. glabrata**: YPELE-SECHREGWKSFALEAQQNFSEDADTPFVP---LRKNLISAIAGTDN 326
- **K. lactis**: YPLH--EMLFREKSNVLQLMHMQISEQ---TKKMNMMFLISGSXN 350
- **A. gossypii**: YPRE-LEPMYPREKDSMLVQGLQHLKLSWQESHEKISIC---LASSBCDCLLMNGCNS 327
- **D. hansenii**: YTMSSMKHQFESLCSFCERDRFLFJACKANT-DSKDLINLQYDLCFLTCGLP 352
- **C. albicans**: YTSYSQSKQFTMKLACEFRRDSSLVSMKNETTCKLHKLIDQYIADLCITFTGLP 347
- **Y. lipolytica**: CPGN-VYSFYROLHISSTGVQSKLLKIN---DAX---JQKCTLTLVLALAGDE 327

**H**

- **S. cerevisiae**: KLIQLSYRTFYEFSGCFLYYI-PSSLESEAYLQMSLEANVVDITND---W 360
- **C. glabrata**: KLIELQYAVLCAYSTKFT---IDTVNP--- 372
- **K. lactis**: KICETQSFYSCVYCLIMYYI-PITIIFSQYECGATKHNADVCNN---W 376
- **A. gossypii**: KIIYYSTKYTECVCMILYI-PSSLSLEAYLQVLKHEPKLDVTSP 386
- **D. hansenii**: TIAHAKIXCYTKVAIPKTVIVESCQTVGQYVLGDVSIS--- 436
- **C. albicans**: KSTEPNYEYVLYLALSQYRDRSDNLEYIDYFTKASVSKPFSIDQDENENJLQ 403
- **Y. lipolytica**: AMCTNNHTDIALTVQODYPDCPSRMCYTYDAVANYY-V-DTFTV---W 373
Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats
Chapter 2: Structural evidence for common ancestry of the nuclear pore complex and vesicle coats
Figure 2.15 - Sequence alignments

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

Based on distance constraints and stoichiometric considerations, the heptameric Y-shaped Nup84 complex has been placed in two concentric eight-membered rings on the nucleoplasmic and cytoplasmic faces of the NPC (Alber et al., 2007). But how is it oriented and how is it connected to the inner ring of the scaffold? Nup133 is anchored to the structural scaffold by its interaction with Nup84, positioning it at the periphery of the pore (Boehmer et al., 2008). Nup84 is the link between Nup133 and Nup145C. Thus, we position the extended arm of the Y composed of Nup145C•Sec13, Nup84, and Nup133 facing outward (Fig 2.16). Excluding the Nup133•Nup84 pair, the remaining pentamer forms a roughly symmetrical triskelion that conceptually resembles the vertex elements that form polygonal cages in vesicle coats. Cryo-EM samples showed that the triskelion
measures approximately 20 nm between the tips (Lutzmann et al., 2002). An eight-membered ring of the Y-complex around the central transport channel has a ~50 nm diameter if the edges were to touch at the tips. Alternatively, the Y complexes might connect through a yet unidentified adaptor protein.

Figure 2.16 - Lattice model for the Nup84 complex and the structural scaffold of the nuclear pore complex

The ACE1 proteins Nup85, Nup145C, Nup84, Sec31, and Nic96 are colored according to Figure 2.12. (A) Schematic diagram of COPII outer coat organization. The Sec31-Sec31 cuboctahedron composed of 24 edge elements (Sec31-Sec31 heterotetramers) is shown unwrapped and laid flat in 2 dimensions. The Sec31-Sec31 crown-crown interactions make edge elements while propeller-propeller interactions are vertex elements (Fath et al., 2007). (B) Schematic diagram of the predicted lattice-like organization of the structural scaffold of the NPC. The entire scaffold (8
spokes) is illustrated unwrapped and laid flat in two dimensions. The Nup84 complex comprises
the nuclear and cytoplasmic rings, while the Nic96-containing complex makes up the inner ring.
The relative position and interactions between the seven proteins in the Nup84 complex are
shown with Sec13, Seh1, Nup133, and Nup120 colored in gray. The remainder of the complex
(Nup157/170, Nup188, and Nup192) is illustrated in gray. The illustration is not meant to predict
relative positions of proteins or structure of the inner ring per se, but shows the lattice-like
organization of the structural scaffold similar to vesicle coating complexes.

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

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

**Methods**

**Protein Production and Purification**

Full-length Seh1 and a proteolytically-mapped Nup85 fragment (residues 1-564) from *S. cerevisiae* were cloned into the bi-cistronic pCOLADuet bacterial expression plasmid (EMD Biosciences) using the BamHI/NotI and Ndel/XhoI restriction sites, respectively, resulting in N-terminally His-tagged Seh1 and untagged Nup85 and transformed into *E. coli* BL21(DE3)-RIL. Nup85_{1-564}•Seh1 is referred to as Nup85•Seh1 in the paper for simplicity. Cells were grown at 30 °C in Luria-Bertani broth supplemented with 0.4% glucose to OD_{600} = 0.8 and induced with 0.2 mM IPTG at 18°C for 18 hours. Cells were harvested by centrifugation, resuspended in 40 mM potassium phosphate pH 8.0, 500 mM NaCl, 20 mM imidazole, and 3 mM β-mercaptoethanol, and lysed using a french press. The crude lysate was centrifuged at 15,000g for 15 minutes. The soluble fraction was then incubated with 1 ml Ni-NTA per 1000 ODs for 30 minutes at 4°C and loaded onto a disposable column (Pierce). The column was washed with four bed volumes of 50 mM potassium phosphate pH 8.0, 400 mM NaCl, 30 mM imidazole, and 3 mM β-mercaptoethanol and the Nup85•Seh1 complex eluted in 4 bed volumes of 50 mM potassium phosphate pH 8.0, 250 mM NaCl, 250 mM imidazole, and 3 mM β-mercaptoethanol. Eluted protein was dialyzed against 10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.1 mM EDTA, and 1 mM DTT and the 6xHis-tag cleaved with PreScission protease. The protein was purified by anion exchange chromatography on a HiTrapQ column (GE Healthcare) via a linear NaCl gradient and twice by size exclusion.
chromatography using a Superdex S200 26/60 column (GE Healthcare) run in 10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.1 mM EDTA, and 1 mM DTT.

Selenomethionine derivatized Nup85•Seh1 was prepared by growing cells in M9 Medium (Sigma) supplemented with 1 mM MgSO₄, 6.6 µM CaCl₂, 1 ml FeSO₄ 4.2 g/L, 0.4% glucose, and 100 µl 0.5% (w/v) thiamine at 37°C (modified from (Doublié, 1997)). At OD₆₀₀ = 0.5 solid amino acid supplements were added (100 mg/ml L-lysine, L-phenylalanine, and L-threonine; 50 mg/ml L-isoleucine, L-leucine, L-valine, and L-selenomethionine). After 30 minutes, cultures were transferred to 22°C for 20 minutes and then induced with 0.2 mM IPTG for 18 hours. The selenomethionine-derivatized Nup85•Seh1 complex was purified as described for the native version. Incorporation of selenomethionine was confirmed by mass spectrometry (data not shown). Both the native protein and selenomethionine derivative were concentrated to 30 mg/ml for crystallization.

Full length Nup85 in complex with Seh1 (Nup85₁-744•Seh1, referred to as Nup85(fl)•Seh1) was cloned and purified identically to Nup85₁-564•Seh1.

Full-length Sec13 and Nup145C₁₉₀-₇₁₂ from S. cerevisiae were cloned into the bi-cistronic pET-Duet bacterial expression plasmid using the BamHI/NotI and NdeI/Xhol restriction sites, respectively, resulting in N-terminally His-tagged Sec13 and untagged Nup145C. In order to stabilize the complex, the C-terminus of Sec13 and N-terminus of Nup145C were linked with a short flexible linker to generate a single chain. This linked version behaved identically to the complex made from separate chains with the advantage of increased stability and is here referred to as Nup145C•Sec13 for simplicity. The C-terminal tail module of Nup145C was removed from the construct to generate Nup145C₁₀₉-₅₅₅•Sec13 by PCR and is referred to as Nup145CΔtail in the text. Nup145C-ELIEA•Sec13 and Nup145C(Q691G)•Sec13 were generated by PCR mutagenesis.
Nup145C•Sec13 was produced identically to Nup85•Seh1 except that gel filtration was performed in 10mM Hepes/NaOH pH 7.4, 150mM NaCl, 1mM DTT, 0.1mM EDTA.

