Strukturna i funkcijska analiza veznog mjesta kvaščevog proteina Sec13

Gogala, Marko

University of Zagreb

Faculty of Science

Department of Biology

Marko Gogala

Structural and Functional Analysis of the Yeast Protein Sec13 With an Emphasis on its Protein Interaction Interface

Graduation Thesis

Zagreb 2009

ISKAZ

Ovaj rad izrađen u Institutu za Tehnologiju u Massachusettsu (Boston, Massachusetts, Sjedinjene Američke Države), pod vodstvom asst. prof. Thomasa U. Schwartza, predan je na ocjenu Biološkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu radi stjecanja zvanja dipl. ing. biologije, smjer molekularna biologija.

Marko Gogala

Zagreb, 2009.

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. Thomas Schwartz, without whose expertise, understanding, and patience I would not have been able to produce this thesis.

I am thankful to my co-supervisor, Dr. Ivana Đurašević-Weygand, for her advice and help with writing this thesis.

A very special thanks goes out to Stephen Brohawn, who helped me considerably by sharing his experience and providing me with advice when I needed it most.

I would also like to thank Silvija Bilokapić, whom I consider both a role model and a dear friend and who helped me immensely with making this research possible.

I am grateful to Nina Leksa, James Partridge and James Whittle, for their help and for making long hours in the lab a fun experience.

I thank Dr. Marija Luić and Nives Ivić who helped me make my first steps in the field of structural biology.

And finally, thanks to my family for the support they provided me through my entire life and whose love and encouragement made writing this thesis easier.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek Diplomski rad

STRUKTURNA I FUNKCIJSKA ANALIZA VEZNOG MJESTA KVAŠČEVOG PROTEINA SEC13

Marko Gogala

Massachusetts Institute of Technology, Department of Biology, 31 Ames Street, 68-132, Cambridge MA 02139

Jezgra omeđena ovojnicom temeljna je osobina eukariota. Kako je kompleks jezgrine pore (NPC) jedini posrednik u transportu između jezgre i citoplazme, rasvjetljavanje njegove strukture važno je za razumijevanje eukariotskih stanica. Sec13 je komponenta NPC čija je uloga u održavanju i stabilizaciji kompleksa nejasna. Njegova dvojna uloga u vezanju proteina Nup145C u sklopu NPC te proteina Sec31 u sklopu vezikula tipa COPII otežava mutacijsku analizu. Cilj ovog projekta bio je riješiti strukturu Sec13 u kompleksu sa insercijskom regijom Nup145C, kako bi se dobila jasna slika veznog mjesta Sec13/Nup145C. Riješena struktura uspoređena je sa strukturama kompleksa Sec13/Sec31. Na temelju razlika među veznim mjestima osmišljen je mutirani Sec13 koji veže isključivo protein Sec31. Posljedica ove mutacije potvrđena je pokusima *in vitro*.

(77 stranica, 27 slika, 13 tablica, 46 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u središnjoj biološkoj knjižnici

Ključne riječi: NPC, Sec13, Sec31, Nup145C, rentgenska strukturna analiza

Voditelj: Thomas, Schwartz, asst. Prof.

Ocjenitelji: Ivana Weygand-Đurašević, prof. dr. sc., Dunja Leljak-Levanić, doc. dr. sc., Dijana Škorić doc. dr. sc.

Rad prihvaćen: 08.04.2009.

BASIC DOCUMENTATION CARD

University of Zagreb

Faculty of Science

Department of Biology

Graduation Thesis

STRUCTURAL AND FUNCTIONAL ANALYSIS OF THE YEAST PROTEIN SEC13 WITH AN EMPHASIS ON ITS PROTEIN INTERACTION INTERFACE

Marko Gogala

Massachusetts Institute of Technology, Department of Biology, 31 Ames Street, 68-132, Cambridge MA 02139

Possession of an envelope-coated nucleus is the defining trait of eukaryotes. Since the nuclear pore complex (NPC) is the only known mediator of nucleo-cytoplasmic traffic, elucidating its structure is important for understanding eukaryotic cells. Sec13 is a component of the NPC whose importance for the stability of the complex is unclear. Its dual role in binding proteins Nup145 in the NPC and Sec31 in COPII vesicles makes mutational analysis of Sec13 function difficult. The goal of this project was solving a structure of Sec13 in complex with the Nup145C insertion blade that would provide a clear view of the Sec13/Nup145C interface. The obtained structure was compared to the structures of the Sec13/Sec31 complex. Based on the differences between the interaction sites of the structures, a mutant Sec13 protein was designed, that binds specifically to Sec31, but has lost affinity for Nup145C. The effect of this mutation was confirmed by *in vitro* experiments.

(77 pages, 27 figures, 13 tables, 46 references, original in English)

Thesis deposited in Central biological library

Key words: NPC, Sec13, Sec31, Nup145C, x-ray crystallography

Supervisor: Thomas, Schwartz, asst. Prof.

Reviewers: Ivana Weygand-Đurašević, prof. dr. sc., Dunja Leljak-Levanić, doc. dr. sc., Dijana Škorić doc. dr. sc.

Thesis accepted: 08.04.2009.

Abbreviations

ACE 1	Ancestral Coatomer Element 1
COPII	Coat Protein Complex 2
DNA	Deoxyribonucleic acid
FG-repeats	phenylalanine-glycine repeats
GTP	Guanosine-5'-triphosphate
His_6-tag	Hexahistidine tag
iPCR	inverse polymerase chain reaction
mRNA	Messenger ribonucleic acid
NES	Nuclear Export Signal
NLS	Nuclear Localization Signal
NPC	Nuclear Pore Complex
NTA	Nitrilotriacetic acid
Nup	Nucleoporin
PCR	polymerase chain reaction
PISA	Protein Interfaces, Surfaces and Assemblies
SDS PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sec	Secretory mutant
WD40-repeat	Approximately 40 amino acid-long motif, often terminating in a tryptophan-aspartate (W-D) dipeptide

Table of Contents

Introduction

There are three domains of life: archaea, bacteria and eukaryotes. The eukaryotes form the most anatomically diverse group, ranging in size and complexity from unicellular fungi to land plants, insects and vertebrates (Alberts et al., 2002; Campbell and Reece, 2002). Due to this evolutionary plasticity of body plan, eukaryotes inhabit a large variety of biotopes and are the only one of the three domains to include organisms above the micrometer scale (Campbell and Reece, 2002). Unlike archaea and bacteria, eukaryotes possess numerous cellular compartments; these, in turn, necessitate elaborate protein sorting and shuttling mechanisms. Understanding these mechanisms is essential for studying the biology of eukaryotic cells (Alberts et al., 2002).

The cell nucleus

The hallmark of eukaryotic evolution is compartmentalization, especially regarding genetic material (D'Angelo and Hetzer, 2008). Indeed, possession of an envelope-coated nucleus, which contains the cellular genome, is the defining trait of the eukaryotic domain (Alberts et al., 2002). The nuclear environment enables more precise regulation of gene transcription and alternative splicing of gene transcripts. This, in turn, makes multicellularity and complex developmental patterns possible. Therefore, tight regulation of nuclear import and export is of critical importance (Alberts et al., 2002; Moreno et al., 2005; Elliott and Grellscheid, 2006).

Functions of the NPC

The primary function of the NPC is selective transport of cargo, in and out of the nucleus (D'Angelo and Hetzer, 2008). Indeed, the nuclear pore complex is the only known mediator of nucleo-cytoplasmic traffic. Nuclear proteins have to be imported from the cytoplasm. On the other hand, transcription products have to be exported from the nucleus (Suntharalingam and Wente, 2003).

A remarkable property of the NPC is its selectivity regarding transport of large molecules. Molecules with a diameter of about 9 nm or less (i.e. ions and biomolecules with a molecular mass of roughly 20-40 kDa or less (Schwartz, 2005)), pass through the NPC by diffusion. Larger molecules can only pass through the pore via an active transport mechanism, requiring a specialized nuclear transport signal (Alberts et al., 2002). Several models have been proposed to explain the nature of the NPC selectivity (D'Angelo and Hetzer, 2008; Rout et al., 2000; Peters, 2005; Bayliss et al., 2000).

Although selective upstream transport of molecules is an active process, the NPC itself lacks any known structures capable of active transport. Instead, the gradient of localized molecules is maintained by separate systems of transport factors coupled to guanosine-triphosphate (GTP) hydrolyzing molecules (Rout et al., 2000). The most extensively studied family of transport factors is the karyopherin-β family of transport factors coupled to a small GTP hydrolyzing protein, Ran. Different members of the karyopherin-β family recognize either a nuclear localization signals (NLS) or a nuclear export signal (NES) on the surfaces of proteins intended for nuclear import or export.

Export of mRNA is mediated by a different family of transport factors: the TAP-NXF family (Pemberton and Paschal, 2005; Fribourg et al., 2001; Matzat et al., 2008).

Although nucleo-cytoplasmic transport is the best known function of the NPC, it is far from being the only one. The NPC has been shown to function as a docking site for enzymes that regulate important processes like transcription, DNA replication and repair, cell death and senescence, chromosome segregation and genome stability. Some nucleoporins have been implicated in nuclear envelope mitotic disassembly and re-assembly processes and kinetochore binding (D'Angelo and Hetzer, 2008; Loïodice et al., 2004; Nagai et al., 2008; Orjalo et al., 2006; Menon et al. 2005).

Structure of the NPC

The NPC structure has been studied through proteomics, biochemistry, bioinformatics, electron microscopy and x-ray crystallography (Siniossoglou et al., 1996; Cronshaw et al., 2002; Rout et al., 2000; Alber et al., 2007; Yang et al., 1998; Kiseleva et al., 2004; Jeudy and Schwartz, 2007; Hsia et al., 2007; Boehmer et al., 2008; Brohawn et al., 2008; Beck et al., 2007; Beck et al., 2004). The molecular mass of this large, protein complex is estimated to be ~50 MDa and 60-80 MDa in yeast and vertebrates, respectively. It is comprised of ~500 individual proteins, but consists of only ~30 different proteins, called nucleoporins or 'Nups' (Cronshaw et al., 2002; Rout et al., 2000; Alber et al., 2007). The NPC has a height of 35-38 nm in yeast and 80 nm in vertebrates, making it higher than the nuclear envelope (25-30 nm in yeast, 50-60 nm in vertebrates), and diameter of 96 nm and 145 nm in yeast and vertebrates respectively.

The NPC is a tripartite structure, consisting of a cytoplasmic part, nuclear part and a central cylinder. The cytoplasmic and nuclear parts are formed by bundles of filaments. The filaments on the cytoplasmic side radially emanate into the cytoplasm. The nuclear filaments join into a structure that has been termed the 'nuclear basket'. The central cylinder has an octagonal symmetry axes perpendicular to the nuclear envelope plane, dividing the structure into eight 'spokes'. The most striking feature of the central cylinder is the central channel. It has a diameter of 49 nm and is filled with filamentous material. The central channel is the route of transport in

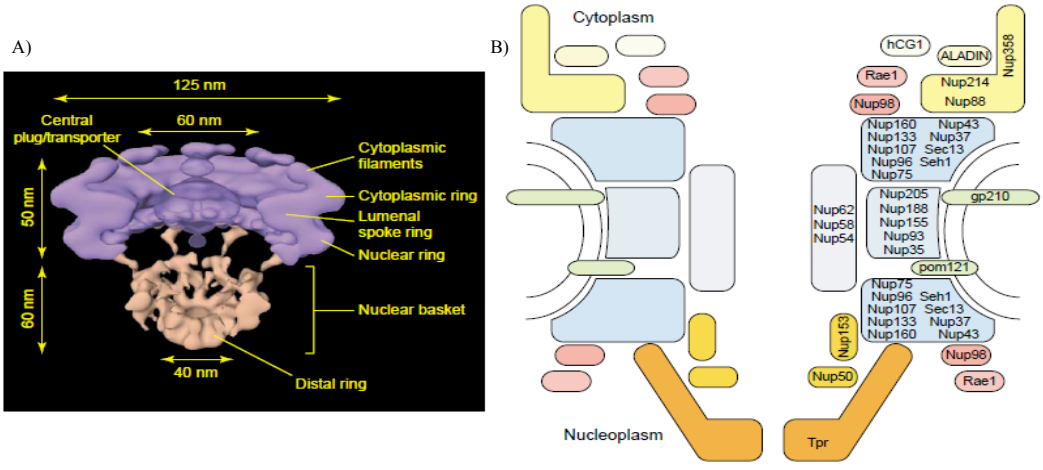

Figure 1. A) A cryo-electron microscopic model of the Dyctiostelium NPC. Major structural features are outlined. B) A simplified scheme of the metazoan NPC. Approximate localizations of nucleoporins are outlined. Schwartz 2005

and out of the nucleus for soluble molecules. Historically, electron microscopic studies have reported the existence of a cloud of electron density within the central channel (Yang et al.,

1998). This had been termed either the 'central plug' or 'central transporter', depending on the perceived function of the structure. Recent studies dismiss the 'central plug/transporter' as various cargo particles, trapped in the central channel during the electron microscopy preparation process (Peters, 2005; Beck et al., 2007; Tran and Wente, 2006) (Figure 1).