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

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

**Crystallization**

The Nup85•Seh1 complex was crystallized in 18 % (v/v) PEG 3350, 0.2 M sodium citrate, and 0.1 M bis-tris-propane pH 8.5 by the hanging drop vapor diffusion method at 18°C. Crystals grew within 4-10 days forming rods with dimensions of 150 μm x 150 μm x 400 μm. The selenomethionine derivative crystallized in the same conditions. Native crystals were cryoprotected in reservoir solution with 15% (v/v) PEG 400. Se-Met crystals were cryoprotected in the reservoir solution supplemented with 4% (v/v) additional PEG 3350 before flash freezing in liquid nitrogen. Both the native and selenomethionine Nup85•Seh1 complex crystallized in space group P4\(^{1}\)2\(^{1}\)2 with two molecules per asymmetric unit. Crystal screening was performed at beamline X6A at National Synchrotron Light Source (NSLS) and final X-ray data was collected at the NE-CAT beamline 24ID-C at Argonne National Laboratory.
**Structure Determination**

Data reduction was carried out using HKL2000 (Otwinowski and Minor, 1997). The structure was solved with the single anomalous dispersion (SAD) technique using the SeMet derivative. The positions of 2*16 selenium sites (out of 2*20 possible) were found with the program SHELXD (Sheldrick, 2008; Adams et al., 2002) and were used for phasing. The NCS-averaged, solvent-flattened 3.7 Å experimental electron density map was of sufficient quality to trace the backbone of most of the model. The selenium positions served as markers to unambiguously assign the sequence for Nup85. Assigning the sequence of Seh1 was assisted by superimposing the structure of the homologous Sec13. The final model was refined against native data extending to 3.5 Å. Model building was carried out with Coot (Emsley and Cowtan, 2004), for refinement the PHENIX suite was used (Adams et al., 2002). Only few packing interactions exist in the crystal, resulting in a relatively high Wilson B-factor of 118 Å². Thus, B-sharpened maps were generated with CNS (Brünger et al., 1998; DeLaBarre and Brunger, 2006) and were used to assist side chain placement in the early stages of model building. NCS-restraints were applied throughout the refinement process. The final model has good stereochemistry and no Ramachandran outliers (84.1% of residues in preferred regions, 14.4% in additional allowed regions) according to Molprobity (Davis et al., 2007). All secondary structure elements of the 100 kDa heterodimer have been traced, however several loops connecting either helices in trunk and crown of Nup85 or strands in Seh1 are omitted due to poor electron density in those regions.

**Analytical Gel Filtration**

For Nup145C•Sec13 and Nup84 binding experiments, equimolar amounts of Nup145C•Sec13 (or Nup145C-ELIEA•Sec13) and Nup84 were mixed and incubated at
4 °C for 5 minutes. Similarly, equimolar amounts of Nup145CΔtail•Sec13 and Nup84Δtail (or Nup84Δtail-DSICD) were mixed and incubated at 4 °C for 5 minutes. Reactions (500 µl) were injected onto a Superdex S200 10/300 column (GE Healthcare) and run in 10 mM Hepes/NaOH pH 7.4, 150 mM NaCl, 1 mM DTT, 0.1 mM EDTA at a flow rate of 0.8 ml/min. Nup85•Seh1 complex at various concentrations (1, 5, 10, and 20 mg/ml) was loaded onto a Superdex S200 10/300 column and run in 10 mM Tris/HCl pH 8.0, 150 mM NaCl, 0.05 mM EDTA, and 0.5 mM TCEP at a flow rate of 0.8 ml/min.

In Figure 2.6A, equimolar amounts of Nup145C•Sec13 and Nup85(fl)•Seh1 were mixed and incubated at 4 °C for 5 minutes before injection on a Superdex200 HR26/60 column. In Figure 2.6B, equimolar amounts of Nup120 and Nup85•Seh1 were mixed and incubated at 4 °C for 5 minutes before injection on a Superose 6 HR10/300 column (GE Healthcare). In Figure 2.6C, Nup120•Nup85(fl)•Seh1 was mixed with a 2-fold molar excess of Nup145C•Sec13 or Nup145CΔtail•Sec13 and incubated at 4 °C for 5 minutes before injection on a Superdex S200 HR10/300 column. In Figure 2.7, Ni-NTA elutions of Nup120•Nup85(fl)•Seh1 co-expressed with Nup145C-ELIEA•Sec13 or Nup145C(Q691G)•Sec13 (where both Nup120 and the Nup145C•Sec13 variant are 6xHis-tagged) were run on a Superdex200 HR26/60 column. Gel filtration was carried out in 10 mM Hepes/NaOH pH 7.4, 150 mM NaCl, 1 mM DTT, 0.1 mM EDTA at the columns recommended flow rates.

**Analytical Ultracentrifugation**

Purified Nup85•Seh1 complex was gel-filtered in 10mM Tris/HCl pH 8.0, 150mM NaCl, 0.5mM TCEP, and 0.05 mM EDTA immediately prior to the experiments. Analytical ultracentrifugation experiments were carried out with an Optima XL-A centrifuge using An50Ti (6 hole, equilibrium runs) or An60Ti (4 hole, velocity runs) rotors.
Samples for sedimentation velocity (440 µl sample or 450 µl buffer) were loaded into Epon-charcoal filled 2 channel centerpieces, fit with sapphire windows, and spun at 42,000rpm. Concentrations of 6.3, 3.2, and 0.63 µM Nup85•Seh1 were used. Sedimentation velocity data was analyzed globally to generate a c(s) distribution using the hybrid local continuous distribution and global discrete species model in SEDPHAT (Schuck, 2000). The data was fit from 2 to 10 s⁻¹ with Sedanal calculated \( v_{\text{bar}} = 0.7319 \) cm³/g, \( \eta = 1.0182 \) cP, and \( \rho = 1.00472 \) g/cm³.

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

**Isothermal Titration Calorimetry**

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

**CD Spectroscopy**

Nup145Δtail•Sec13, Nup145Δtail-ELIEA•Sec13, Nup84Δtail, and Nup84Δtail-DSICD were purified as described above and gel filtered into 5 mM Hepes/NaOH pH 7.4 and 150 mM NaCl immediately prior to the experiment. An Aviv Model 202 CD spectrometer was used for all experiments. CD signal at 208nm of 1.3 μM protein in a 1mm pathlength cell was recorded at every degree during a 25-80°C temperature ramp with two minutes of equilibration time at each step.

**In vivo Localization Experiments**

Strains were grown in YPD (1% yeast extract, 2% yeast peptone, 2% glucose) at 24°C to an OD600 of 0.5-0.8. Cells were added to an equal volume of phosphate buffered saline (PBS, pH 7.3) with 20 mg/ml Hoechst dye (Invitrogen) for 30 minutes at room temperature to stain DNA. Cells were harvested by centrifugation, washed once in phosphate buffered saline (PBS) and viewed using a Nikon E800 fluorescent microscope (Melville, NY) mounted with a Hamamatsu digital camera (Bridgewater, NJ). Images were captured using Improvision OpenLabs 2.0 software (Lexington, MA) with identical exposure times for each sample. Final images were constructed using Adobe Photoshop CS3 and Adobe Illustrator CS3 (Adobe Systems).

**Yeast Plasmid Construction**

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

Strain Construction

Yeast strains used in this study are listed in Table 2.2. Genomic tagging of NUP133 and NUP84 in a NUP145/nup145::KANMX4 diploid (BY4743 background, Saccharomyces cerevisiae deletion consortium) was done by homologous recombination (Janke et al., 2004), 2004 using pYM28 (eGFP: HIS3MX6) as template. Strains were selected on SMM-histidine plates and verified by western blotting with anti-GFP antibody and nuclear rim localization of the Nup-GFP chimeric proteins. The resulting diploids were transformed with SBYp115 (NUP145, URA3, CEN) to allow viable colonies following sporulation and tetrad dissection. Haploid strains were then transformed with SBYp116 (NUP145, LEU2, CEN) or SBYp117 (nup145C-ELIEA, LEU2, CEN), grown in medium lacking leucine, and plated on 5-fluoroorotic acid plates (Fig 2.10).
Table 2.2 - Yeast strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>YS221</td>
<td>MATα nap145::KANMX4 NUP133-GFP; HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0 [SBYp116; (NUP145, LEU2, CEN)]</td>
</tr>
<tr>
<td>YS222</td>
<td>MATα nup145::KANMX4 NUP133-GFP; HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0 [SBYp117; (nup145C-ELIEA, LEU2, CEN)]</td>
</tr>
<tr>
<td>YS223</td>
<td>MATα nap145::KANMX4 NUP133-GFP; HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0 [SBYp116; (NUP145, LEU2, CEN)]</td>
</tr>
<tr>
<td>YS224</td>
<td>MATα nup145::KANMX4 NUP133-GFP; HIS3MX6 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0 [SBYp117; (nup145C-ELIEA, LEU2, CEN)]</td>
</tr>
<tr>
<td>YS225</td>
<td>MATα nap145::KANMX4 met17Δ0 leu2Δ0 his3Δ1 ura3Δ0 [SBYp115; (NUP145, URA3, CEN)]</td>
</tr>
</tbody>
</table>