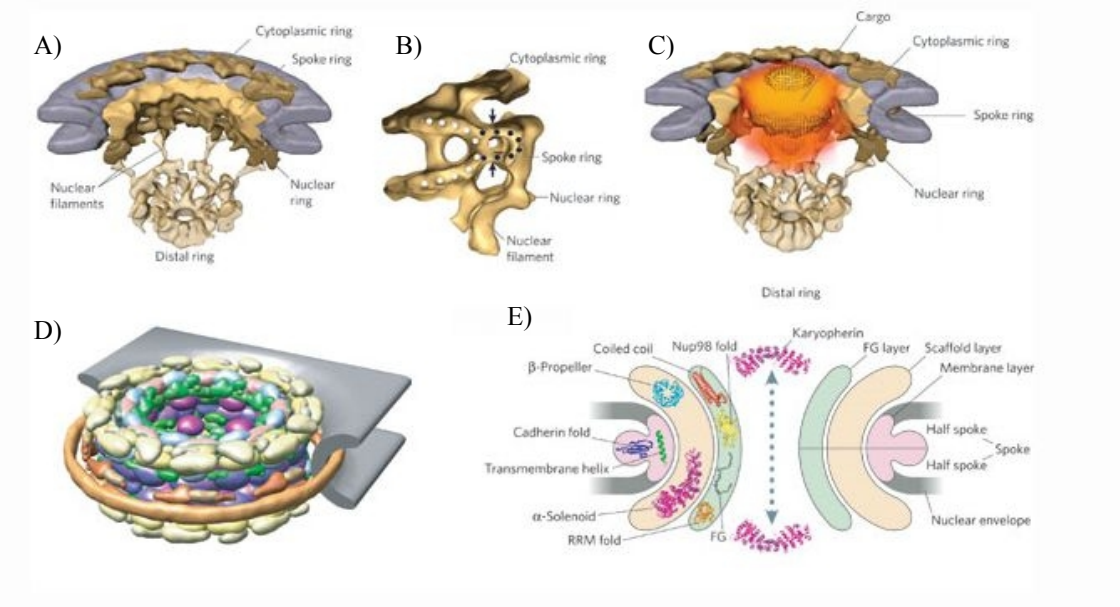

Figure 2. Model of the NPC. A) A cryo-electron microscopic reconstruction of the Dyctiostelium NPC. B) A highlight of the NPC ring system . C) The approximate path of the nucleo-cytoplasmic cargo through the central channel. D) A model of the composition of the NPC. E) A suggestion of the fold distribution throughout the NPC. (Alber et al., 2007)

The central cylinder has been shown to consist of four ring systems: a central 'spoke ring', a 'membrane ring', partially located in the NE lumen and transversing the pore membrane and a pair of peripheral rings, located on the cytoplasmic and nucleoplasmic sides of the central cylinder and accordingly named the 'cytoplasmic ring' and the 'nuclear ring' (Akey and Radermacher, 1993; Beck et al., 2007; Kiseleva et al., 2004).Overall, the volume of the yeast NPC is five times smaller than its vertebrate counterpart. However, symmetry, ring systems and other key structural elements (e.g. a central channel and filaments) of the NPC are conserved between yeast and vertebrates (Cronshaw et al., 2002; Rout et al., 2000; Yang et al., 1998; Kiseleva et al., 2004).

Although the detailed arrangement of proteins within the NPC has yet to be unambiguously determined (Alber et al., 2007; Hsia et al., 2007; Brohawn et al., 2008), it is well known that the NPC is composed of several subcomplexes such as the Nsp1, Nup84, Nup159/82 and Nic96 subcomplexes in yeast (Nup62, Nup107-160, Nup214/88 and Nup93 subcomplexes in vertebrates, respectively) (D'Angelo and Hetzer, 2008; Tran and Wente, 2006). These subcomplexes, in addition to being biochemically stable, remain stable throughout mitosis, when

the NPC as a whole is disassembled (Loïodice et al., 2004). Localization of subcomplexes within the NPC using immunoelectron microscopy has revealed that the Nup107-160 subcomplex is a part of the central spoke rings and Nup214/88 forms the cytoplasmic ring. Additionally, another nucleoporin, Tpr has been localized to the nuclear basket, providing a map of approximate positions of Nups within the NPC (Schwartz, 2005) (Figure1).

Recently, a computational study based on a wealth of biochemical, biophysical and structural data, has produced a model of the distribution of different proteins within the yeast nuclear pore complex that is in line with electron microscopic measurements (Alber et al., 2007; Figure2). According to this model, the central cylinder of the NPC is built in a highly modular fashion, essentially consisting of 32 repetitions of a single structural module.

NPC proteins are predicted to possess eight types of folds, mostly composed of multiple, simple, repetitive units. These fold types are, in turn, partitioned into three groups, according to their localization (Devos et al., 2006). Core scaffold proteins are believed to belong to three domain architectures: α-solenoids, β-propellers or a binary combination of both, with the β-propeller located amino-terminally. X-ray crystallographic studies showed, however, that the predictions do not match the experimental data very well (Boehmer et al., 2008; Jeudy and Schwartz, 2007; Hsia et al., 2007; Brohawn et al., 2008).

The seven-bladed β-propeller was first identified in proteins containing a WD40 motif. However, additional nucleoporins without a detectable conservation of this motif in their sequence have been found to have the seven-bladed β-propeller fold, allowing for improved prediction of nucleoporin structures and suggesting a significant evolutionary adaptability of this fold, regarding primary structure (Berke et al., 2004).

The structural core of the central cylinder is believed to consist of a network of α-solenoids. The flexibility of this fold is thought to insure the ability of the NPC to accommodate cargo of different sizes (Alber et al., 2007; Devos et al., 2006). However, the recent determination of structures for many of the components of the core scaffold revealed a novel α-helical fold that is rigid, rather than flexible, putting the pre-existing theories of NPC fold composition and flexibility in question (Boehmer et al., 2008; Jeudy and Schwartz, 2007; Hsia et al., 2007; Brohawn et al., 2008).

Proteins lining the central channel possess an unstructured region, consisting of repeats rich in phenylalanine and glycine (termed the FG-repeats). The FG-repeats fill the interior of the central channel and extend into the cytoplasm and nucleoplasm. They form the filamentous

matter observed in EM studies. Most of the FG-repeat proteins are anchored to either the inner rings or the linker nucleoporins of the central cylinder via coiled-coil-type anchoring domains. The FG-repeats are generally agreed to form a meshwork that serves as a selective barrier for nucleo-cytoplasmic transport (D'Angelo and Hetzer, 2008; Rout et al., 2000; Peters, 2005; Bayliss et al., 2000).

Nup84 subcomplex

The best characterized subcomplex of the NPC is the Nup84 subcomplex in yeast and its metazoan counterpart, the Nup107-160 subcomplex (referred to as yNup84 and mNup107-160 subcomplex, respectively). The yNup84 subcomplex consists of seven proteins: yNup133, yNup84, ySec13, yNup145C (the carboxy-terminal part of yNup145 after *in vivo* autoproteolysis), yNup120, yNup85 and ySeh1 (Sec13 homolog 1). Their counterparts in the mNup107-160 subcomplex are: mNup133, mNup107, mSec13, mNup96, mNup160, mNup75 and mSeh1, respectively. Two additional Nups exist in the mNup107-160 subcomplex: mNup37 and mNup43 (Boehmer et al., 2008; Rout et al., 2000; Loïodice et al., 2004; Belgareh et al., 2001; Vasu et al., 2001; Ratner et al., 2007).

Each NPC contains 16 copies of the yNup84 subcomplex, distributed equally between the cytoplasmic and nucleoplasmic side of the complex in the outer scaffold rings of the central cylinder. (Alber et al, 2007). It is a 500 kDa , Y-shaped structure, about 42 nm long and 25 nm in diameter (Brohawn et al., 2008; Lutzmann et al., 2002). The yNup84 subcomplex has been reconstructed *in vitro*. The reconstitution experiment revealed the subcomplex to be composed of several subcomponents. Proteins yNup145C and ySec13 form a soluble heterodimer when isolated from the rest of the yNup84 complex, as do yNup85 and ySeh1. Another pair of closely associated proteins is formed by yNup133 and yNup84. In the yNup84 subcomplex, yNup84 binds the ySec13/yNup145C pair making a yNup133/yNup84-ySec13/yNup145C heterotetramer. Both this heterotetramer and the ySeh1/yNup85 heterodimer bind yNup120 to form the heptameric yNup84 subcomplex (Lutzmann et al., 2002).

Progress in characterizing the yNup84 complex has been made via crystallographic studies of the bound protein pairs mNup133/mNup107 (Boehmer et al., 2008), yNup145C/mSec13 (Hsia et al., 2007) and yNup85/ySeh1 (Brohawn et al., 2008).

As predicted for WD40 containing proteins (Berke et al. 2004; Devos et al. 2004), mSec13 and ySeh1 assume the fold of a β-propeller, with each blade consisting of four antiparallel β-strands; however, unlike the 'closed' structure of a typical seven-bladed β-propeller, both of these proteins form an 'open', six-bladed β-propeller. The seventh, 'insertion' blade, composed of three β-strands, is provided by the interacting partner (yNup85 and yNup145C for ySeh1 and ySec13 respectively) (Brohawn et al., 2008; Hsia et al., 2007) .

An interesting pattern comes to attention when structures of yNup145C and yNup85 are compared. Secondary structures of both of these proteins are, with the exception of their amino-terminal insertion blades, entirely α-helical, however, they do not assume the predicted α-solenoid fold. Instead, in yNup85, helices α1-α3 (counting from the amino terminus) and α12-α19 form two strings of α-helices, zig-zagging in opposite directions, forming a structure termed 'the trunk'. In between the two antiparallel strings, helices α4-α11 form a tight bundle capping the trunk, with helices α5-α10 positioned almost perpendicularly to the meandering helices of the trunk. This central bundle has been named 'the crown'. Together, the trunk and the crown form a cuboid, U-shaped structure with a profoundly irregular pattern of α-helices, very different from the ordered α-solenoid. A very similar arrangement of α-helices, with a distinct crown and trunk is visible in yNup145C. Both of these proteins possess a structurally uncharacterized, carboxy-terminal domain. This domain is predicted to be mostly α-helical and is termed 'the tail' and is necessary for binding to yNup120 (Brohawn et al., 2008; Hsia et al., 2007).

The structure of another NPC component, yeast protein Nic96, shows striking similarity to yNup85 and yNup145C and conforms well to the crown-trunk-tail architectural model. Nic96 is an important linker nucleoporin and a component of the inner scaffold ring (Jeudy and Schwartz, 2007; Alber et al., 2007). All three parts of Nic96 have been structurally characterized, providing an insight into the shape of the 'tail' part of the protein. The Nic96 tail consists of 11 α-

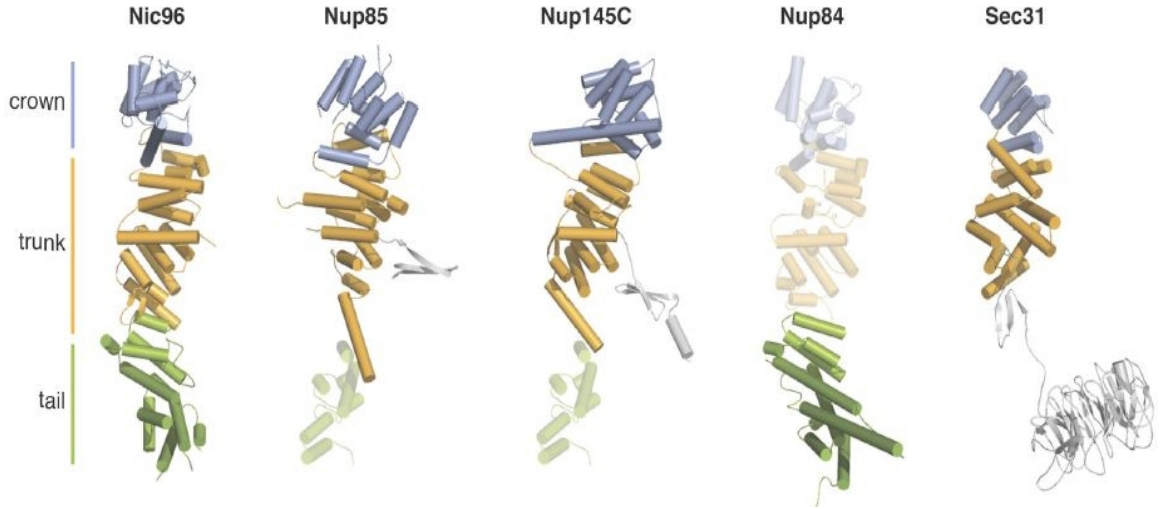

Figure 3. ACE-1 domains of Nic 96, Nup85, Nup145C, Nup84 and Sec31. (Brohawn et al., 2008)

helices that pair in a similar manner to the helices in the trunk, although they do not make a U-turn, but rather an irregular, kinked structure (Jeudy and Schwartz, 2007; Brohawn et al., 2008).

Comparison of the Nic96 structure to the structure of interacting carboxy-terminal

domains of mNup133/mNup107 reveals that the carboxy-terminal domain of mNup107 shares the arrangement of α-helices similar to that seen in the Nic96 tail, confirming the consistency of the crown-trunk-tail model and suggesting that the tails of yNup85 and yNup145C are organized in the same fashion. Computer models based on the data from the Nic96, yNup85 and yNup145C structures predict the rest of the yNup84 carboxy-terminal domain to assume the crown and trunk folds (Brohawn et al., 2008).

In addition to binding yNup120 and ySec13, yNup145C also binds yNup84. This interaction has been localized to conserved, hydrophobic α-helices in the crown segments of both proteins. Thus, yNup145C forms three contacts: to Sec13 via its amino-terminal insertion blade, to yNup120 via its tail and to yNup84 through a crown-crown interaction (Brohawn et al., 2008).

A further examination of known crystal structures reveals that the yeast protein Sec31 also possesses a crown and a trunk. This vesicular protein has an amino-terminal seven-bladed β-propeller domain, followed by an insertion blade (serving as a contact site for ySec13), followed by a helical domain, assuming the crown and trunk folds. The carboxy-terminal region of Sec31 is predicted to be α-helical, but contains a proline-rich region implicated in activating the GTPase activity of Sar1 (another component of COPII vesicles), making the exact nature of its fold uncertain (Fath et al., 2007; Brohawn et al., 2008; Bi et al., 2007).

Sec31 dimerizes through crown-crown binding. But unlike the yNup145C/yNup84 interaction, this is a domain-swap interaction, creating a 'mixed crown', composed of crowns from the two proteins and topologically similar to the single crowns of yNup145C, yNup85 and yNic96 (Bi et al. 2007; Brohawn et al. 2008). Binding of Sec13 is important in Sec31 oligomerization, as the Sec13 β-propeller is inserted between the α-helical an the β-propeller domains of Sec31 and interacts with both. In particular, Sec13 helps position the Sec31 β-propeller into a orientation that enables Sec31 β-propeller dimerization (Bi et al. 2007).

Because the crown-trunk-tail domain architecture is present in several components of a complex that coats the nuclear pore and in Sec31, which coats the COPII vesicles, it has been named the ACE 1 domain (Ancestral Coatomer Element 1) (Brohawn et al. 2008) (Figure 3).

Models of the NPC scaffold structure based on the Nup84 complex

Two radically different models of the NPC core scaffold have recently surfaced (Hsia et al., 2007; Brohawn et al., 2008).

Based on the structures of ySeh1/yNup85, mSec13/yNup145C, mNup107/mNup133 complexes and a wealth of biochemical and biophysical information on protein interactions within the yNup84/mNup107-160 complex, the Schwartz group (Brohawn et al. 2008) proposes a

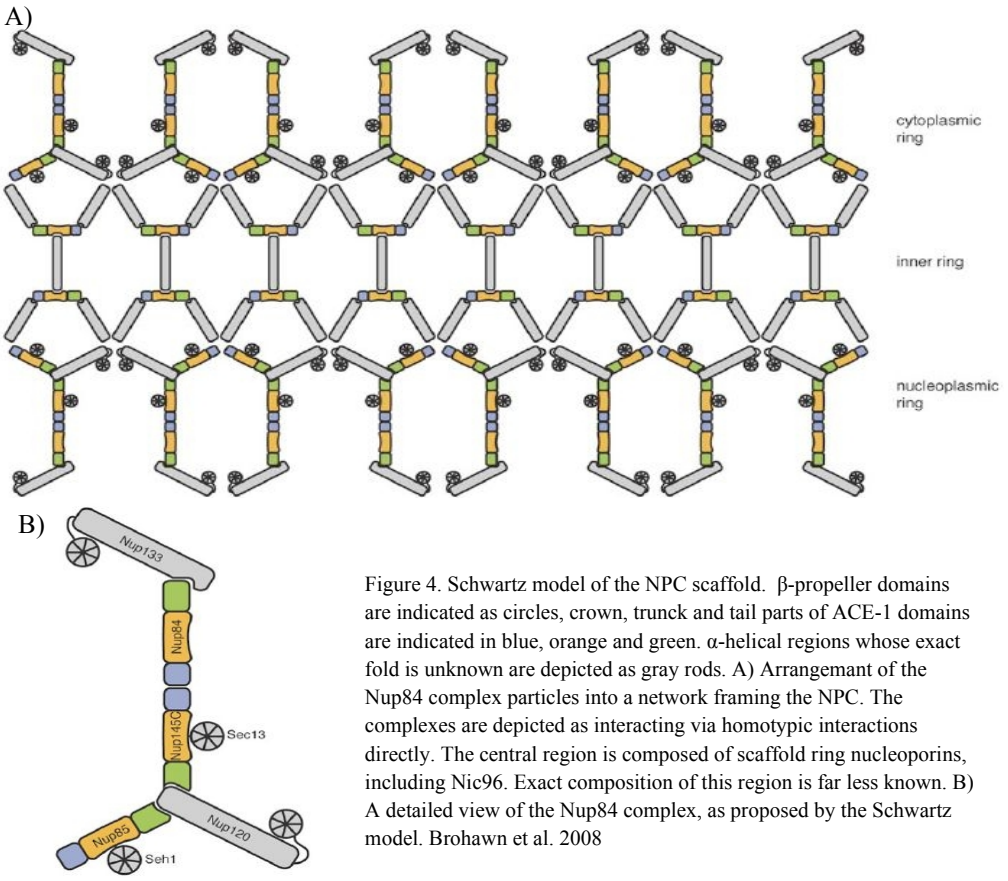

Figure 4. Schwartz model of the NPC scaffold. β-propeller domains are indicated as circles, crown, trunck and tail parts of ACE-1 domains are indicated in blue, orange and green. α-helical regions whose exact fold is unknown are depicted as gray rods. A) Arrangemant of the Nup84 complex particles into a network framing the NPC. The complexes are depicted as interacting via homotypic interactions directly. The central region is composed of scaffold ring nucleoporins, including Nic96. Exact composition of this region is far less known. B) A detailed view of the Nup84 complex, as proposed by the Schwartz model. Brohawn et al. 2008

model of the Y-shaped yNup84 complex and of the structural scaffold of the NPC as a whole (Figure 4). According to this model, the longer end of the Y-shaped complex is formed by yNup133 bound via its carboxy-terminal domain to the tail of yNup84. In turn, yNup84 binds yNup145C by a crown-crown interaction. Through its tail, yNup145C also interacts with yNup120, which forms one of the shorter ends of the Y shaped structure. Also bound to yNup120

is yNup85, forming the second shorter end of the Y-structure. Peripherally bound to this predominantly α-helical complex would be the β-propeller folds of ySec13, ySeh1 and the predicted amino-terminal β-propeller domain of yNup133.

Hsia et al. suggest an altogether different model (Hsia et al., 2007) (Figure 5). The rationale for this model is based on the structure of human Sec13 bound to yNup145C. The unit cell of the hSec13/yNup145C contains a heterooctamer formed by interactions of four vSec13/yNup145C heterodimers and has a shape of a 285 Å long, slightly bent rod.

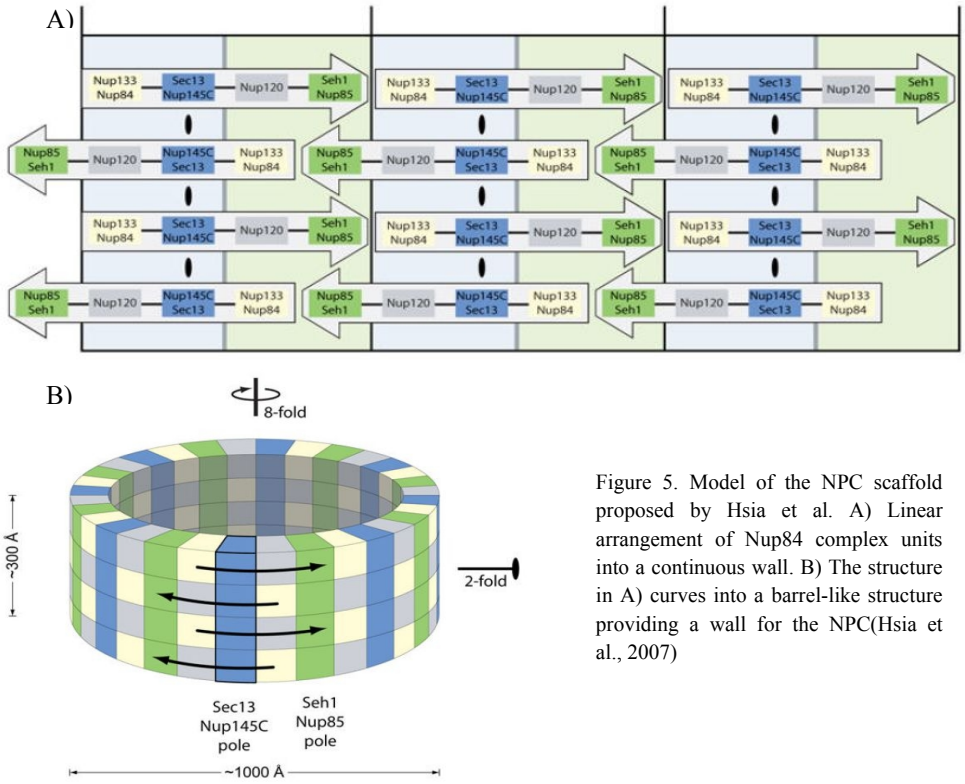

Figure 5. Model of the NPC scaffold proposed by Hsia et al. A) Linear arrangement of Nup84 complex units into a continuous wall. B) The structure in A) curves into a barrel-like structure providing a wall for the NPC(Hsia et al., 2007)

According to the model proposed by Hsia et al., eight of these rods, placed vertically with respect to the pseudo-mirror symmetry plane of the NPC, support four stacked rings composed of seven yNup84 complexes, bound in a head-tale fashion. The authors propose that ySeh1/yNup85 form an additional rod, increasing the number of vertical poles to sixteen.

Importance of Sec13 for NPC function

Strikingly different roles that Sec13 plays in the two proposed models of NPC structure make the question of its importance for NPC function especially interesting. Due to its double function as part of the NPC and of the COPII vesicle coat, deletion of Sec13 is lethal in yeast (Pryer et al., 1993), rendering loss of function analysis impossible. Therefore, very little is known about the significance of Sec13 for NPC structure, function or stability. A Sec13 mutation ($AA224^{S\rightarrow K}$) that specifically disrupts localization of ySec13 to COPII vesicles in a temperature-dependent fashion without disrupting localization to the NPC is lethal, indicating that loss of ySec13 function in vesicle transport alone is enough to cause lethality (Siniossoglou et al., 1996). For this reason, a mutant ySec13 that would specifically bind Nup145C, but would not interact with Sec31 might be a good tool for probing the role of Sec13 in the NPC. An orthogonal mutant would be equally desirable.

Although *in vitro*, Nup145C is insoluble without Sec13, this does not prove that the same holds true *in vivo*. Sec13 homolog, Seh1 is necessary for Nup84 solubility *in vitro* (Lutzmann et al., 2002), but the Seh1 gene is not essential in yeast (Siniossoglou et al., 1996), demonstrating that Seh1 protein is not essential for functioning of yNup85 *in vivo*. The Sec13/Nup145C pair may prove to be similar. Further, this suggests that the role β-propeller nucleoporins playing the structure of the yNup84 complex could be marginal.

Finding that a single mutation in the Sec13-Nup145C interface is sufficient for complete disruption of interaction might indicate that Sec13 binds the Nup145C insertion blade without strongly interacting with its trunk domain, suggesting high mobility of the Sec13 β-propeller in a manner reminiscent of a medieval flail weapon. Effects of such a mutant could be studied by in vivo localization of Sec13 labeled with a fluorescent probe (Osmani et el., 2006). Loss of Sec13 localization to the NPC without a significant impact on cell growth would support a marginal role of Sec13 in the NPC. Lethality of the mutation disrupting Sec13 localization to the NPC would signify a critical role of Sec13 in NPC structure. Localization of the mutant Sec13 to the NPC despite the loss of Nup145C binding would be suggestive of additional, as yet unknown interactions (Siniossoglou et al., 1996; Boehmer et al., 2008). Such interactions are not wholly inconceivable as β-propeller folds have already been implicated in protein interactions that do not utilize insertion blade binding (Bi et al. 2007).

One of the approaches for designing a mutant Sec13 that would bind exclusively to Sec31 is to compare interfaces of Sec13-Sec31 and Sec13-Nup145C and find amino acids that can be mutated to selectively affect Sec13-Nup145C binding.

However, as the only extant structure of the Sec13/Nup145C complex is a hybrid between a human vSec13 and yeast yNup145C (Hsia et al., 2007), it is of little relevance for mutant design. Since two structures of ySec13 bound to ySec31 are known (Fath et al., 2007), a structure of the ySec13-yNup145C complex is needed for binding site comparison.

The goal of this project was obtaining a crystal structure of the yeast Sec13 in complex with the yNup145C insertion blade that would provide a clear view of the ySec13/yNup145C interface, i.e. a structure with a resolution of <2.8 Å (Berg et al., 2002). The obtained 2.5 Å structure of the ySec13/yNup145C complex was compared to the structures of ySec13/ySec31. Based on the differences between the interaction sites of the two structures, a mutant ySec13 protein was designed, that binds specifically to ySec31, but has lost affinity for yNup145C. The effect of this mutation was confirmed by *in vitro* experiments.

Materials and methods

Materials

Standard chemicals

agar (Sigma), agarose (EMD), acetic acid (VWR), acrylamide (National Diagnostocs), ammonium persulphate, APS (VWR), ampicillin (VWR), bromophenol blue, BPB (VWR), chloramphenicol (VWR), Coomassie Brilliant Blue R250 (Sigma), dipotassium phosphate (Sigma), dithiothreitol, DTT (Gold Biotechnology), ethanol, 100% p.a. (VWR), ethidium bromide (Sigma), ethylenediaminetetraacetic acid, EDTA (Sigma), glycerol (Sigma), glycine (), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, HEPES (Sigma), imidazole (Sigma), isopropyl alcohol (VWR), isopropyl β-D-1-thiogalactopyranoside, IPTG (Gold Biotechnology), lithium chloride (Sigma), monopotassium phosphate (Sigma), kanamycin (VWR), hydrochloric acid, 36% p.a. (VWR), 2-mercaptoethanol (VWR), N,N'-methylenebisacrylamide (National Diagnostocs), phenylmethanesulphonylfluoride, PMSF (Sigma), polyethylene glycol 4000, PEG 4K (Sigma), sodium acetate (Sigma), sodium dodecyl sulfate, SDS (Sigma), sodium hydroxide (Sigma), sodium chloride (Sigma), tetramethylethylenediamine, TEMED (Sigma), tris(hydroxymethyl)aminomethane , Tris (Sigma), urea (Sigma), xylene cyanol (Sigma)

Enzymes and other protein reagents

BamHI (*NEB*), BSA (*NEB*), DpnI (*NEB*), NdeI (*NEB*), NotI (*NEB*), PfuUltra II Fusion HS DNA Polymerase (*Stratagene*), PrescissionProtease (*GE healthcare*), T4 Ligase (*NEB*), T4 Polynucleotide Kinase (*NEB*), XhoI (*NEB*)

Strains and plasmids

For plasmid amplification and cloning, the *Escherichia coli* DH5α (*supE44*, Δ*lac*U169 (Φ80*lac*ZΔM15), *rec*A1, *end*A1, *hsd*R17, *thi*-1, *gyr*A96, *rel*A1) strain was used. For protein overexpression, the *E. coli* BL21-RIL (F^-, *ompT* , *hsdS*(r_B^- m_B^-), *dcm*$^+$, Tet^r, *gal, endA,* Hte [*argU ileY leuW* Cam^r]) strain was used.