Cell fractionation

Strains expressing Nup133-GFP or Nup84-GFP were grown to log phase (OD₆₀₀ = 0.5-0.7) in YPD at 30 °C. Twenty-five OD₆₀₀ units were harvested by filtration, washed with cold water, and collected by centrifugation. Cells were pre-treated with 100 mM Tris/HCl, pH 9.4, 0.5% 2-mercaptoethanol for 10 minutes at 30 °C and spheroplasted in S buffer (40 mM HEPES pH 7.5, 5 mM MgCl₂, 1.2 M sorbitol) containing 0.2 mg/ml Zymolyase (100T) for 1 hour at 30 °C. Spheroplasts were washed 3 times with S buffer, resuspended in 0.5 mL lysis buffer containing 20 mM HEPES, pH 7.5, 5 mM MgCl₂, 1 mM PMSF and protease inhibitor cocktail (Roche) and lysed on ice by Dounce homogenization. A portion of the total lysate was removed (T), and the remaining lysate was centrifuged at 16,000 xg for 30 minutes at 4 °C resulting in a soluble (S) and pellet (P) fraction. Equal cell equivalents were resolved by SDS-PAGE, transferred to nitrocellulose, and proteins were detected using antibody against GFP (1:20,000) or the cytosolic protein Pgk1p (1:3,000).
CHAPTER 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

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


*These authors contributed equally to this work.

N.C.L. designed and conducted all crystallographic and S. cerevisiae experiments and wrote the manuscript. S.G.B. designed and conducted binding experiments and wrote the manuscript. T.U.S. advised on the project and wrote the manuscript.
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

Introduction

The main feature that distinguishes eukaryotes from prokaryotes is the confinement of the genetic material into a membrane-enveloped nucleus. Since gene transcription and mRNA processing occur inside the nucleus while protein translation is restricted to the cytoplasm, transport across the double-layered nuclear envelope (NE) is essential for cellular homeostasis. The exchange of all molecules, including ions, proteins, and RNAs is facilitated exclusively by nuclear pore complexes (NPCs) (D'Angelo and Hetzer, 2008; Lim et al., 2008; Tran and Wente, 2006; Weis, 2003). NPCs are large protein assemblies of 40-60 MDa that are embedded in the nuclear envelope and exhibit an 8-fold rotational symmetry around a central axis in addition to an imperfect two-fold symmetry across the plane of the NE (Beck et al., 2007; Stoffler et al., 2003). Composed of multiple copies of ~30 proteins, termed nucleoporins (nups), the NPC has an outer diameter of ~100 nm while the central channel measures ~40 nm in width. Transmembrane nups directly connect the NPC to the NE, while the phenylalanine-glycine (FG) repeat-containing nups line the interior of the pore. These FG-filaments mediate nucleocytoplasmic transport of cargo molecules across the NE. FG-filament bearing nups are anchored to the NPC scaffold built from architectural nucleoporins arranged in two large multiprotein complexes that form a membrane-proximal layer. The scaffold structure is very stable and undergoes virtually no turnover in the quiescent cell (D'Angelo et al., 2009), while many other nucleoporins have variable dwell times at the NPC (Rabut et al., 2004). In consequence, the NPC is a highly modular structure (Schwartz, 2005). Understanding the structure of the NPC therefore depends upon elucidating its basic scaffold.

The two essential architectural building blocks of the NPC are the Nup84 subcomplex and the Nic96 subcomplex. The components of the Nic96 subcomplex likely include Nic96, Nup53/59, Nup157/170, Nup188 and Nup192 (yeast nomenclature), as
inferred from co-immunoprecipitations (co-IPs) (Alber et al., 2007; Hawryluk-Gara et al., 2005; Marelli et al., 1998; Onischenko et al., 2009) and yeast-two-hybrid screens (Wang et al., 2009; Yu et al., 2008). Judged by immunolabeling, the Nic96 subcomplex might form a central ring within the NPC sandwiched between peripheral rings formed by Nup84 subcomplexes (Alber et al., 2007). In comparison to the Nic96 subcomplex, the Nup84 subcomplex is substantially better understood. It has 7 universally conserved members (yeastNup84/humanNup107, yNup85/hNup75, yNup120/hNup160, Nup133, yNup145C/hNup96, Sec13, and Seh1) and three additional members (Nup37, Nup43, and ELYS/Mel-28) to date found mainly in metazoan (Cronshaw et al., 2002; Gillespie et al., 2007; Loiodice et al., 2004; Rasala et al., 2006). In the fungus Aspergillus nidulans, distant Nup37 and ELYS orthologs have been described recently (Liu et al., 2009). The heptameric core Nup84 complex assembles tightly as shown by co-IPs and in vitro assembly (Harel et al., 2003; Lutzmann et al., 2002; Siniossoglou et al., 2000; Walther et al., 2003). Negatively-stained electronmicrographs of the assembled Nup84 complex reveal a branched Y-shaped structure, with two short arms and a kinked stalk connected at a central hub (Lutzmann et al., 2002).

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

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

Results

Structure Determination

After systematic C-terminal truncation, a stable fragment comprising most of Nup120 (residues 1-757 of 1037 total) from S. cerevisiae was recombinantly expressed in E. coli and purified. The protein is a monomer in solution (data not shown). Native protein readily crystallized and selenomethionine derivatized crystals were obtained after microseeding with native crystals. Though both crystal forms were optically identical, the selenomethionine crystals diffracted better and were used exclusively in structural analysis. The structure of Nup120 was solved with one molecule per asymmetric unit by single-wavelength anomalous dispersion (SAD) on very strong Se-Peak data (all 9 Se sites are well ordered). The model is complete except for 27 residues at the C-terminus and 7 flexible loops (out of 43 total loops) and was refined to $R_{\text{work}} / R_{\text{free}}$ of 24.4 % / 29.9 % (Table 3.1).
Table 3.1 - Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>Nup120.757</th>
<th>Selenomethionine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2;2;2</td>
<td></td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>114.6, 153.7, 53.0</td>
<td></td>
</tr>
<tr>
<td>α, β, γ (*)</td>
<td>90, 90, 90</td>
<td></td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>50.0 - 3.0</td>
<td></td>
</tr>
<tr>
<td><strong>Rsym (%)</strong></td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td><strong>Rmerge / R_m</strong></td>
<td>20.4 (1.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>99.8 (99.6)</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.0 (3.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.0 - 3.0</td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>35841</td>
<td></td>
</tr>
<tr>
<td>R_m / R_free</td>
<td>24.4 / 29.9</td>
<td></td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>5305</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>B-factors (Å²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-Propeller</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>α-helical insertion</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>α-helical domain</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td><strong>Ramachandran plot (%)</strong></td>
<td>93.6 / 5.3 / 1.1</td>
<td></td>
</tr>
<tr>
<td>RMSD bond lengths (Å)</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>RMSD bond angles (*)</td>
<td>1.915</td>
<td></td>
</tr>
</tbody>
</table>

*Values in parentheses are for highest-resolution shell.

**Crystal Structure of Nup120**

Nup120 folds into a continuous, prolate disk with overall dimensions of 90 Å x 55 Å x 35 Å. One half of the structure is formed by an N-terminal β-propeller domain that is intimately connected to a compact central domain built from two closely packed α-helical segments (Fig 3.1). Overall, the structure is better resolved in the α-helical segment than the β-propeller, likely a result of a paucity of packing contacts involving the latter. The β-propeller of Nup120 contains 7 consecutive blades that fan out from a central axis. The blades are formed by a β-sheet of 4 consecutive antiparallel strands, labeled A-D. Blade 7 is built from the very N-terminus of the polypeptide chain forming strand 7D and joining strands 7A-C to close the propeller in a velcro-like closure commonly observed in β-propeller domains (Chaudhuri et al., 2008). Blade 1 is 5-stranded, with
strand 7D extending to form the additional strand 1E before connecting to strand 1A (Fig 3.1E). Blade 3 is somewhat unusual in that the outermost strand 3D is only loosely connected to strand 3C with a hydrogen-bonding network hardly visible in our structure and the sequence could only be tentatively assigned for strand 3D residues 204-216.