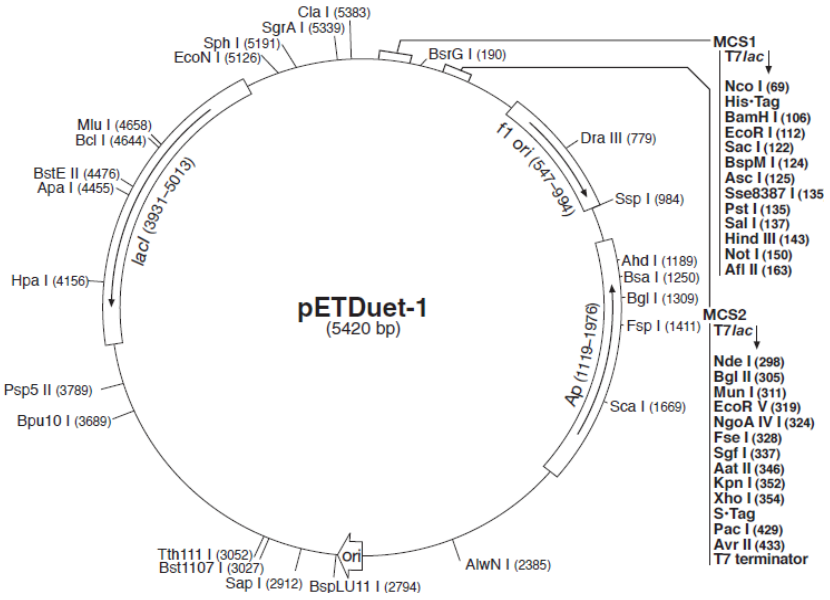

Figure 6. pETDuet-1 plasmid chart

The **pETDuet-1** (Figure 6) vector with cloned genes of interest was used in all experiments. pETDuet-1 is suitable for coexpression of two genes of interest. This plasmid contains two multiple cloning sites (MCS), each of which is preceded by a T7 promoter/*lac* operator and a ribosome binding site (rbs), thereby comprising two separate expression cassettes. The first MCS is also preceded by a sequence coding for a polyhistidine (His_6) tag. The pETDuet-1 vector also carries the pBR322-derived ColE1 replicon, *lacI* gene and ampicillin resistance gene.

Following plasmids were derived from pETDuet-1:

pETDuet1-Sec13-Nup145C plasmid carrying a gene coding for Sec13 in the first expression cassette and a gene for Nup145C in the second cassette.

pETDuet1-Sec13/Nup145C (109-712) plasmid carrying the gene coding for Sec13 fused to Nup145C amino acids 109-712 through a glycine (4x)-serine-glycine(4x) linker .

pETDuet1-Sec13/Nup145C (109-179) plasmid carrying the gene coding for Sec13 fused to the gene coding for Nup145C amino acids 109-179 through a glycine (4x)-serine-glycine(4x) linker .

pETDuet1-Sec13(I12F)-Nup145C plasmid carrying a gene coding for Sec13 with a 12I→F mutation in the first expression cassette and the gene for Nup145C in the second cassette.

pETDuet1-Nup145C-Sec13 plasmid carrying the gene coding for Nup145C in the first expression cassette and the gene for Sec13 in the second cassette.

pETDuet1-Nup145C-Sec13(I12F) plasmid carrying the gene coding for Nup145C in the first expression cassette and the gene for Sec13 with the $12^{I\rightarrow F}$ mutation in the second cassette.

pETDuet1-Sec13-Sec31(372-411) plasmid carrying the gene coding for Sec13 in the first expression cassette and a gene for Sec31 amino acids 372-411 in the second cassette.

pETDuet1-Sec13(I12F)-Sec31(372-411) plasmid carrying the gene coding for Sec13 with the $12^{I\rightarrow F}$ mutation in the first expression cassette and the gene for Sec31 amino acids 372-411 in the second cassette.

pETDuet1-Sec31(1-748)-Sec13 plasmid carrying a gene coding for a gene for Sec31 amino acids 1-748 in the first expression cassette and Sec13 in the second cassette.

pETDuet1-Sec31(1-748)-Sec13(I12F) plasmid carrying the gene coding for a gene for Sec31 amino acids 1-748 in the first expression cassette and Sec13 with the $12^{I\rightarrow F}$ mutation in the second cassette.

Methods for working with DNA

Polymerase chain reaction

Polymerase Chain Reaction (PCR) is a method for *in vitro* amplification of DNA sequences. DNA primers and the DNA polymerase are pre-added to a PCR mixture, while cyclic variations in temperature serve as a substitute for *in vivo* replication machinery: different temperatures are defined for strand separation, primer annealing and DNA polymerization. The DNA polymerase used is usually a thermostable protein isolated from thermophilic organisms Several modifications of PCR were used in this project.

iPCR

Inverse PCR (iPCR) was used in producing the pETDuet1-Sec13/Nup145C(109-179) plasmid from the pETDuet1-Sec13/Nup145C (109-712) plasmid (kindly donated by S. Brohawn), for removal of parts of the Sec13/Nup145C fusion gene coding for regions of Nup145C that were not part of the crystallization construct. Unlike most PCR reactions that amplify a small segment of a large DNA molecule, iPCR amplifies a majority of the plasmid, leaving out a small part of the sequence. The reaction mix for a 50 µl reaction is given in (table 1), the thermal cycling parameters are in (table 2). Primers used are in (table 3).

Component	**Volume**
H_2O	30.25 µl
DNTP mix (2mM each dNTP)	6.25 µl
Forward primer (5 µM)	2 µl
Reverse primer	2 µl
Vector DNA	4 µl
Pfu Buffer (10x)	5 µl
PfuUltra II Fusion HS DNA Polymerase	0.5 µl

Table 1. reaction mix for iPCR

Temperature	**Time**
95°C	2 min
Following steps repeated in 30 cycles:	
95°C	20 sec
5°C less than the lowest melting temperature	20 sec
72°C	15 sec/kb-pairs of the vector
After 30 cycles, the reaction ends with following steps	
72°C	3 min
4°C	until the reaction is terminated

Table 2. Thermal cycling parameters for iPCR

Forward primer	5'-GCTTTAGAGAAGAACTATTCG-3'
Reverse primer	5'-GATTTAATGAAGTGTACATATAAG-3'

Table 3. Primers used for iPCR of the Sec13/Nup145C fusion gene

After iPCR, remnants of PCR reagents were removed from the amplified DNA using the *QIAquick PCR purification kit* (*Qiagen*).

The purified iPCR product was DpnI digested (50 µl of purified iPCR product, 5.7 µl NEB-buffer 4 (10x), 1.3 µl DpnI) for 1h at 37°C. DpnI specifically recognizes a methylated restriction site. Since PCR products are not methylated, DpnI restriction ensured that the parental plasmid DNA would be degraded. The digest was again purified using the QIAquick PCR purification kit, but into 30 µl of elution buffer. Purified, DpnI digested iPCR product was than phosphorylated at 5' ends: 11.5 µl of iPCR product was mixed with 1.5 µl of T4 ligase buffer (10x) and 1 µl of T4 polynucleotide kinase and incubated for 1h at 37°C. After 1h, 1 µl of T4 ligase was added to the mix. Ligation proceeded at room temperature overnight. After the incubation, *E. coli* DH5α cells were transformed with 5 µl of ligated vectors.

Site-directed mutagenesis

Plasmids pETDuet1-Sec13(I12F)-Nup145C, pETDuet1-Nup145C-Sec13(I12F), pETDuet1-Sec13(I12F)-Sec31(372-411) and pETDuet1-Sec31(1-748)-Sec13(I12F) were created from pETDuet1-Sec13-Nup145C, pETDuet1-Nup145C-Sec13, pETDuet1-Sec13-Sec31(372-411) and pETDuet1-Sec31(1-748)-Sec13 plasmids via site-directed mutagenesis. The PCR for mutagenesis was done following the protocol for iPCR (table 1)(table 2). Unlike primers for iPCR, forward and reverse primer pairs for site directed mutagenesis overlapped by several bases. The mutant base was located roughly in the center of the overlap (table 4).

Forward primer	5'-GAATTA**TTC**CATGACGCTGTTCTAG-3'
Reverse primer	5'-CGTCATG**GAA**TAATTCGTTGTGC-3'

Table 4. Primers used for the Sec13 AA12Ile→Phe mutagenesis. Overlapping ends of the primers are underlined. The mutated codon (ATC TTC) is in bold.

PCR products were analyzed on agarose gels and PCR purified using the QIAquick PCR purification kit, into 30 µl of elution buffer. The purified PCR products were than DpnI digested (30 µl purified PCR product, 3.5 µl NEB-buffer 4 (10x), 1.5 µl DpnI) for 1h at 37°C. After the incubation, *E. coli* DH5α cells were transformed with 5 µl of DpnI digested vectors. Since primers for site-directed mutagenesis overlap, the PCR products had overlapping ends. Therefore, *in vitro* phosphorilation and ligation were unnecessary as *E. coli* enzymes performed the ligation reaction.

Genomic PCR

DNA fragments coding for Sec31(372-411) and Sec31(1-748) proteins were PCR amplified from *S. cerevisiae* genomic DNA (kindly donated by S. Brohawn). Two PCR runs were done: in the first run (table 5)(table 2),

Component	
H_2O	33.75 µl
DNTP mix (2mM each dNTP)	6.25 µl
Forward primer (5 µM)	2 µl
Reverse primer	2 µl
Vector DNA	0.5 µl
Pfu Buffer (10x)	5 µl
PfuUltra II Fusion HS DNA Polymerase	0.5 µl

Table 5. reaction mix for the first round of genomic PCR

the DNA fragments were amplified directly from *S. cerevisiae* genomic DNA. The PCR product was analyzed by agarose electrophoresis and purified using the QIAquick PCR purification kit. The resulting product, free of both remnants of *S. cerevisiae* DNA and of PCR reagents, was submitted to a second run of PCR (table 6)(table 7), optimized for amplification of sequences from smaller templates. Same primers were used for both runs.

Component	
H_2O	33.25 µl
DNTP mix (2mM each dNTP)	6.25 µl
Forward primer (5 µM)	2 µl
Reverse primer	2 µl
Vector DNA	1 µl
Pfu Buffer (10x)	5 µl
PfuUltra II Fusion HS DNA Polymerase	0.5 µl

Table 6. reaction mix for the second round of genomic PCR

Temperature	**Time**
95°C	1 min
Following steps repeated in 30 cycles:	
95°C	20 sec
5°C less than the lowest melting temperature	20 sec
72°C	30 sec/kb-pairs of the vector
After 30 cycles, the reaction ends with following steps	
72°C	3 min
4°C	until the reaction is terminated

Table 7. Thermal cycling parameters for the second round of genomic PCR

The primers used for amplification had additional, overhang sequences containing sites recognized by restriction enzymes (table 8).

Sec31(372-411) forward primer	5'-GTGGTGTCCATATGCAAGCCCCAACTTGGTATG-3'
Sec31(372-411) reverse primer	5'-CCTGCTCTCGAGTCAGCCTGAAATTTTTGGGTTTGT TATAGATAC-3'
Sec31(1-748) forward primer	5'-ACCAGGATCCATGGTCAAACTTGCTG-3'
Sec31(1-748) reverse primer	5'-GCA AGG CGG CCG CTA ATG ATT TGC-3'

Table 8. Primers used for amplification of sequences coding for Sec31(372-411) and Sec31(1-748). Recognition sites for NdeI (Sec31(372-411) forward primer), XhoI (Sec31(372-411) reverse primer), BamHI (Sec31(1-748) forward primer) and NotI (Sec31(1-748) reverse primer) are underlined.

Isolation of plasmid DNA

Plasmid DNA was isolated using the QIAprep spin miniprep kit, from *E. coli* DH5α grown in 5ml of medium with added ampicillin (0.1 mg ampicillin/ml) and incubated for ~7 h at 37°C with shaking (250 rpm). DNA yield was estimated from absorbency at 260 nm, using the approximation

$$\gamma(\mathrm{DNA}) \approx A_{260} \cdot 50 \text{ mg/ml}$$

Cloning into the plasmid vectors

Plasmids pETDuet1-Sec13-Sec31(372-411) and pETDuet1-Sec31(1-748)-Sec13 were created from plasmids pETDuet1-Sec13-Nup145C and pETDuet1-Nup145C-Sec13 through cloning. The pETDuet1-Sec13-Nup145C vector and genomic PCR product Sec31(372-411) were digested with restriction enzyme pair NdeI, XhoI. The pETDuet1-Nup145C-Sec13 vector and Sec31(1-748) insert were digested with the BamHI, NotI restriction enzyme pair. In both cases digestion of the vector removed a DNA fragment carrying the gene coding for Nup145C, leaving a place for Sec31 fragment insertion. The incubation time and volumes of substrate DNA were determined by total mass of the DNA (determined from A_{260}) and by enzyme concentration (supplied by the manufacturer). Incubation temperature and the buffer used for double digest were determined according to instructions from the manufacturer. BSA was added to reaction mixtures; BSA reduces enzyme loss on tube and pipette tip surfaces, BSA has also been shown to stabilize enzymes in during reactions. The general reaction scheme is given in table 10.

Component	**Volume**	**Final concentration**
Vector/Insert DNA	variable	variable
Restriction enzyme 1	1 µl	variable
Restriction enzyme 2	1 µl	variable
BSA	variable	100 µg ml^{-1}
NEB buffer	variable	1x
H_2O	variable	\

Table 9. The reaction scheme for restriction of DNA fragments

The insert digestions were purified by QIAquick PCR purification kit into 30 µl of elution buffer. Loss of insert DNA was regarded as insignificant. Entire digested vector DNA was loaded onto an agarose

gel. A band corresponding in size to a digested plasmid was excised and purified using the QIAquick Gel Extraction Kit. Concentration of purified plasmid was again determined from A_{260}. Stoichiometry of the vector/insert DNA was calculated using the formula:

$$n(\text{vector})/n(\text{insert}) = [\gamma(\text{vector})/\gamma(\text{insert})] \cdot [l(\text{insert})/l(\text{vector})]$$

The ligation reactions were always prepared by the scheme showed in table 11.

vector/insert ratio	3/1	1/1	1/3	Only vector (same volume as in 3/1 vector/insert ratio
Final concentration of T4 ligase buffer	1x	1x	1x	1x
T4 ligase	1 µl	1 µl	1 µl	1 µl
H_2O	0 µl	0 µl	0 µl	Same volume as insert in 3/1 vector/insert ratio

Table 10. The reaction scheme for ligation of restricted DNA fragments

After a 2h incubation at room temperature, *E. coli* DH5α cells were transformed with 5 µl of each ligation mix.

Transformation of E. coli DH5α and BL21-RIL strains

For transformation, a 50 µl aliquot of *E. coli* DH5α cells would be thawed on ice. Than, 5 µl of DNA would be added to the aliquot and incubated on ice for 20 min. The bacteria would first be exposed to a 42°C heat-shock for 45s and than incubated on ice for another 2 min. Afterward, 300 µl of SOC medium (pre-warmed to 37°C) would be added to the bacteria which would than recover for 1h at 37°C with shaking (220 rpm). Recovered bacteria were evenly spread across LB-agar plates with 100 µg ml^{-1} ampicillin. For transformation of *E. coli* BL21-RIL cells a modified protocol was used: first incubation on ice was shortened to 15 min and recovery was shortened to 30 min. Transformed cells were plated on LB-agar plates with 100 µg ml^{-1} ampicillin and 30 µg ml^{-1} chloramphenicol. In both cases, inoculated plates were incubated at 37°C for a minimum of 12 h.

Electrophoresis on agarose gels

DNA samples were analyzed by agarose gel electrophoresis. In agarose gels, molecules of agarose interweave to produce a gel of hydrogen-bonded filaments. Differences in macromolecular migration through an agarose gel in an electric field are primarily a consequence of differences in size, shape and charge of molecules. Since DNA molecules are mostly linear and have the same amount of negative charge per unit of mass (due to the repeating orthophosphate groups within the molecule), they are separated primarily by difference in size. Non-linearized plasmids are a notable exception, due to the effect of supercoiling on their shape.

Agarose gels were prepared by mixing agarose with TAE-buffer (40 mM Tris, 1 mM EDTA, titrated with concentrated hydrochloric acid to pH 8,0) and heating to ~100°C. When the agarose was fully melted in the TAE buffer, ethidium bromide was added to the mixture, gels were cast and left to cool down to room temperature. For analysis of DNA fragments coding for Sec31(372-411), 1.2% agarose gels were used. For all other procedures, 0.75% agarose gels were used. For analytical procedures, 1-5 µl of DNA sample was mixed with ~ 2 µl of loading buffer (2,5 % bromophenol blue, 10 mM EDTA pH 8,0, 3 % glycerol, 2,5 % xylene cyanol) and loaded onto gels. Analytical electrophoresis were run in TAE buffer at 135 V for 15 min. For gel extraction procedures, 25 µl of restriction product would be mixed with ~5 µl of loading buffer and loaded onto the gel. Electrophoresis for gel extraction procedures were run in TAE buffer at 100 V for 20 min. Gels were analyzed under UV light, λ=314 nm, which is the excitation wavelength for ethidium bromide bound to DNA. The *1 kb DNA Ladder (NEB)* and *100 bp DNA Ladder (NEB)* were used for determination of DNA fragment sizes.

DNA primer synthesis and construct sequencing

DNA primers were ordered from Integrated DNA Technologies, Coralville, Iowa, USA. Before use, primers were resuspended in water to a working concentration of 5 µM. Sequencing samples (5 µl plasmid DNA, 6 µl sequencing primer 5 µM , 11 µl H2O) were sent to the MGH DNA Core Facility, Cambridge, Massachusetts, USA, for sequencing.

Bacterial cultures

LB medium

LB (Lysogeny Broth) was used for growing all bacterial cultures:

5 g/l yeast extract

10 g/l tryptone

10 g/l NaCl

For LB-agar, 1.2% agar was added to the LB medium

Both the liquid LB medium and LB-agar were autoclaved prior to use.

Antibiotics for selective media: 100 μg ml^{-1} ampicillin

30 μg ml^{-1} kanamycin

SOC medium

SOC (Supra Optimal broth with Catabolite repression) was used for transformed cell recovery:

2% w/v bacto-tryptone

0.5% w/v bacto-yeast extract

10mM NaCl

2.5mMKCl

Autoclaved

0.4% glucose

Storage of bacterial cultures

For long term storage, samples of overnight cultures were mixed with equal amounts of glycerol (40% w/v) and stored at -80°C.

Protein purification

Protein expression

For protein overexpression, sterile LB media with 100 μg ml^{-1} ampicillin, 30 μg ml^{-1} kanamycin and 0.4% glucose were inoculated with a ~7h old preculture of BL21-RIL cells carrying one of the plasmids used in experiments. Cultures were grown at 30°C, 220 rpm, for ~3 h to an optical density of $OD_{600}\approx$ 0.6 before being transferred to 18°C, 220 rpm for 20 min. than, cells were induced with 200 μM of IPTG and left to grow over night at 18°C, 220 rpm. When grown to an optical density of $OD_{600}\approx$ 6, cells were pelleted by centrifugation (6 min, 6000g) and stored at -20°C until further processing.

Cell disruption

Cells were resuspended in resuspension buffer (50 mM potassium phosphate buffer pH 8.0, 500 mM NaCl, 20 mM imidazole, 3 mM BME) with stirring for a minimum of 20 min or until a homogeneous resuspension was obtained; 1 ml of resuspension buffer/100 OD_{600} (as determined before cell pelleting) was added to the pelleted cells. Cells were disrupted using a French press cell disrupter (Constant Systems). Immediately after lysis, 0.2 mM of PMSF was added to the lysate to inhibit serine-proteases. The lysate was fractionated by centrifugation (20 min, 8000g). Supernatant was collected for further use. All steps were done at 4°C to prevent protein denaturation.

Affinity chromatography on Ni-NTA agarose

Affinity chromatography is a group of methods for protein separation based on differential affinity of proteins for binding certain molecules. Affinity chromatography on Ni-NTA agarose is one such method that exploits preferential binding of imidazole (or molecules carrying an imidazole group, e.g. histidine) to Ni^{2+} ions. The Ni-NTA beads are composed of agarose carrying nitrilotriacetic acid (NTA), a chelating molecule that binds four out of six coordination bond sites on a nickel Ni^{2+} cation. The remaining two coordination sites are free to bind two histidine side chains from a His_6-tag, covalently bound to a protein of interest. When the column is washed in buffer with a low concentration of imidazole (a low concentration of imidazole is usually present to prevent unspecific binding of non-tagged proteins to the beads), proteins without a His_6-tag are washed away, while the tagged protein, together with any protein bound to it, remains attached to the Ni-NTA beads. The protein can than be eluted by washing the

Ni-NTA agarose in a buffer containing a high imidazole concentration. Imidazole competes with His_6-tags for binding nickel cations, thereby releasing the protein of interest. Any protein bound to the His_6-tagged protein will also co-elute in this step. Because the His_6-tag has a small molecular mass and is not charged at pH 7.5-8.0, it usually has no effect on protein structure. In some constructs, the His_6-tag is followed by a protease recognition motif and can be cleaved.

In this experiment, the Ni-NTA beads were first equilibrated with resuspension buffer (50 mM potassium phosphate buffer pH 8.0, 500 mM NaCl, 20 mM imidazole, 3 mM BME). Imidazole concentration in the resuspension buffer was too low to compete with His_6-binding, but high enough to prevent most of the non-specific binding. BME prevented protein oxidation and formation of disulfide bonds. The beads were mixed with the buffer and pelleted by centrifugation (3 min, 2000g). The process was repeated four times. The beads were than mixed with an equal volume of the resuspension buffer and added to the bacterial lysate supernatant. About 2 ml of Ni-NTA beads were added for each 1000 OD_{600}/(L of bacterial culture). The supernatant-Ni-NTA mix was incubated for ~45 min, with stirring, at 4°C. After the incubation, nickel was collected by centrifugation (3 min, 2000g). The beads were washed three times by mixing with resuspension buffer and centrifugation (3 min, 2000g). After the last wash, beads were mixed with an equal amount of resuspension buffer and loaded onto a plastic column (Pierce). Another volume of resuspension buffer was added for wash. The proteins were eluted with elution buffer (50 mM HEPES buffer pH 7.4, 250 mM NaCl, 250 mM imidazole, 3 mM BME). All steps were done at 4°C to prevent protein denaturation.

Dialysis and His_6-tag cleavage

Dialysis is a method used for selective exchange of molecules in a solution based on the difference in their rates of diffusion through a porous membrane. The difference in diffusion rates is a consequence of molecular size: molecules that are smaller than the pores of the dialysis membrane, such as small molecules and ions, can transverse the membrane freely while macromolecules, being significantly larger than the pores in the dialysis membrane remain in the dialysis bag.

Ni-NTA affinity purified protein samples were dialyzed in dialysis tubing (Thermo) to remove excess imidazole. Dialysis bags with the protein were put into 2L of dialysis buffer (20 mM HEPES buffer pH 7.4, 200 mM NaCl, 1 mM DTT, 0.1 mM EDTA) and left to stir overnight at 4°C. DTT in the buffer prevented protein oxidation and formation of disulfide bonds. EDTA is a chelating agent that inhibits protein degradation by binding potential cofactors for metalloproteases.

An additional 200 µl of PrescissionProtease was added into the dialysis bag if the protein sample was intended for crystallization experiments. PrescissionProtease is a fusion protein of glutathione-S-transferase (GST) and human rhinovirus (HRV) type 14 3C protease. PrescissionProtease specifically recognizes the core amino-acid sequence Leu-Phe-Gln-Gly-Pro and cuts between glutamine and glycine residues.

Ion exchange chromatography

Ion exchange chromatography is a group of methods for protein separation based on different charges. The stationary phase contains charged molecules (usually having a carboxy group or an amino group conferring negative or positive charge respectively) termed ion exchangers that preferentially bind molecules depending on their overall charge. Proteins that have the opposite overall charge to that of the stationary phase will bind the ion exchangers, while other molecules flow through the column. Elution can be achieved by changing either the pH or the ionic strength of the chromatography buffer. If the pH is changed, proteins will be eluted from the column as they reach their pI. If the ionic strength is increased, proteins will release the ionic exchangers due to competitive binding of ions.

PrescissionProtease was removed from the dialyzed protein solution by ion-exchange chromatography on HiTrap™ IEX SP Fast Flow columns (*GE Healthcare*). Before chromatography, columns were equilibrated with washing buffer (20 mM HEPES buffer pH 7.4, 1 mM DTT, 0.1 mM EDTA). After equilibration, samples were run through the column. The weakly charged sample protein

would be collected in the flow through while the more strongly bound His_6-tags would elute from the column with rising ionic strength, i.e. by gradually mixing the washing buffer flowing through the column with an increasing percentage of ion exchange elution buffer (20 mM HEPES buffer pH 7.4, 1 mM DTT, 0.1 mM EDTA, 2 M NaCl).

Gel filtration chromatography

Gel filtration chromatography is a method for separating molecules based primarily on their hydrodynamic volume. The stationary phase of the column is composed of small beads of different, chemically inert, polymers containing pores of different sizes. As various molecules flow through the column, smaller ones fit into the pores and flow slowly while larger molecules cannot fit into the pores and move through the column more quickly. The pore size is specified for a given column, but is far from exact. In columns with larger pores, molecules will move more slowly than in columns with smaller pores.

For purification of the Sec13/Nup145C(109-179) fusion construct and analysis of Sec13-Sec31(372-411) and Sec13(I12F)-Sec31(372-411) binding, *HiLoad 26/60 Superdex 75pg* columns (*GE Healthcare*) were used. For analysis of Sec13-Nup145C, Sec13(I12F)-Nup145C, Sec13-Sec31(1-748) and Sec13(I12F)-Sec31(1-748) binding *HiLoad 26/60 Superdex 200 pg* columns (*GE Healthcare*) were used. Fractions containing protein were analyzed by SDS PAGE and pooled together. Due to the variability in pore size and inherent flexibility of protein shape, the samples did not elute in sharp, narrow concentration peaks but rather in Gaussian shaped concentration curves covering several fractions. To reduce the possibility of sample contamination with misfolded proteins, fractions at the edges of concentration peaks were not pooled with the rest of the protein in samples intended for crystallization experiments.

Protein concentration determination

Protein concentration was determined by UV-spectrophotometry from absorbency at λ= 280 nm. Extinction coefficients for different proteins were calculated using ProtParam (*Expasy*) (table 11)

Protein sample	**Abs 0.1% (=1 g/l)**
Sec13/Nup145C(109-179)	2.11
Sec13 (with His_6 tag)	2.56
Sec13 (without His_6 tag)	2.81
Nup145C	1.01
Sec31(372-411)	1.64
Sec31(1-748)	1.30

SDS PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) is a method for protein separation in an electric field based on their molecular weight. A polyacrylamide gel is formed by a network of polyacrylamide molecules interconnected by N,N'-methylenebisacrylamide linkers. Pore size of a polyacrylamide gel is constant and is defined bya ratio of concentration of acrylamide molecules to the concentration of N,N'-methylenebisacrylamide molecules. SDS is an anionic detergent that denatures proteins by inserting its hydrocarbon tail into the hydrophobic interior of the protein while conferring a uniform, negative charge to the surface of the protein. Since all proteins in an SDS PAGE sample have equivalent shapes (spherical denatured globules) and mass/charge ratios (approximately 1.4 g of SDS binds 1g of protein with very little dependence on the protein type), differences in migration of proteins within the gel are mostly a consequence of the difference in molecular weight. When a proper molecular weight marker is used and a solution that is analyzed is relatively pure, SDS PAGE can serve as a semi-quantitative method for identification of proteins. In SDS PAGE, two gel types are used: a stacking gel that stacks all the macromolecules from a solution into a single, narrow band and a separating gel (positioned underneath the stacking gel), that separates this narrow band into separate bands based on molecular weight of component proteins.