**Figure 3.1 - Overall topology of Nup120**

(A) Current model of the Y-shaped Nup84 subcomplex. The relative position of Nup120 is highlighted. (B) Schematic of full-length Nup120 from *S. cerevisiae*. Residues that form the β-propeller are colored blue, those that form the α-helical domain are purple, and those not present in the crystallized construct are in gray. (C, D) The overall topology of Nup120 (residues 1-757 of 1037) is shown in two views rotated by 90°. The structure is gradient-colored from blue to white to magenta from N- to C-terminus. At its N-terminus, Nup120 forms a 7-bladed β-propeller. A 4-helix bundle (α1-α4) between blades 6 and 7 packs against the remainder of the helical domain (α5-α15), composed of helices wrapping around a central hydrophobic stalk of the two long helices α11 and α12. Unstructured loops absent from the final model are shown in gray. (E) A
topological diagram of the Nup120 structure is shown, illustrating the 4-helix insertion between blades 6 and 7 of the propeller as well as the two central helices of the helical domain.

The α-helical domain that forms the second half of the molecule is constructed in a unique discontinuous manner. In total the domain contains 15 helices, labeled α1-α15. The first 4 helices form a compact bundle and are inserted between blades 6 and 7 of the β-propeller. The remaining 11 helices are C-terminal to the β-propeller and pack tightly against the 4-helix bundle to form one compact entity. The arrangement of the helices is highly irregular. The most prominent feature of the domain are two long helices, α11 and α12, which pack against each other and form a central stalk, defining the long axis of the domain. Helices α5-α9 wrap up and around this element, with helices α6/α7 and α8/α9 arranged in two stacked braces oriented perpendicular to the stalk. Helices α1, and α13-α15 meander back down and around the other side to bury most of the hydrophobic stalk. The remaining surface area of the two central helices is closed by the 4-helix insertion bundle. Taken as a whole, the structure of Nup120,757 consists of a bipartite helical domain that is interrupted by a β-propeller.

**The Main Crystal Contact is Formed by a Domain Swap**

Other than a collection of spurious small contacts crystal packing is mainly achieved by a domain swap of the terminal helices α15 and α15' exchanging between two neighboring molecules (Fig 3.2A). The interface measures 1355 Å², is entirely hydrophobic and highly complementary (Fig 3.2B). Domain swaps are regularly found in crystals (Liu and Eisenberg, 2002) and, as stated above, we do not observe dimerization of Nup120 in solution. We cannot rule out the possibility that the interface is physiologically relevant; sterically the domain swap is conceivable in the context of the entire molecule including the C-terminal 280 residues omitted in our construct. It is
however more likely that the exposed hydrophobic patch is artificially generated by the truncation of the domain, since we also do not observe particularly high sequence conservation within helix α15. We speculate that in vivo the patch likely accommodates one of the additional helices from the C-terminal domain, or alternatively, is involved in interaction with a neighboring molecule. Whether the C-terminal domain is rigidly or flexibly tethered to Nup120-1.757 is an open question.

Figure 3.2 - Crystal contacts between two symmetry-related molecules

(A) One molecule of Nup120 in blue, and its symmetry mate, related by a 2-fold rotation, in orange. The β-propellers are at opposite ends while the helical domains engage in a putative domain swap between helices α15 of both molecules. (B) Close-up of the domain-swapped region, illustrating the hydrophobic nature of helix α15 and the surrounding pocket (hydrophobic residues are shown in white). In monomeric Nup120, helix α15 likely folds under (arrow) and occupies the position taken by helix α15' (orange) of the symmetry-related molecule in the crystal.

Conservation of Nup120 and Comparison to the Human Ortholog Nup160

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

We analyzed the charge distribution on the surface of Nup120 (Fig 3.3B). Since Nup120 is part of the scaffold structure of the NPC, we asked whether it may be possible that it directly juxtaposes the pore membrane. This would also be consistent with a membrane-curvature sensing ALPS motif, predicted in helix α5-6 of Nup120 (Drin et al., 2007). The surface charge of Nup120 1-757, however, is fairly mixed without conserved positive patches that might suggest direct membrane interaction. The ALPS motif is embedded in the structure and it is rather unlikely that it would swing out and insert in the membrane. Thus we suggest that Nup120 1-757 does not directly touch the nuclear membrane.

Structure-guided sequence comparison of Nup120 and its human ortholog Nup160 strongly suggests that both proteins adopt the same unique fold despite a low sequence identity of ~10%. Both non-canonical characteristics of Nup120 (the helical insertion between blades 6 and 7 of the N-terminal β-propeller and the long central stalk helices forming the hydrophobic core of the central domain) are clearly conserved in Nup160. The 279 additional residues of Nup160 are dispersed over several regions and mostly correspond to different loop lengths connecting α-helices and β-strands. Of note, the C-terminal domain of Nup160, which is not present in the Nup120 crystal structure described here, has 5 additional predicted helices, possibly indicating a vertebrate-specific extension. Despite these differences, the Nup120 crystal structure is likely
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

generally representative of all Nup120/Nup160 orthologs.

Figure 3. 3 - Surface conservation and electrostatics of Nup120

(A) Surface conservation of Nup120 is shown from three different views. To illustrate the conservation of residues on the surface of Nup120, a multiple sequence alignment sampling the phylogenetic tree of budding yeasts was generated and mapped onto the surface, colored from white (not conserved) to orange (highly conserved). The view in the middle panel corresponds to the view shown in Figure 1C. A patch of highly conserved residues is apparent on the outer face of the propeller domain of Nup120. (B) The electrostatic surface potential of Nup120 is shown in the same views as in (A) and is colored from red (-10 kT/e) to blue (+10 kT/e).

The C terminus of Nup120 Directly Binds Nup145C and Nup85

We sought to map the interaction of Nup120 with its direct binding partners in the Y-complex, Nup145C and Nup85. In a gel filtration assay, we tested for the formation of
a pentameric Sec13•Nup145C•Nup120•Nup85•Seh1 complex (Fig 3.4). Incubating Nup120\textsubscript{766-1037} or Nup120\textsubscript{1-757} with both Nup145C•Sec13 and Nup85•Seh1 resulted in complex formation only for the C-terminal Nup120 domain, but not for the crystal construct. In combination with previous interaction mapping experiments (Brohawn et al., 2008), we conclude that the helical tails of the ACE1 domains of both Nup145C and Nup85 each interact directly with the helical Nup120\textsubscript{766-1037}. This positions the C-terminus of Nup120 at the center of the hub of the Y-complex.
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

Figure 3.4 - The C-terminus of Nup120 is necessary for binding Nup85·Seh1 and Nup145C·Sec13

(A) Nup85·Seh1 (red), Nup145C·Sec13 (gray), and Nup1201-757 (green) were run individually and in combination (blue) on a Superdex S200 10/300 gel filtration column. (B) Nup85·Seh1, Nup145C·Sec13, and Nup120766-1037 were incubated together and run on Superdex S200 26/60 and eluted in a single peak. (C) Fractions from the gel filtration experiment in B were analyzed by SDS-Page. Co-migration of Nup85·Seh1, Nup145C·Sec13, and Nup120766-1037 indicates that the C-terminus of Nup120 is necessary for the formation of the pentameric complex that comprises the hub of the Y-shaped complex.

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

Having established that Nup120766-1037 is sufficient to bind both Nup145C·Sec13
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

and Nup85•Seh1 in vitro, we sought to examine the integration determinants of Nup120 into the NPC in vivo. NUP120 is not essential in yeast but nup120Δ cells exhibit a pore clustering phenotype (Aitchison et al., 1995; Heath et al., 1995) that is reminiscent of but less severe than the pore clustering observed for other scaffold nucleoporins including Nup84 and Nup133 (Li et al., 1995; Pemberton et al., 1995; Siniossoglou et al., 1996).

We genomically GFP-tagged full length Nup120 and replaced the C-terminal 280 residues of genomic Nup120 with an in frame GFP-tag to create strains expressing Nup120-GFP or Nup120\textsubscript{1-757}-GFP in a BY4741 background and examined the localization of the proteins via immunofluorescence (Fig 3.5). Nup120-GFP properly localizes to the NPC and shows typical nuclear rim staining, superimposing well with mAb414-staining of FG-nups (Aris and Blobel, 1989). Nup120\textsubscript{1-757}-GFP, on the other hand, does not properly localize to the nuclear envelope and shows staining throughout the cell. This result is consistent with our in vitro data and suggests that the integration into the Y-complex is important for proper localization of Nup120.
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

**Figure 3.5 - Nup120\textsubscript{1-757} does not localize to the nuclear envelope**

(Aa-Ad) Nup120-GFP is targeted to the nuclear envelope, as confirmed by co-localization with mAb414 (staining FG-nups), while Nup120\textsubscript{1-757}-GFP (Ba-Bd) is distributed throughout the cell. Mislocalization indicates that the C-terminus of Nup120 is necessary for proper recruitment to the NPC. Nup120\textsubscript{Δ}, nup133\textsubscript{Δ}, and nup84\textsubscript{Δ} cells (in the same BY4741 strain background) are shown for comparison (Cb-Cd, Db-Dd, Eb-Ed). Nuclear rim was visualized using mAb414, GFP-tagged Nup120 using goat α-GFP, and DNA using DAPI. Merged images are shown on the right.

*Nup120 is Topologically Different from other Scaffold Nucleoporins*

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

The structural subunits of the NPC were initially predicted to be composed of simple combinations of regular $\alpha$-helical solenoids and $\beta$-propellers (Devos et al., 2006). Experimental data now allows to specify these broad classifications, which should help to more specifically address the ancestry of the NPC. Both Sec13 and Seh1 form open, 6-bladed propellers that are completed in trans by the N-terminal insertion blades of their binding partners Nup145C and Nup85, respectively. Furthermore, helical nucleoporins Nic96, Nup145C, Nup85, and Nup84 are built around a common and distinct ancestral coatamer element (ACE1) shared with Sec31 of the outer coat of COPII vesicles (Brohawn et al., 2008). In ACE1 proteins, a specific N-terminal elaboration is followed by a tripartite helical domain composed of a trunk, a crown and a tail element. The ~30 helices within ACE1 follow a J-like pattern, zig-zagging up on one side of the trunk, making a U-turn within the crown domain, and then following down on the opposite side of the trunk (Fig 3.6A, right panel). The tail domain is often attached with modest flexibility to the trunk and is missing in most crystal constructs. In the case of Nup145C and Nup85 the N-terminal elaborations are the aforementioned insertion blades that bind to Sec13 and Seh1. Nup145C•Sec13 and Nup85•Seh1 heterodimers form the two proximal segments of the Y-shaped complex and are tethered together by Nup120 (Brohawn et al., 2008).
Based on structure predictions and its overall size, it was reasonable to suggest that Nup120 may take on a structure similar to Nup145C*Sec13 and Nup85*Seh1, with the only major difference being that the β-propeller and the α-helical domains are fused into one polypeptide chain. However, comparison between the structure of Nup120 and the Nup85*Seh1 heterodimer reveals a marked difference in topology (Fig 3.6A). Whereas the ACE1 architecture of Nup85 forms an elongated α-helical domain, the central α-helical domain of Nup120 is nearly as wide as it is long, forming an almost globular structure. The ACE1 trunk module covers the bottom face of the Seh1 β-propeller, while in Nup120 the helical domain is attached to and integrated into an edge of the β-propeller. Further, the ACE1•β-propeller interaction is accomplished by the addition of an insertion blade N-terminal to ACE1, while in Nup120 the β-propeller domain inserts a 4-helix bundle into the central α-helical domain. This helical insertion fits snugly into a pocket formed by helices α5-α7 and α11-α13 and creating an interface of nearly 600 Å² (Fig 3.6B).

The extensive interaction between the β-propeller and α-helical domain of Nup120 creates a large, rigid interface of 2175 Å². In contrast, the largest contact area between ACE1 and its β-propeller partner is at the insertion blade/β-propeller interface. Additional contact areas in ACE1•β-propeller complexes are smaller in comparison to the corresponding interfaces in Nup120 and, importantly, far less hydrophobic. Thus, for the ACE1•β-propeller assembly one has to consider substantial flexibility about the interaction joint, while the Nup120 structure presented here is very likely inflexible. Not only does the structure of Nup120 significantly differ from the ACE1•β-propeller heterodimers, but additional emerging evidence suggests that it also lacks similarity to Nup170 and Nup133, the two other scaffolding nucleoporins of similar size and domain composition with an N-terminal β-propeller followed by an α-helical domain (Whittle and
Figure 3.6 - Nup120 is composed of a combined β-propeller - α-helical domain distinct from ACE1-β-propeller

(A) The overall architectures of Nup120 and the ACE1 motif of Nup85·Seh1 are distinctly different. Nup120 is characterized by a bipartite helical domain (blue to white from N- to C-terminus) that is interrupted by a β-propeller domain (gray). The Nup85 ACE1 motif is characterized by an elongated helical stack (colored blue to white from N- to C-terminus) that makes a U-turn in the crown domain of the molecule. At its N-terminus, Nup85 inserts a blade (in red) into the open, 6-bladed Seh1 β-propeller. In contrast, the β-propeller of Nup120 contributes a helical insertion bundle (red) to the helical domain. The view of Nup120 is the same as that in Fig. 3B. (B) Surface representations of intact Nup120 are shown on the left, while on the right the three modules of Nup120 – the propeller, the helical insertion, and the helical domain – are shown pulled apart to illustrate the buried surface areas in between. Interacting surfaces between the propeller and the insertion bundle are outlined in green, between the propeller and the helical domain in yellow, and between the insertion bundle and the helical domain in orange. The molecule is N-to-C gradient-colored from blue-to-white-to-magenta.
Discussion

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

\textit{Nup120 in the Context of the NPC Scaffold}

Nup120 forms one of the two short arms of the universally conserved, 0.6 MDa Y-complex, the essential building block of the NPC scaffold. The assembly of the Y-complex from its 7 members is fairly well understood and has been studied using many different techniques. All of these studies profit from generally very high affinities observed between the interacting proteins within the Y, which generated largely consistent co-immunoprecipitation and yeast two-hybrid results and facilitated the crystallization of several complex crystal structures. While there is general agreement on the overall topology of the NPC, as determined by electron microscopic techniques, different models for the assembly of the NPC structural scaffold and the integration of the Y-complex are being discussed, as more detailed information is becoming available.

Based on a combination of computational, structural, biochemical, and \textit{in vivo} experiments, a model was proposed where the Y-complex is positioned in two 8-membered rings located at the periphery of the NPC sandwiching two equally wide rings
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

composed of Nup157/170, Nup188 and Nup192 in between (Alber et al., 2007). One exiting aspect of the combinatorial approach is, that going forward to a higher resolution it may allow the integration of crystallographic data as well, in which case it could come close to a detailed molecular description of the NPC.

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

In contrast to the concentric cylinder model, we proposed a lattice-like model for the NPC, extrapolated from the assembly of COPII vesicle coats and substantiated by the structure and assembly principles of core components of the NPC scaffold (Brohawn et al., 2008; Brohawn and Schwartz, 2009a). We propose the Y-complex does not directly coat the pore membrane (in analogy to the COPII outer coat), but is anchored by another set of proteins, likely involving the essential transmembrane nucleoporin Ndc1 and/or its direct binding partners (Onischenko et al., 2009). It is of interest to discuss this issue in respect to the membrane-inserting ALPS motif that was experimentally characterized within a loop structure in hNup133-NTD and that was predicted to occur as well in yNup85 and yNup120 (Drin et al., 2007). Based on the structural data now available on both yNup85 and yNup120, it appears unlikely that the predicted ALPS motif in both proteins is functional in membrane-binding since neither is in an exposed region of the protein, or is likely to become exposed. This is in contrast to the ALPS motif
in hNup133-NTD, where it is well exposed in the crystal structure, and also highly conserved in metazoan Nup133 (Berke et al., 2004). Taking all the available data together, it appears more reasonable to suggest a specific function for the ALPS motif in metazoan Nup133 rather than a general function in anchoring of the NPC to the pore membrane. Since Nup133-ALPS is only poorly conserved in yeast, it is tempting to speculate that it may have a specific role in NPC assembly in open mitosis (Güttinger et al., 2009).

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

The fact that inter-subunit interactions are still obscure suggests that these interactions are rather weak and hard to establish. Each short arm of the Y-complex contains one \( \beta \)-propeller domain, while the stalk contains two (Fig 3.1). For the assembly of the extensions of the Y, direct interactions between the \( \alpha \)-helical domains is essential, however this does not exclude the participation of the \( \beta \)-propellers. It is reasonable to suggest, that the \( \beta \)-propellers are prime candidates for the elusive inter-subcomplex contacts. The vertices of the outer coat of COPII vesicles are assembled exclusively via \( \beta \)-propeller interactions, which have still only been inferred by fitting crystal structures into EM maps (Fath et al., 2007; Stagg et al., 2008). \( \beta \)-propellers make excellent protein-protein interfaces due to their inherent ability to pair with a binding partner in multiple modes. Binding to peptides via the face of the \( \beta \)-propeller is well known (Jawad and Paoli, 2002). Additionally, each blade exposes on its edge (typically on strand D) a stretch of \( \sim 6-8 \) residues available for intermolecular \( \beta \)-sheet formation, which can be likened to one half of a zipper. In Nup120, 5 of the 7 blades are exposed this way, two are buried in the hydrophobic core shared with the attached \( \alpha \)-helical domain. In addition
to these interactions being relatively weak, another inherent difficulty in identifying them is that they are likely very poorly conserved at the sequence level because the contacts are mediated via the backbone rather than side chains. Based on the available data, it is conceivable that the Nup120 β-propeller is involved in inter-Y contacts. It is also possible that it is used to bridge to the Nic96 complex, but we can also not exclude that it may be an anchor for dynamic nucleoporins or other accessory proteins. The relatively mild nup120Δ phenotype (Fig 3.5) compared to nup133Δ or nup84Δ and the behavior of Nup1201-757-GFP suggests that if the Nup120 β-propeller has an integral role in the NPC scaffold, it is either redundant or can be functionally replaced by another nucleoporin.