In this experiment samples for SDS PAGE were prepared by mixing 5 µl of protein sample with 2 µl of sample buffer (50 mM Tris-HCl pH 6,8, 1 % (w/v) SDS, 5 mM BME, 5 % glycerol, 0,02 % BPB) and heated to 96°C for ~5 min. Alternatively, for visualization of lower concentrations of proteins, 20 µl of protein sample was mixed with 5 µl of sample buffer and heated to 96°C for ~5 min. The pellet from the fractionation of the cell lysate was resuspended in 20 ml of urea and homogenized using a Dounce homogenizer before taking samples for the electrophoresis.

Two types of separating gels were used; 10 % acrylamide gels (table 12) were used for analyzing all protein samples, additionally, 16.5% acrylamide gels (table 12) were used for analyzing Sec13-Sec31(372-411) and Sec13(I12F)-Sec31(372-411) complexes. In both cases, the gels were poured into *Mini PROTEAN III Electrophoresis Cell* casts (*BioRad*) and overlaid with isopropyl alcohol. When the separating gel hardened, stacking gel (table 12) was poured over it.

	Separation gels		Stacking gels
	10%	16.5%	5.3%
30% acrylamide/bisacrylamide	13.33 ml	22 ml	1.77 ml
Separating buffer	10 ml	10 ml	0 ml
Stacking buffer	0 ml	0 ml	2.5 ml
H_2O	To 40 ml	To 40 ml	To 10 ml
10% APS	400 µl	400 µl	100 µl
TEMED	40 µl	40 µl	10 µl

Table 12. compositions of SDS PAGE gels

Electrophoresis were run at 200 mV for 10 min followed by 250 mV for 35 min. Gels were fixed in a fixing solution (25% isopropyl alcohol, 10% acetic acid) for ~10 min before staining with staining solution (10% acetic acid, 100 mg/l Coomassie Brilliant Blue R250) for ~2 days. Gels were destained in 10% acetic acid for ~2 days before photographs were taken. The *Protein Marker, Broad Range (2-212 kDa) (NEB)* was used for determination of DNA fragment sizes.

Solution of the protein crystal structure

To determine the crystal structure of the Sec13/Nup145C(109-712) protein complex, crystals of the purified construct were obtained. Diffraction data was collected and an electron density map was calculated. After structural refinement, the protein structure was analyzed.

Crystal growth

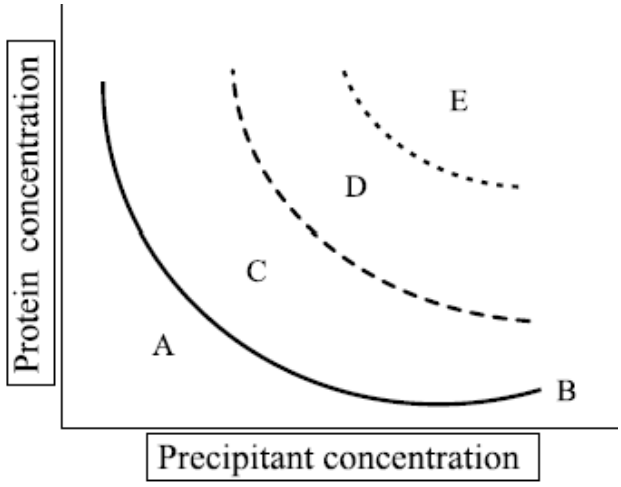

Figure 7. Four zones of crystal growth: **A** stable zone, **B** limit of solubility, when the solid and liquid phase of the protein are in equilibrium, **C** metastable zone, **D** labile zone, **E** precipitation zone (Bergfors, 2003)

Proteins are usually dissolved in an unsaturated solution containing water, buffer and salts. In an unsaturated solution, no crystals will nucleate, nor will crystals grow from preexisting seeds. In a crystallization experiment, a precipitant is added to the protein solution. The precipitant decreases the solubility of the protein, thereby increasing the saturation of the mixture. Depending on the level of saturation, four zones of crystal growth can be defined (Figure 7). In the stable zone, the solution is still unsaturated, as before the experiment. In the metastable zone, preexisting crystallization seeds will grow, but no new nucleation of crystals will occur. In the labile zone spontaneous crystal nucleation and growth happen. In the precipitation zone, the protein becomes completely insoluble and precipitates from the solution.

There are several approaches for setting a crystallization experiment. In vapor diffusion methods, a droplet containing the protein solution with added precipitant is placed in a closed system with a reservoir of pure precipitant. As the concentration of precipitant increases through diffusion from the reservoir, solubility of the protein gradually decreases.

A common method for executing vapor diffusion experiments is a hanging drop method in which a droplet containing the protein solution is suspended over a reservoir of the precipitant.

When crystals grown in the labile zone are not useful for structure determination, one of the approaches available is microseeding. In a microseeding experiment, a crystal is pulverized into

submicrometer sized fragments. These fragments are than transferred into a protein solution that is in the metastable zone, where they serve as seeds for crystal growth. Since no nucleation sites can spontaneously occur in a metastable solution, the number of crystals that can grow in such an experiment is limited by the number of seeds introduced. If the number of seeds is small enough, only a few crystals will grow. With less competition for dissolved protein, crystals can grow larger and in a more regular fashion. Since nucleation seeds are invisible, making determination of the exact amount of the seeds that are introduced into a solution impossible, a dilution series is made. In addition to seeding, there are numerous techniques for crystal quality improvement. Despite developments in understanding the process of protein crystallization, finding a proper condition for crystallization remains a trial and error process. Automated systems for setting crystallization experiments using commercially available screens make identifying initial conditions for crystal growth a much faster and less labor intense task.

In this experiment, initial crystals were grown by the hanging drop vapor diffusion method. A 0.5 µl drop of protein (γ= 85.3 mg ml^{-1}) was mixed with 0.5 µl of precipitant (26.5 % PEG 4K, 0.25 M LiCl, 0.1 M Tris pH 8.3). Crystals were grown at 16°C, over a reservoir containing 400 µl of precipitant. After three days, the crystals were prepared for microseeding using the *Seed Bead* (*Hampton*) seeding kit. The precipitant solution served as a stabilizing solution for the seeds. A series of dilutions (10^0-10^7) of the seed stock was made.

In a repeated crystallization experiment, 0.5 µl of different precipitant solutions (26.5 %-25% PEG 4K, 0.25 M LiCl, 0.1 M Tris pH 8.3) and 0.2 µl of seed dilutions were added to 0.5 µl drops of protein (γ= 85.3 mg ml^{-1}). Crystals were grown at 16°C, over reservoirs containing 400 µl of precipitant. Crystals from this experiment (grown over a precipitant containing 26% PEG 4K, 0.25 M LiCl, 0.1 M Tris pH 8.3) were used in a second round of seeding. The crystal used for obtaining the structural data was grown for three days from protein concentration (γ= 37.9 mg ml^{-1}), at 16°C using a precipitant solution containing 22% PEG 4K, 0.25 M LiCl, 0.1 M Tris pH 8.3. The crystal was collected using a nylon loop, dipped into 25% glycerol for cryopreservation and stored under liquid nitrogen.

Data collection

Diffraction data was collected by S. Brohawn and T. Schwartz at the Argonne National Laboratory Advanced Photon Source synchrotron facility, DuPage County, Illinois, USA, at the beamline 24-ID-C, operated by the Northeastern Collaborative Access Team (NE-CAT).

Structure determination and refinement

Diffraction data was processed using the HKL-2000 software package. Phases were solved using the CCP4i software package. PDB file 2pm6 was used as a template for molecular replacement. The electron density map was refined using the phenix.refine program package. Manual refinement was done using Coot and phenix.refine.

Tools for protein structure visualization and sequence analysis

Protein structures were visualized in PyMol and Coot. Sequencing results were evaluated in MacVector. EnzymeX was used for primer design and protein and DNA sequence analysis. Protein complexes observed in crystal structures were evaluated using PISA (*EBI*). Protein sequence conservation was analyzed by T-Coffee (*EBI*). Sequences of Nup145C from *Saccharomyces cerevisiae*, *Vanderwaltozyma polyspora*, *Homo sapiens* and of Sec13 from *Drosophila melanogaster*, *Aedes aegypti*, *Neurospora crassa*, *Arabidopsis thaliana*, *Homo sapiens* and *Saccharomyces cerevisiae* were obtained from the SwissProt database using Blast (NCBI) and SRS (EBI).

Results

Production of the crystallization construct

The Sec13/Nup145C(109-179) fusion construct that was used for crystallographic experiments was produced through iPCR. Success of PCR amplification was judged from agarose gel electrophoresis (Figure 8). The amplified sequence was further verified by DNA sequencing. The construct is composed of an amino-terminal His_6-tag followed by a full length yeast Sec13 protein and a carboxy-terminal insertion blade consisting of yeast Nup 145C amino acids 109 to 179. The His_6-tag is separated from Sec13 by a PrescissionProtease site. Sec13 is separated from the carboxy-terminal domain by an amino acid linker.

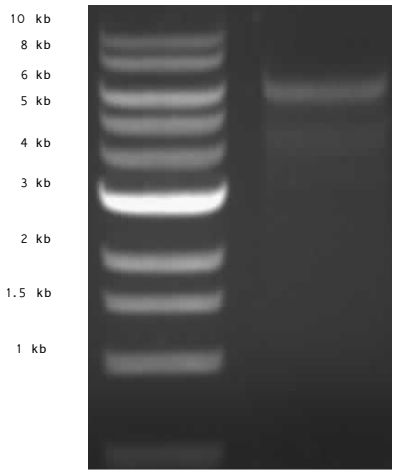

Figure 8. Result of iPCR production of the Sec13/Nup145C(109-179) fusion construct. Lane 1 is the DNA ladder. The band ~6 kb visible in lane 2 belongs to iPCR amplified construct and vector

Expression and purification of the crystallization construct

The Sec13/Nup145C(109-179) fusion protein was expressed in BL21-RIL cells. Induction of construct protein expression was successful. Majority of the protein was soluble in the lysis buffer. The fusion protein was purified via affinity purification on Ni-NTA agarose. Several contaminants were present in the eluate (Figure 9). The His_6-tag was removed from the protein by PrescissionProtease cutting and subsequent ion exchange chromatography.

Remaining contaminants were removed by gel filtration chromatography (Figure 10). Protein-rich fractions were chosen from A_{280} data and analyzed by SDS PAGE. Fractions containing the Sec13/Nup145C(109-179) fusion protein were identified by protein size, as judged from gel filtration chromatograms and SDS PAGE. These fractions were pooled and concentrated. Edge fractions of the peaks were left out of the pools to eliminate possible perturbations that cause the protein to travel on a gel filtration column differently from the most of the sample (Roe, 2001).

The pooled, concentrated protein samples were immediately used in crystallographic experiments as one day old samples produced crystals of significantly lower morphological quality (smaller, thinner, more irregular crystals, as judged from visual observation).

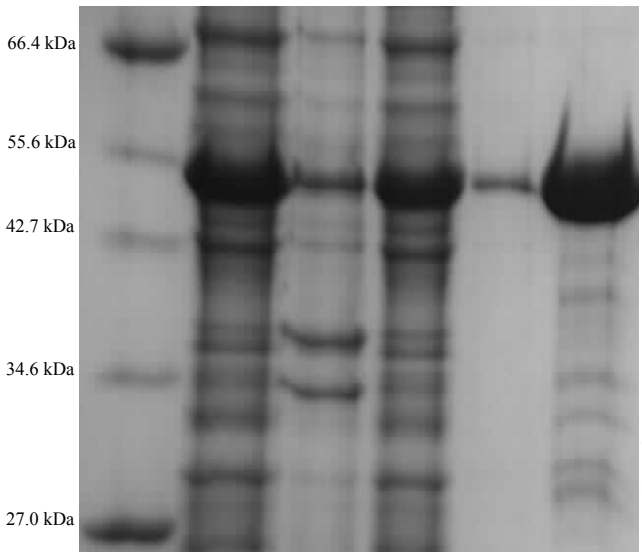

Figure 9. SDS PAGE of fractions from purification of the Sec13/Nup145C(109-179) fusion protein on NiNTA agarose. From left to right: marker, soluble fraction of the cell lysate, insoluble fraction of the cell lysate, column flow-trough, wash, elution. The construct forms the ~55 kDa band. Other bands belong to contaminants.

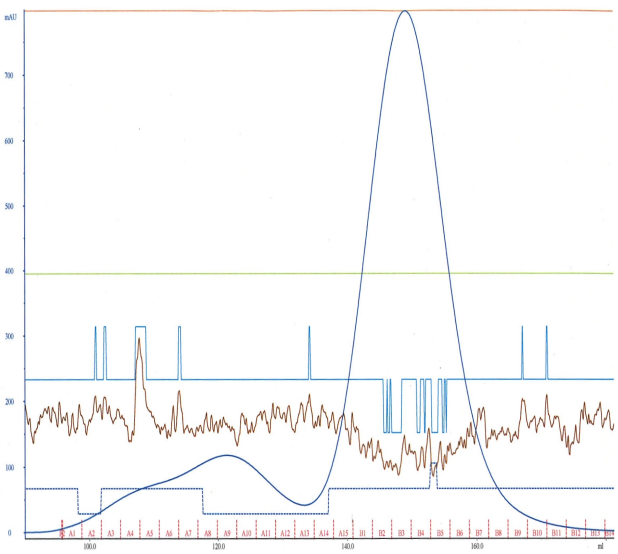

Figure 10. Gel filtration chromatogram of the Sec13/Nup145C(109-179) fusion protein. Fractions B1-B6 were pooled and concentrated for crystallographic experiments. A_{280} is in blue.

Crystallization experiments

Initial crystals were grown in the 'sitting drop' arrangement, from two kinds of solutions, both part of the commercially available PEGs II Suite (*Qiagen*) crystallization condition screen:

- A combination of plate-like crystals and proteinaceous spherulites grew at 16°C in a mix of 0.15 µL of an aqueous solution of 0.2 M Sodium acetate, 0.1 M Tris pH 8.5 and 30% PEG 4000 with an equal volume of the purified Sec13/Nup145C(109-179) fusion protein construct (γ = 8.8 mg mL^{-1}).
- Plate-like crystals grew at 16°C, in a mix of 0.15 µL of an aqueous solution of 0.1 M Lithium chloride, 0.1 M Tris pH 8.5 and 28% PEG 6000 with an equal volume of the purified Sec13/Nup145C(109-179) fusion protein construct (γ = 8.8 mg mL^{-1}).