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

**Experimental Procedures**

*Protein Expression and Purification*

Nup120 from *S. cerevisiae* (residues 1-757 of 1037) was expressed at 18°C in *E. coli* strain BL21(DE3)-RIL as a 6xHis N-terminal fusion protein from a pET-Duet-derived plasmid. Cells were pelleted and resuspended in lysis buffer (50mM potassium phosphate pH 8.0, 500mM NaCl, 40mM imidazole, 5mM β-mercaptoethanol). Cells were lysed using a french press and the clear lysate incubated in batch with Ni-affinity resin. After washing the resin in batch with lysis buffer, the protein was eluted with lysis buffer containing 250mM imidazole. After cleavage of the purification tag, Nup120 was subjected to size exclusion chromatography on Superdex S200 equilibrated in 10mM
Tris/HCl pH 8.0, 150mM NaCl, 0.1mM EDTA, and 1mM DTT. Nup120 eluted as a monomer of 88 kDa. Selenomethionine-derivatized protein was prepared as previously described (Brohawn et al., 2008) and Nup120-SeMet was purified identically to the native version.

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

**Protein Crystallization**

Nup120 concentrated to 20 mg/ml was crystallized in 15% (w/v) PEG 3350, and 0.1M Tris/HCl pH 7.5, 0.2M KSCN by the hanging drop vapor diffusion method at 18°C in 2μl drops. Crystals grew within 3-6 days forming rhomboid prisms with dimensions of 60μm x 60μm x 20μm. The selenomethionine derivative crystallized in the same condition, while the highest quality crystals were obtained by microseeding with native crystals. Both native and derivative crystals were cryo-protected by serial transfer of the crystals into reservoir solutions supplemented with increasing amounts of PEG200
(10%-25% (v/v), 5% steps) before flash freezing in liquid nitrogen. Both native and
derivative protein crystallized in space group P2\textsubscript{1}2\textsubscript{1}2\textsubscript{1} with one molecule per asymmetric
unit. Data was collected at beamline 24ID-C at Argonne National Laboratory.

**Structure Determination**

Although the native crystals were larger and optically superior, the
selenomethionine-derivatized crystals diffracted significantly better and were exclusively
used for data analysis. A complete dataset was collected at the Se-Peak wavelength
and data reduction was carried out using the HKL2000 package (Otwinowski and Minor,
1997). All 9 selenium sites were found using SHELXD (Sheldrick, 2008). After
refinement of the Se positions and density modification with SHARP, an adequate
experimental electron density map was obtained, allowing for the assignment and
building of the majority of the structure. Sequence assignment was aided by using the
selenium positions as markers. Model building was done with Coot (Emsley and
Cowtan, 2004) and refinement was carried out using the PHENIX suite (Adams et al.,
2002). The model is complete except for residues 31-52, 188-200, 303-313, and 731-
757 for which only spurious electron density was observed. Blades 3 and 4 of the β-
propeller have the highest temperature factors and are not as well packed as the
remainder of the molecule. Sequence assignment in this region, particularly in strand 3D
residues 204-216, is tentative.

**Analytical Size Exclusion Chromatography**

For Nup120 (residues 1-757), Nup145C•Sec13, and Nup85•Seh1 binding
experiments, equimolar amounts of each component were incubated alone or in
combination for 30 minutes at 4°C in binding buffer (10mM Tris/HCl pH 8.0, 250mM
NaCl, 1mM DTT, 0.1mM EDTA). Reactions were injected onto a Superdex S200 hr10/300 column (GE Healthcare) equilibrated in binding buffer, and run at a flow rate of 0.8 ml/min (Fig 3.4A).

**Yeast Strain Construction**

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

**Fluorescence Microscopy**

Strains were grown overnight in YPD (1% yeast extract, 2% yeast peptone, 2% glucose) at 30°C, diluted 20-fold into fresh YPD, and grown for 4-5 hours at 30°C to OD₆₀₀ ~0.5. Cells were harvested by centrifugation, fixed for 3 minutes in 3.7% formaldehyde/0.1M potassium phosphate pH 6.5, and prepared for immunofluorescence as previously described (Kilmartin and Adams, 1984). Samples were incubated with mAb414 (abcam, 1:1000) alone or in combination with goat anti-GFP (1:500) for 90 minutes at room temperature. Bound antibodies were detected by incubation with Cy5-conjugated anti-mouse (Jackson Labs, 1:500) alone or in combination with Cy2-conjugated donkey anti-goat (Jackson Labs, 1:200) for 45 minutes at room temperature. DNA was stained with 0.05 µg/ml 4',6-diamidino-2-phenylindole (Sigma-Aldrich) and samples were mounted for imaging in 1 mg/ml p-phenylenediamine and 90% glycerol. Fluorescence microscopy was performed on a Zeiss AxioImager.Z1 microscope and images were taken with a Zeiss AxioCam HRm camera.
Chapter 3: The structure of the scaffold nucleoporin Nup120 reveals a new and unexpected domain architecture

ACCESSION NUMBER

The atomic coordinates of the Nup120 structure have been deposited in the Protein Data Bank (PDB) with accession number 3HXR.

ACKNOWLEDGEMENTS

We thank staff at the NE-CAT beamline 24 at Argonne National Laboratory for excellent assistance with data collection; Iain Cheeseman for providing antibodies; Eric Spear for advice on in vivo experiments; Dennis Kim for use of his fluorescence microscope. Supported by NIH Grant GM77537 (T.U.S.), a Pew Scholar Award (T.U.S.), a Koch Fellowship Award (S.G.B.) and a Vertex Scholarship (S.G.B.).
CHAPTER 4: Conclusion

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


N.C.L. and T.U.S. wrote the manuscript.
Summary

The body of work presented in this thesis recapitulates the path taken to understanding the assembly and structure of the Nup84 subcomplex of the NPC, and covers a broad spectrum of biochemical, structural, and in vivo techniques. The approach taken to study this complex can now be applied to the remaining subcomplexes of the NPC, in particular the Nic96 complex. The past few years have been productive ones, not only for our lab, but also for the entire field in general. From a flurry of crystal structures, to continually improving EM reconstructions and emerging interaction networks from a combination of in vitro and in vivo methods, our knowledge about NPC structure and assembly is rapidly expanding. Here, I will discuss recent advances as well as some of the most intriguing questions that currently remain unanswered.

"Y" it needs more work

To date, at least a portion of each of the 7 constituent proteins of the Nup84 complex from S. cerevisiae have been structurally characterized, covering over 500 kDa of the ~575 kDa Y-shaped complex (Brohawn et al., 2009). Still lacking, and perhaps of greatest interest, is the hub region where the three arms of the Y (consisting of Nup85, Nup145C, and Nup120) meet. In vitro binding assays have shown that each of these three nups can bind to one another in a binary fashion (Brohawn et al., 2008; Leksa et al., 2009). Not only is a structure of the composite hub not available, but the C-terminal ~180-250 residues of each of the individual proteins have not been solved either. EM data suggest that there is an inherent flexibility within the assembled Y-shaped complex (Kampmann and Blobel, 2009). In addition, limited proteolysis experiments suggest flexibility within each of the individual proteins, as limited proteolytic cleavage generates
smaller stable domains, corresponding to the ACE1 tail domain for Nup85 and Nup145C, and a \(~250\) residue C-terminal domain of Nup120 (unpublished data). Such flexibility is known to be detrimental to crystal formation, and explains why solving the structures of the full-length proteins poses additional challenges. A crystal structure of the full hub, which may include additional stabilizing contributions from the Sec13 and Seh1 \(\beta\)-propellers, would be the most telling. However, such a large and extended complex is even less amenable to crystallization. To date, docking crystal structures into the single particle EM reconstruction (Kampmann and Blobel, 2009) of the Y-shaped Nup84 complex has provided the greatest insight into the overall structure of this complex.

**Building the Nic96 complex**

While the Nup84 Y-shaped complex is currently the best understood, the Nic96 complex is an integral component of the NPC scaffold as well. It is composed of Nic96, Nup53/59, Nup157/Nup170, Nup188, and Nup192 (Marelli et al., 1998; Hawryluk-Gara et al., 2005; Amlacher et al., 2011). So far, structures of individual proteins have been solved, including full length Nup35 from mouse (\(y\)Nup53) (Handa et al., 2006), as well as domains of Nic96 and Nup170 from yeast (Jeudy and Schwartz, 2007; Schrader et al., 2008; Whittle and Schwartz, 2009). Nup53 forms a dimer in solution, and has been shown to interact with Nic96 (Fahrenkrog et al., 2000). In addition, pull-down experiments have shown that Nup188 and Nup192 bind to Nic96 competitively (Theerthagiri et al., 2010; Amlacher et al., 2011). Recently, additional interactions have been pinned down using nups from a thermophilic yeast (Amlacher et al., 2011). However, the interactions within the Nic96 complex appear to be weaker and potentially more flexible than those of the Y-shaped complex, as none of the interactions have been recapitulated using more stringent means, such as gel filtration, or in crystal structures.
If a stable Nic96 complex could be formed, whether from *S. cerevisiae* or any other species, it would be very interesting to see whether it assumes an equally extended structure as is observed with the Y-shaped complex. Also, questions of stoichiometry and different pools of the Nic96 complex still exist, given that Nup53 forms a dimer and binds Nic96, and Nup188 and Nup192 appear to bind Nic96 in a competitive fashion (Handa et al., 2006; Theerthagiri et al., 2010).

**Models of NPC assembly**

In addition to understanding each of the individual subcomplexes in molecular detail, it is equally important to build an understanding of their arrangement within the context of the assembled NPC. Currently, three different assembly models exist. The first is a computational model based on general distance restraints derived from vast amounts of biochemical and proteomic data to generate a 3-dimensional model (Alber et al., 2007). As discussed previously, while the positioning of complexes relative to one another is plausible, the arrangement of proteins within individual subcomplexes is incorrect. The biggest debate between assembly models, though, is between the head-to-tail model proposed by the Blobel group, and the lattice model, proposed by our lab.

The head-to-tail model is characterized by four stacked rings, each containing 8 horizontal Y-shaped complexes arranged in a head-to-tail fashion around the central axis of the pore (Fig 1.7). In this model, four Nup145C·Sec13 dimers and four Nup85·Seh1 dimers form vertical poles, spanning the height of the NPC. The initial driving force behind this model was the crystal contacts observed between two Nup145C·Sec13 dimers, mediated through the crown domain of Nup145C (Hsia et al., 2007). Aside from this crystal contact, there is no further supporting evidence for this interaction. Instead, we were able to show that the suggested dimerization interface on Nup145C instead serves as a binding site for Nup84 within the context of the Y-shaped
complex (Fig 2.5) (Brohawn et al., 2008; Brohawn and Schwartz, 2009b). The
dimerization of Nup145C in the crystal stems from an exposed hydrophobic patch on the
crown of Nup145C that, in the concentrated crystallization environment, forms a binding
interface that is nothing more than a crystal artifact.

Additional experimental evidence for the head-to-tail model includes binding
between Nup120 and an N-terminal peptide of Nup133 (Seo et al., 2009), as well as
polarized fluorescence microscopy experiments (Kampmann et al., 2011), which position
the Y-shaped complex in a similar fashion. While the Nup120-Nup133 binding data,
representative of a head-to-tail interaction between two Y-shaped complexes, cannot be
directly refuted, it can also be explained by other models. Additionally, in the polarized
fluorescence microscopy experiments, the authors fluorescently tag a minimal complex
between Nup84 and Nup133 (corresponding to an existing crystal structure) and
determine its orientation within the context of the NPC, and assume it to be
representative of the orientation of the entire Y-shaped complex. There are a few
problems with this assumption. First, the tagged construct used for the fluorescence
experiments constitutes less than 20% of the length of the entire Y-shaped complex.
Additionally, it cannot be unequivocally placed within the EM structure of the Y since the
resolution in this region is not of high enough quality for accurate docking of the crystal
structure (Fig 1.6). Thus, the accuracy with which the minimal complex represents the
orientation of the entire Y is dubious. Secondly, this particular region of the Y-shaped
complex is one of the most flexible, and exhibits large, distinct hinge movements as
evidenced by EM, again calling into question whether the minimal complex can truly be
used to assume the orientation of the entire Y-shaped complex. Taken together, this
evidence is somewhat questionable, and cannot be taken at face value.

In contrast, our lab has proposed a lattice-like model of the NPC, based on the
analogy to the COPII outer vesicle coat. The most distinguishing feature between this
model and the head-to-tail model is the orientation of the Y-shaped complex. While the head-to-tail model places the Y's horizontally around the central axis of the pore, our model positions them vertically, with the long arms of the Y pointing outwards toward the nucleo- and cytoplasm (Fig 1.7). The supporting experimental evidence from our lab has been discussed in the previous chapters. More recently, ultra high-resolution microscopy experiments by another group provided data consistent with the vertical orientation of Y complexes around the central axis of the pore (unpublished data, ASCB meeting 2011), supporting our lattice-like model of NPC assembly.

**Lattice model of the NPC and of COPII coats: similarities and differences**

Upon discovery of ACE1, we proposed a model of the NPC scaffold based on and analogous to the lattice-like assembly of COPII vesicle coats (Fig 4.1). COPII vesicle coats are composed of two layers. The membrane-proximal or inner layer is composed of the three proteins Sar1, Sec23, and Sec24, while the membrane-distal or outer layer can be exclusively constructed from the ACE1-protein Sec31 and the β-propeller protein Sec13 (Russell and Stagg, 2010). In the COPII outer coat, one assembly unit is composed of two Sec31-Sec13 heterodimers that interact in a homotypic fashion via the crowns of Sec31’s ACE1 domain, positioning the N-terminal β-propeller of Sec31 and the adjacent β-propeller of Sec13 as a tandem pair at either end of this extended rod-shaped structure (Fig 4.1) (Fath et al., 2007). Thus, ACE1 domains form the symmetrical edge elements of the lattice, while four Sec31-β-propellers from adjacent edges form the vertex elements in the assembled COPII coat. Beyond spacing the N-terminal Sec31 β-propeller and ACE1-domain, the role of the Sec13 β-propeller is still largely unclear. In our lattice model of the NPC, we predicted that the ACE1 domains of Nup84 and Nup145C form a heterotypic edge element analogous to the homotypic
edge element of the COPII coat, indicating that COPII and the NPC share not only building blocks, but also assemble at least in part by similar mechanisms (Brohawn et al., 2008). We were able to provide the evidence for this edge element in the Nup84•Nup145C•Sec13 crystal structure (Brohawn and Schwartz, 2009b). While Sec31•Sec13 is the principal building block of the COPII outer coat, the Y-complex is equally essential for the construction of the NPC coat. Biochemical and structural studies have characterized the interactions between nucleoporins within the Y-complex (Brohawn et al., 2009). Interactions between the stacked helical domains of the components Nup133, Nup84, Nup145C, Nup120, and Nup85 appear sufficient to assemble its branched Y-shaped structure. The β-propeller domains, of which there are four (N-terminal domains of Nup133 and Nup120, Seh1, and Sec13), are not central to these interactions, leaving them as candidates for mediating interactions outside of the Y.
Chapter 4: Conclusion

Figure 4.1 - Membrane coating scaffolds of the NPC and COPII vesicle coats

Schematics of the lattice-like assemblies of the NPC and COPII scaffolds are shown, with ACE1 modules colored as in Figure 1. Enlarged views of the ACE1 edge elements are shown for both the NPC and COPII, in addition to the vertex element of the COPII lattice (with Sec31 shown in green and Sec13 shown in gray). The assembly of the vertex element of the NPC, including which β-propellers are used, is currently unknown. While the COPII ACE1 interaction is homotypic, the analogous ACE1 interaction between Nup145C and Nup84 is heterotypic.
Interestingly, structural analogs to the inner COPII coat components Sec23, Sec24, and Sar1 have not been observed in the NPC. Yet, we speculate that the NPC lattice built from Y- and Nic96 complexes is connected to the pore membrane via adaptors, consistent with the ~9nm gap between the scaffold and the pore membrane observed in cryo-tomography of the NPC (Beck et al., 2007). We speculate, that this gap could serve as a conduit for the transport of inner nuclear membrane proteins to their destination after insertion into the endoplasmic reticulum upon synthesis. In contrast to COPII vesicles, which need to be assembled and disassembled quickly and repeatedly, the NPC scaffold is a very stable entity, virtually without turnover in the non-dividing cell (D'Angelo et al., 2009). Thus the adaptor components may be very different, with Ndc1 and its direct interactors being strong candidates (Onischenko et al., 2009). Besides the difference in membrane-anchorage, the NPC-COPII analogy also differs in another way. While both systems coat highly curved membranes, the pore membrane has a decidedly different topology from a spherical vesicle. The COPII coat only needs to stabilize a contiguous, positively curved (convex) membrane. The nuclear pore membrane, like the hole in a doughnut, is convex in the longitudinal direction, but negatively curved (concave) in the latitudinal direction (Antonin and Mattaj, 2005). While the vesicle membrane has neither end nor beginning, the pore membrane ends at the perimeter of the opening. Thus, the NPC likely has elements that are specific to the distinct pore membrane topology. In the lattice model, the Y-complex is oriented relative to the pore membrane such that the Nup145C•Nup84 edge element follows the positive curvature, positioning Nup133 at the periphery of the NPC (Brohawn and Schwartz, 2009b). However, how the Y complexes are connected is unclear. While the interactions may be direct, they may also be mediated by Nic96 complex components, or other factors.