The initial crystallization conditions were successfully reproduced in the 'hanging drop' arrangement, with 2 µL and 1 µL drops.

Crystals used for structure determination were grown in 1µL drops, from protein concentration (γ = 37.9 mg mL^{-1}), at 16°C , with a precipitant solution containing 22% PEG 4K, 0.25 mol dm^{-3} LiCl, 0.1 mol dm^{-3} Tris pH 8.3. A seeding procedure was used. These crystals were soaked in cryoprotectants and frozen in liquid nitrogen. A crystal soaked in 25% glycerol provided highest quality data.

Molecular structure determination

Phase determination and structural refinement

A data set of diffraction maxima of down to 2.45 Å resolution was obtained. The space group of the crystal was determined to be $P2_12_12$. A previously known structure, accessible in the protein data bank under code 2pm6, was used for molecular replacement. The structure was refined to $R_{working}/R_{free}$ values of 23%/28%. Crystal lattice properties and refinement statistics are summarized in table 13.

Crystal properties	
Space group	P 2_1 2_1 2
a, b, c	68.2Å, 93.8Å, 55.0Å
α, β, γ	90.00°, 90.00°, 90.00°
Data collection	
Resolution	50.00-2.45Å
Completeness	94.2%
Refinement	
Resolution	29-2.45Å
No. of unique reflections	12,723
$R_{working}/R_{free}$	23.4%/28.4%
No. of atoms Protein Water	2575 2557 18
Wilson B-factor (Å^2)	63.37
r.m.s.d. Bond lengths (Å) Bond angles (deg)	 0.003 0.701

Table 13. Lattice properties, data set statistics and refinment results.

Overall structure

A single molecule is present in the asymmetric unit. There are several contacts between molecules in adjoining cells (Figure 11). These contacts are presumed to be the reason behind relatively quick formation of crystals (fully-grown crystals were obtained one day after starting crystallographic experiments).

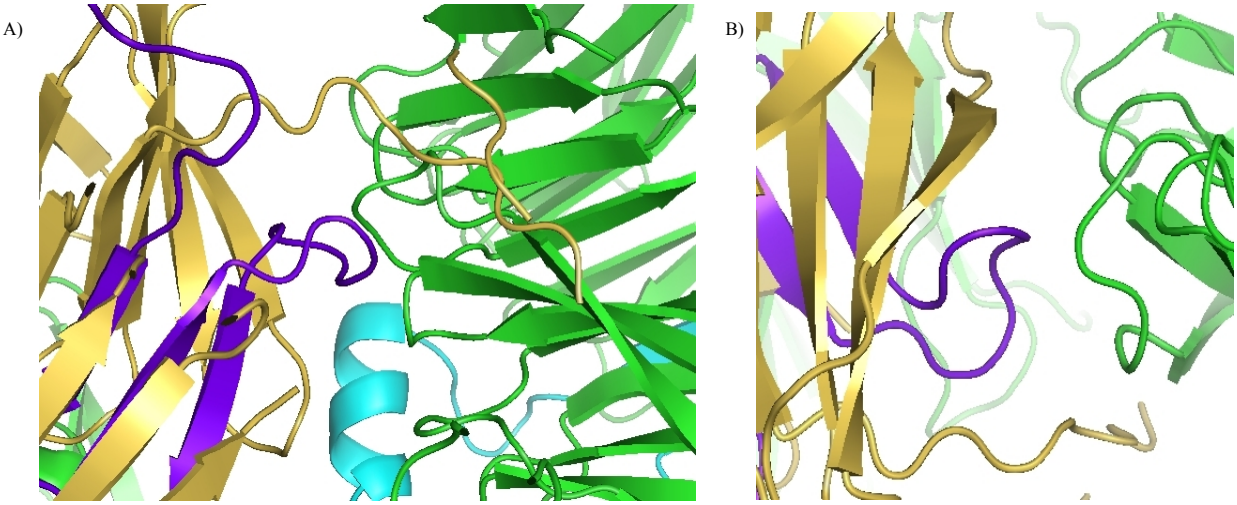

Figure 11. Contacts between neighboring molecules in the Sec13/Nup145C(109-179) crystal. A) Loop connecting Nup145C β-strands 2-3 (violet) binding loop connecting Sec13 β-strands 8-9(green). B) A close-up of A)

The molecule is an overall donut-shaped structure (height ~40Å, diameter ~70Å), with a central cavity (diameter ~15 Å) extending roughly halfway down the axes of the structure, followed by a water-filled central cavern. The opposite end of the molecule is indented toward

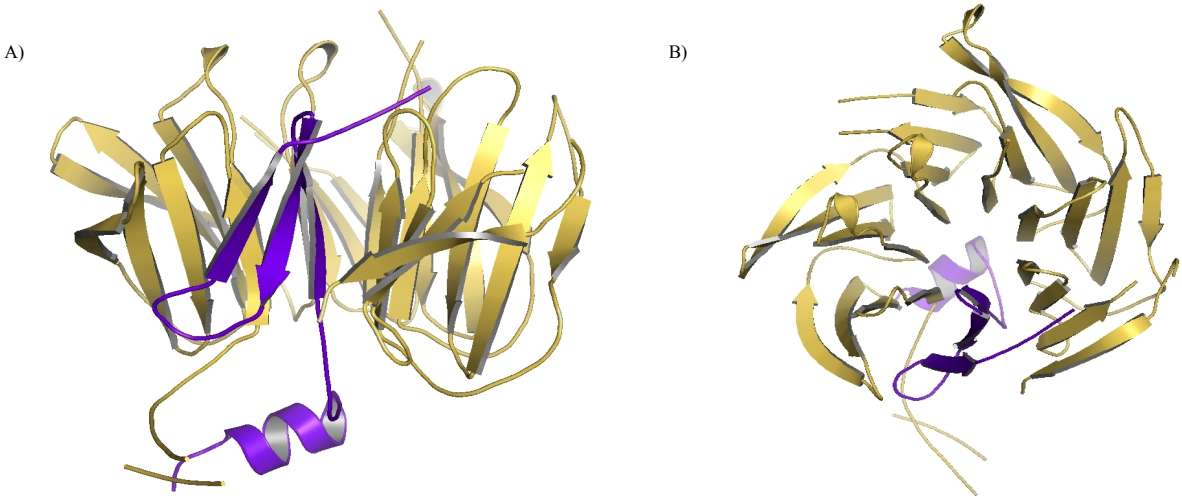

Figure 12. General shape of the Sec13-Nup145C (109-179) complex. A) Side view B) Top view. Sec13 is yellow, Nup145C is violet.

the central cavern. The six amino terminal residues of Sec13 protrude from the donut-like structure, perpendicularly to the axes. The four amino-acids preceding the amino-terminal methionine are linkers to the (removed) His_6-tag and are not a part of wild-type Sec13 in yeast. Another protrusion from the structure is formed by the amino-terminal α-helix of the Nup145C insertion blade. The α-helix protrudes from the molecule on the opposite side to the cavity, under a ~45° angle to the shorter axes (Figure 12).

The secondary structure of the molecule consists predominantly of β-strands, which form a fold that has been termed the seven bladed β-propeller (Berke et al., 2004). The six blades of this fold are each formed by four antiparallel β-strands of Sec13 (numbered 1-24, starting from the amino-terminus). The strands are connected in a β-meander-type supersecondary structure. The last β-strand of each meander is connected to the first β-strand of the next one by a loop. A seventh blade is formed by three antiparallel β-strands of the carboxy-terminal segment of the Nup145C insertion blade. This blade also assumes a β-meander-type supersecondary structure. The only extensive α-helical region is the amino-terminus of the Nup145C insertion blade. The amino-terminal α-helix and the carboxy-terminal β-strand of the Nup145C insertion blade are connected by a nine amino-acid long extended region lacking a defined secondary structure. The two faces of the donut-shaped structure are covered by loops and three short, α-helical regions.

Several features are notably absent from the structure. The nine amino-acid long loop linking the carboxy-terminus of Sec13 to the amino-terminus of the Nup145C insertion blade shows no density and, thus, could not be modeled into the structure. A five amino-acid long segment between Phe^{130} and Thr^{136} (between β-strands 11 and 12) and an eight amino-acid long segment between Tyr^{157} and Gly^{166} (between β-strands 13 and 14) of Sec13 are also absent from the electron density map. Since these regions are located between β-strands, they probably form highly mobile loops.

Surface properties, structural flexibility and sequence conservation

The surface of Sec13 bound to the Nup145C insertion blade features several charged regions. A negatively charged surface is formed by Sec13 amino acids Glu^{250}, Glu^{251}, Glu^{286} and peptide carbonyl oxygen atoms from Sec13 amino acids Leu^{248}, Gly^{287} and Trp^{289}. These amino acids are parts of loops connecting β-strands 20-21 (belonging to the fifth and the sixth blade of the Sec13 six-bladed β-propeller) and 23-24 (within the sixth blade). The amino-terminal α-helix of the Nup145C insertion blade is positively charged. Sec13 amino acids Lys^{244}, Lys^{245} and peptide amino groups from Sec13 amino acids Thr^{235}, Gln^{236} and Thr^{246} form a positively charged patch on β-strands 19 and 20 (fifth blade of the Sec13 six-bladed β-propeller). Similarly, Sec13 amino acids His^{43} and Lys^{44}, together with the peptide amino group of Sec13 amino acid Thr^{42} (β-strand 4) join the Nup145C amino acid Lys^{167} (part of the loop connecting Nup145C β-strands 2-3) in forming another positively charged patch.

The temperature factor, also called the B-factor, is a measure of the dynamic order in the crystal, caused by temperature-dependent vibration of the atoms within the molecules of the crystal. The B-factor is given in $Å^2$ and is proportional to structural flexibility (Drenth, 2007). Within the structure of the Sec13/Nup145C(109-179) fusion protein construct, observation of the B-factors for different regions of the molecule indicates that the central region of the β-propeller is fairly rigid, while the loops above and below the propeller are flexible. The β-strands that are in

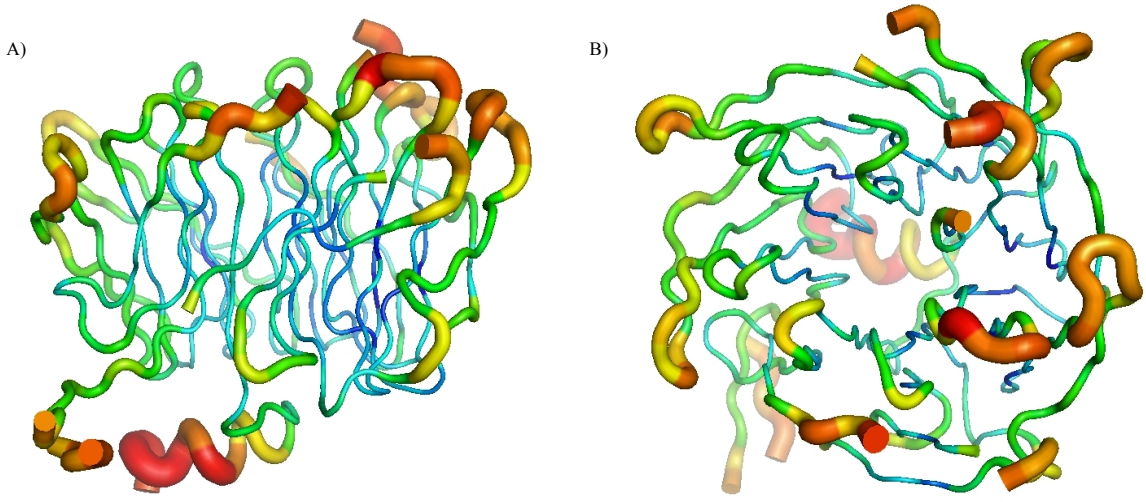

Figure 13. Flexibility of the Sec13-Nup145C (109-179) complex. Thicker amino acid chain indicates higher B-factor and, therefore, higher flexibility. A) Side view. B) Top view.

contact with the solvent (generally the fourth β-strand of each blade) are more flexible than the strands that are buried deeper within the protein. Several segments of the protein are especially flexible, notably the amino-terminal α-helix of the Nup145C insertion blade and loops connecting β-strands 17-18, 19-20 and the visible parts of loops connecting β-strands 11-12 and 13-14. These loops are all located on the same side of Sec13 and probably face the carboxy-terminal domain of Nup145C in yeast. Loops that are located proximally to the amino-terminal α-helix of the Nup145C insertion blade are as rigid as the interior regions of the protein (Figure 13).

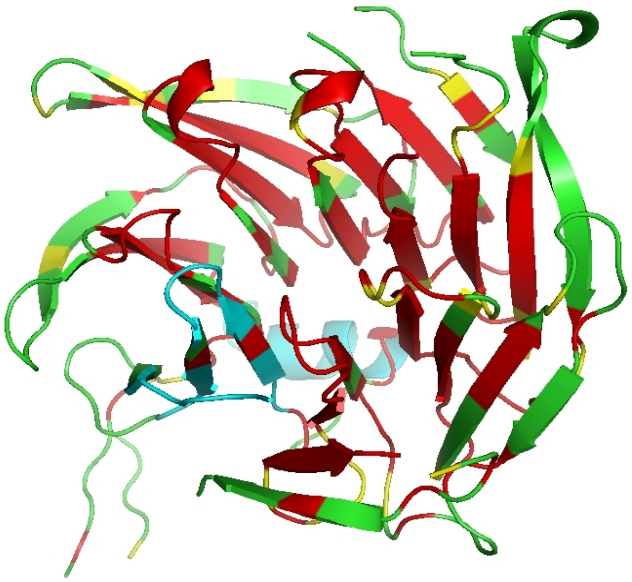

Figure 14. Conservation of the Sec13 and Nup145C insertion blade sequences. Conserved regions are red. Non-conserved regions are green for Sec13 and cyan for Nup145C. Partially conserved residues are yellow.