Additionally, it will be interesting to see whether the lattice model of the NPC can account for some of the elasticity and varying pore sizes that have been observed.
(Hinshaw et al., 1992). It is possible that the NPC lattice may be capable of expanding around the central axis of the pore to include a 9th or even 10th spoke in a way similar to which the COPII coat can expand to accommodate vesicles of varying sizes (Stagg et al., 2008).

**β-Propellers and the Assembly of the NPC**

Regarding the higher-order assembly of the NPC from its subcomplexes, the β-propeller proteins need to be looked at closely. In COPII, the vertices are formed from Sec31 β-propellers exclusively (Fath et al., 2007; Stagg et al., 2006). Although there is no direct equivalent to the Sec31 β-propeller in the NPC, the β-propellers of Sec31, Seh1 (including the insertion blade of Nup85) and Sec13 (including the Nup145C insertion blade) are in fact quite similar. Thus, it is tempting to speculate that these two nucleoporins, Sec13 and Seh1, play a major role in the higher-order assembly of the NPC. Propeller-propeller interactions have several intrinsic features. First, they tend to be less hydrophobic than helix-mediated interactions. Neighboring helices largely interact via exposed side chains, typically establishing a hydrophobic core. The β-propeller fold, however, exposes the peptide backbone on every outermost strand of the blades, allowing for a velcro-like interaction with a neighboring β-propeller. Since such interactions are backbone-mediated and thus largely sequence-independent, they are consequently very difficult to detect by sequence conservation, which is otherwise a useful tool for such predictive analysis.

The prospect of β-propellers mediating inter-subcomplex interactions within the NPC is likely. β-propellers are prominent mediators of protein-protein interactions and serve as versatile binding platforms throughout eukaryotes (Chaudhuri et al., 2008; Smith, 2008). Sometimes, a single β-propeller can even interact with different partners.
in a number of different manners. For instance, the Gemin5 protein contains two β-propellers and interacts with at least 5 different proteins to assemble the SMN (survival of motor neuron) complex (Gubitz et al., 2002). This complex is composed of the SMN protein and Gemin2-7 and coordinates the assembly of small nuclear ribonucleoproteins (snRNPs) important to motor neurons. Another example is the Tup1 protein found in yeast, which together with Cyc8 forms a complex that functions as a general repressor of transcription mediated by a diverse collection of repressor proteins (Zhang et al., 2002). While the propeller domain of Tup1 is not necessary for interaction with Cyc8, it interacts directly with a number of target proteins. In the COPII lattice, the N-terminal β-propeller domains of Sec31 interact with one another via three different contacts to form the vertices (Fath et al., 2007; Stagg et al., 2006). If β-propellers are also involved in lattice formation in the NPC and the interactions are equally homotypic as in the COPII coat, one can ask whether the numerous crystal structures of β-propeller nups may provide clues. Lattice contacts in crystals sometime mimic physiologically relevant interactions. The prime interaction candidate Sec13 has been crystallized in 6 different crystal forms (Fath et al., 2007; Brohawn and Schwartz, 2009b; Hsia et al., 2007; Nagy et al., 2009). Superposition of all neighboring molecules on Sec13 reveals a remarkably diverse ensemble of arrangements, not showing a strongly preferred orientation (Fig 4.2). Thus, it may be indicative of a flexible arrangement of neighboring subcomplexes.
Chapter 4: Conclusion

Figure 4.2 - Crystal contacts of the Sec13 β-propeller

A Sec13 reference molecule is shown in orange, with neighboring molecules from five different crystal structures (PDB accession codes 2PM6, 2PM9, 3BG1, 3JRO, 3JRP) superimposed to indicate the number and variety of β-propeller-β-propeller contacts that the Sec13 molecule can make in the crystal. Neighboring molecules from the same crystal are colored in the same shade of blue.

Challenges

Inter-subcomplex interactions are yet poorly defined, including contacts required to propagate 8 copies of the Y-complex around the circumference of the pore. For instance, how does the Y-complex interact with the Nic96 complex? Is this interaction direct or indirect, and if it is indirect, are there other players that we have not yet identified? Does the Y-complex interact with the Ndc1 complex? It seems unlikely that the Y-complex contacts the membrane directly in an extensive fashion. EM studies have also shown an ~9 nm gap between the NPC structural scaffold and the nuclear membrane (Beck et al., 2007). Is this indicative of an adaptor layer that links the membrane and the Y-complex, similar to the adaptor layer found in COPII? Thus far, we have taken advantage of the modular nature of the NPC (Schwartz, 2005) and the analogy to COPII vesicle coats to study its subcomplexes in molecular detail and to suggest a lattice-like model of the NPC structural scaffold (Brohawn et al., 2008), respectively. One of the next steps in understanding the structure of the NPC in its entirety is to determine the specific interactions linking all of its constituent subcomplexes together.
Not only are the inter-subcomplex interactions within the NPC numerous, but they are likely weaker and more transient in nature than intra-subcomplex interactions. This is especially evident in higher eukaryotes, where NPCs disassemble into their constituent subcomplexes during mitosis. Establishing similar interactions mediating assembly of the Sec31-Sec13 heterotetramers, which are stable in solution, into vesicle coats has challenged researchers studying the assembly of the COPII outer coat. However, due to the possibility to reconstitute the COPII coat in vitro, a model for the assembled COPII coat was inferred with the help of an EM-map of high enough resolution (Stagg et al., 2006) for the crystal structures of Sec31-Sec13 to be confidently placed (Fath et al., 2007). Recently, Kampmann and Blobel have provided an EM map of the Y-complex at 3.5 nm resolution, allowing for the tentative placement of currently solved structures into the Y-shaped density (Kampmann and Blobel, 2009; Brohawn and Schwartz, 2009b). While such single-particle reconstructions are not yet abundant, they, in combination with better-resolved EM-maps of the intact NPC, will be key in solving the structure of the NPC in molecular detail.

Additionally, studying NPC structure and assembly in vivo is likewise challenging due to the possibility of redundant interactions obscuring the function of an individual protein. While mutational and deletion analyses can be useful in studying unknown protein functions or interactions in vivo, the ability of proteins coevolved from common ancestors (i.e. ACE1 and β-propellers of the NPC) to substitute for one another may be sufficient to preserve the integrity of the scaffold. Even without such substitution, enough interactions may still remain to withstand the omission of one individual nucleoporin and may result in no apparent phenotype. For instance, we recently showed that truncation of the C-terminal 280 residues from Nup120 results in the loss of NPC localization of Nup120 (Leksa et al., 2009). However, despite severing Nup120 integration, the NPC appears to be largely functional and intact. Similarly, the dual
function of several nucleoporins is also not uncommon, ranging from Sec13’s role at the NPC and COPII vesicle coats to COPI coatomers playing an essential role in nuclear disassembly (Liu et al., 2003) and Y-complex members’ recruitment to kinetochores during mitosis (Zuccolo et al., 2007). In this case, mutational and deletion analyses can result in difficult to interpret phenotypes.

**Outlook**

It is apparent that there is a multiplicity of challenges that must be overcome to dissect and define the inter-subcomplex contacts that mediate NPC assembly. The structural relationship between the NPC and vesicle coat proteins has given us a good starting point from which to formulate a likely, yet tentative, model of the NPC structural scaffold. Building on this analogy, it seems that the β-propellers of the Y-complex will likely prove to be important players in mediating the inter-subcomplex interactions necessary for NPC assembly. Simultaneously, we must keep in mind that there are some key differences between the NPC and the COPII coat, a necessity of the different membrane topologies that are stabilized in each case. There are still a number of obstacles to overcome on the path to gaining a detailed, molecular understanding of how the individual subcomplexes come together and interact with the membrane to form the intact NPC, and it will require the continued use of multidisciplinary efforts to reach this goal. But if the progress of the last few years gives any indication of what is to come, it will not be long before a basic experimental model of the NPC scaffold will become available.
References


References


References


Drin, G., Casella, J.-F., Gautier, R., Boehmer, T., Schwartz, T. U., and Antonny, B.


References


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


References


Stagg, S. M., LaPointe, P., Razvi, A., Gürkan, C., Potter, C. S., Carragher, B., and