Plotting Sec13 sequence conservation on the Sec13/Nup145C(109-179) fusion protein structure revealed that the inner parts of the protein are more conserved than the solvent-contacting regions. Indeed, in all six blades, the outer β-strands (i.e. strands 4, 8, 12, 16, 20, 24) have few or no conserved amino acids, while the inner strands are conserved either fully or in most of their length. The only inner β-strands to have more than one non-conserved amino acid are strands 1 and 5 that have two non-conserved amino acids each. A significant part of these two strands (including the non-conserved amino acids) is in contact with the solvent in the central cavity. Loops connecting the β-strands show varying degrees of conservation. In general, loops proximal to the axes of the molecule tend to be better conserved than the distal loops.

Since the parts of the protein that are in proximity to solvent tend to be variable across phyla, general surface conservation is low. Large, conserved areas are generally those surrounding the Nup145C insertion blade, the surface facing the amino-terminal α-helix of the Nup145C insertion blade and the surface lining the entrance into the central cavity. A conserved patch is formed by loops connecting β-strands 1-2 and 5-6.

Overall, amino acid sequence of the Nup145C insertion blade is far less conserved across phyla than that of Sec13. Here too, the best conserved segment of the sequence forms the inner-most β-strand (strand 1), while the sequence of the outer β-strand (strand 3) is not conserved and has insertions in other species. The sequence of the amino terminal α-helix is not well conserved, although it lacks insertion mutations between species. Amino acids Arg^{142} and Arg^{143} of the amino terminal region are conserved between humans and yeast (Figure 14).

Protein binding interface

The Nup145C-Sec13 interface covers an area of 1869.5 Å^2 (calculated by PISA). Two salt bridges and 27 hydrogen bonds are predicted by PISA to be responsible for protein-protein interactions. Most of the hydrogen bonds are main-chain/side chain interactions. Several structurally bound water molecules are also visible in the interface, making hydrogen bonds to residues from both Nup145C and Sec13. An extensive network of hydrogen bonds is formed by a water molecule bound to Sec13 amino acids Asp^{18} and Tyr^{19} and Nup145C amino acids Ser^{154} and Thr^{155}.

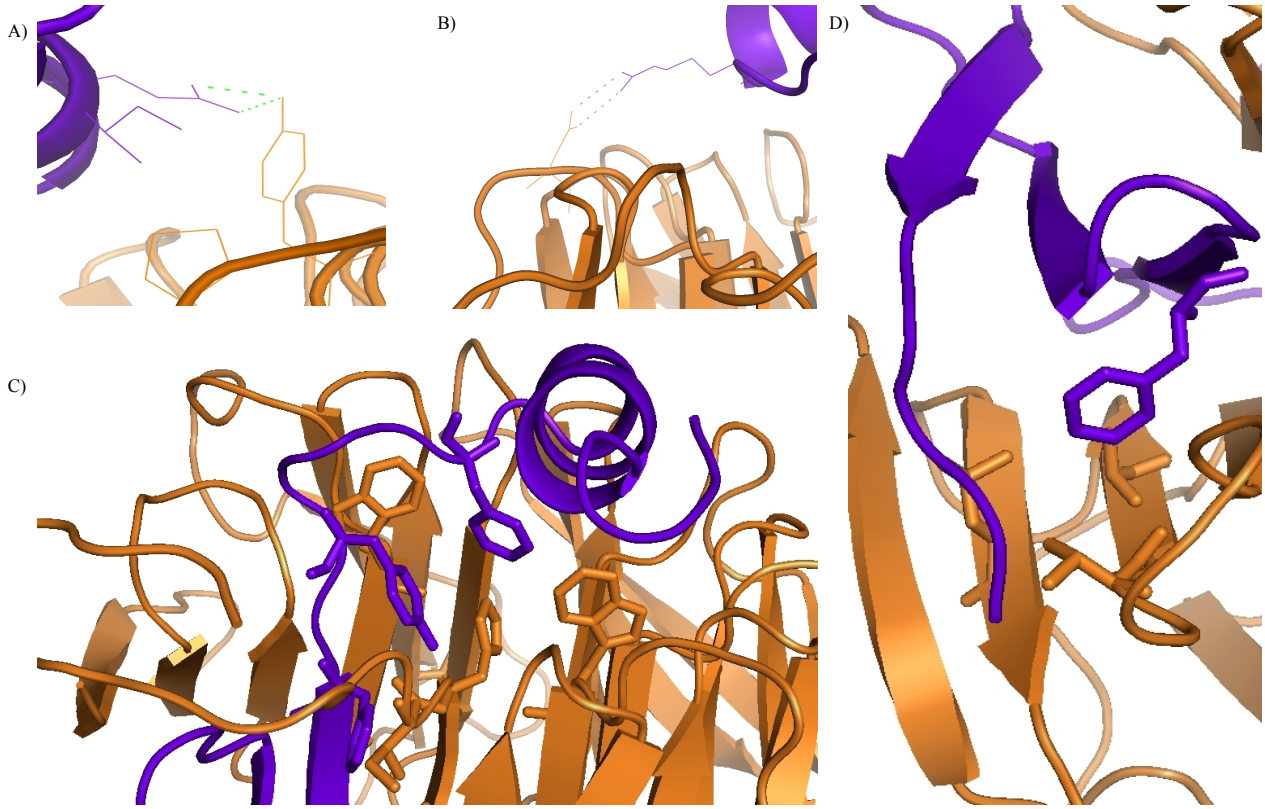

Figure 15. Contacts between Sec13 and Nup145C. A) Nup145C amino acid Ile^{138}, whose contacting sidechains of Sec13 amino acids Pro^{55} and Tyr^{75}. Sec13 Tyr^{75} side chain is hydrogen-bonded to Nup145C Glu^{141}. B) A salt bridge between the Nup145C amino acid Arg^{143} forms and Sec13 amino acid Asp^{205}. C) Nup145C amino acids Phe^{144}, Tyr^{148} and Phe^{150} binding side chains of Sec13 amino acids Ile^{12}, His^{13}, Ala^{15}, Trp^{57} and Trp^{258}. D) Nup145C amino acid Phe^{153} binding Sec13 amino acids Val^{268}, Ala^{270} and Leu^{280}. Sec13 is orange, Nup145C is violet.

Starting from the amino terminus of the Nup145C fragment, the first amino acid to make contact with Sec13 is Ile^{138}, whose side chain is immersed into a hydrophobic pocket formed by side chains of Sec13 amino acids Pro^{55} and Tyr^{75}. Sec13 Tyr^{75} side chain is also hydrogen-bonded to Nup145C Glu^{141}. Both Ile^{138} and Glu^{141} are part of the Nup145C amino-terminal α-helix.

Further downstream, the region linking the α-helical segment to the insertion blade of the Nup145C fragment is attached to Sec13 by a combination of salt bridges and hydrophobic interactions. The Nup145C amino acid Arg^{143} forms a salt bridge to Sec13 Asp^{205}. Three Nup145C amino acids, Phe^{144}, Tyr^{148} and Phe^{150} immerse their large, aromatic side chains into a hydrophobic environment created by side chains of Sec13 amino acids Ile^{12}, His^{13}, Ala^{15}, Trp^{57} and Trp^{258}

In the Nup145C insertion blade, Phe^{153} (part of β-strand 1) is immersed into a hydrophobic pocket created by Val^{268}, Ala^{270} and Leu^{280} of the sixth blade of the Sec13 β-propeller. On the opposite face of the same β-strand, Lys^{152} forms a salt bridge to Asp^{14} of the Sec13 first blade. Nup145C β-strands 2 and 3 are bound to the first and sixth β-propeller strands of Sec13 through hydrophobic interactions involving aliphatic side chains (Figure 15).

Comparison to known structures of Sec13

Previously solved structures of Sec13 include two structures of ySec13 in complex with ySec31 (Fath et al., 2007) and a structure of hSec13 in complex with yNup145C (Hsia et al., 2007). Additionally, a structure of yeast Sec13 paralog Seh1 bound to yNup85 has been solved (Brohawn et al., 2008) (Figure 16). Comparison of these structures to the structure of ySec13 in complex with a fragment of yNup145C revealed conserved features.

Structures of Sec13 bound to the amino terminal β-propeller and the carboxy-terminal domains of Sec31 reveal that Sec13 bound to the Sec31 insertion blade has the same overall shape as the Sec13/Nup145C construct. Some of the more notable differences are the absence of β-strand 20 (this segment of Sec13 lacks a defined secondary structure in Sec13/Sec31 structures) and a shorter β-strand 19 in the blade 5 of Sec13 bound to Sec31. Unlike the Sec13/Nup145C construct structure, the loop connecting β-strands 11-12 is clearly present, although with a high B-factor, indicating flexibility. Like in the Sec13/Nup145C structure, a large segment of the loop connecting β-strands 13-14 could not be modeled into the structure. All of these features are present in both structures of Sec13 bound to Sec31. The β-strand 24 that is completely absent in the structure of Sec13 bound to the amino terminal β-propeller domain of Sec31 and is very short in the structure of Sec13 bound to the carboxy-terminal domain of Sec31 is much longer in the Sec13/Nup145C structure.

Examination of the Sec13-Sec31 interaction interface reveals that the region linking the amino-terminal β-propeller domain to the insertion blade of Sec31 is attached to Sec13 by hydrophobic contacts that are arranged in a different manner than those in the Sec13-Nup145C complex. Tyr^{377} of Sec31 is inserted into a hydrophobic pocket formed by Sec13 amino acids Trp^{57} and Tyr^{75}. Sec13 amino acid Trp^{206} side chain is too far away from the interaction site to make contact with the insertion fragment of Sec31. The side chains of Sec13 amino acids Trp^{258} and His^{13} interact with Sec31 amino acids Gly^{378} and Pro^{280}.

The first β-strand of the Sec31 insertion blade is bound to the sixth blade of the Sec13 β-propeller through two hydrogen bonds that Sec31 Trp^{386} makes with Sec13 Ser^{261} and through hydrophobic contacts between Sec31 Trp^{386} and Val^{268}, Ala^{270} and Leu^{280} of the sixth blade of Sec13. The second and the third insertion blade of Sec31 bind Sec13 through nonspecific contacts between aliphatic amino acid side chains.

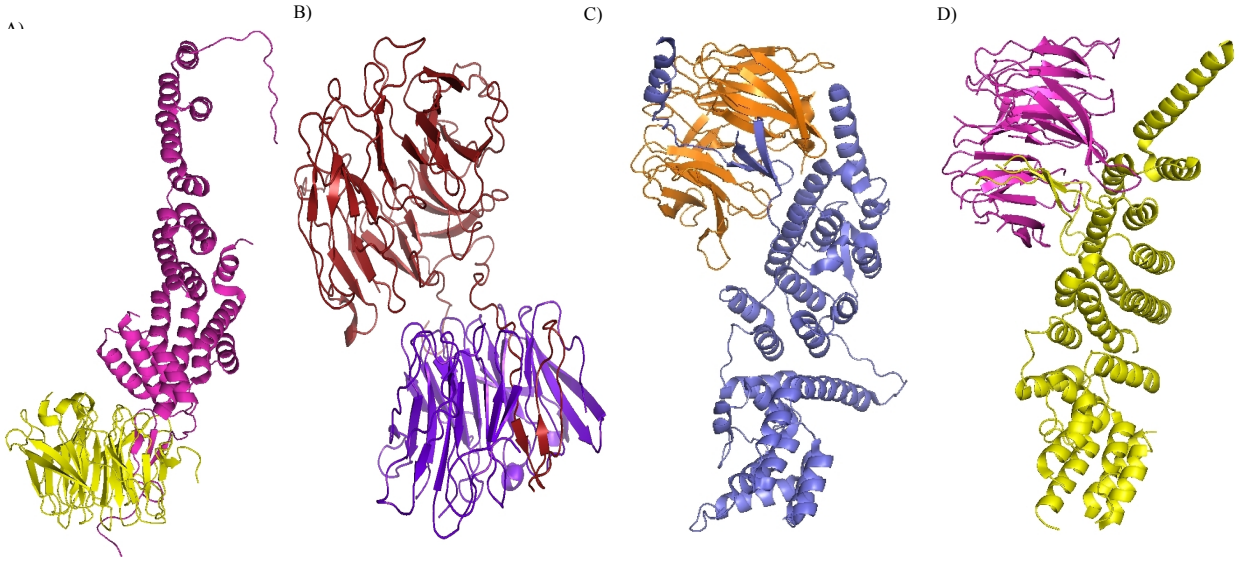

Figure 16. Structures of Sec13 bound to domains of Sec31. A) Sec13 (yellow) bound to the carboxy terminal domain of Sec31 (purple). B) Sec13 (violet) bound to the amino terminal domain of Sec31 (orange). C) Structure oft he human Sec13-yeast Nup145C hybrid. Sec13 is orange, Nup145C is violet D) Structure of yeast Nup85-Seh1 complex. Sec13 is purple, Nup85 is yellow

Analysis of structures of Sec13 bound to Sec31 in PISA revealed several ionic interactions. Two salt bridges are observed in the combined structures of the Sec13-Sec31 interactions. Sec13 Asp^{14} is bound Sec31 His^{385} and Sec13 Asp^{205} is bound to Sec31 Lys^{19}. An additional salt bridge is visible in the structure of Sec13 bound to the amino terminal β-propeller of Sec31: here, Sec13 Arg^{259} is coupled to Sec31 Glu^{379}. In the structure of Sec13 bound to the carboxy terminal domain of Sec31, the side chain of Sec13 Arg^{259} assumes a different rotamer and protrudes into the solvent.

Structure of the human Sec13/yeast Nup145C hybrid, again, displays the same overall donut-shaped structure as the yeast Sec13-Nup145C construct. A notable difference is the far lesser degree of secondary structure in the third and the fourth blade of the human Sec13 six bladed β-propeller: β-strand 16 spans four amino acids, β-strands 12 and 15 are composed of three amino acids each, while β-strands 13 and 14 consist of one amino acid each. In yeast Sec13 strands 16 and 15 span 9 amino acids, strand 14 spans 6, strand 13 spans 5 amino acids, while strand 12 consists of 4 amino acids.

A part of the loop spanning human Sec13 amino acids Trp^{158} to Phe^{182}, connecting β-strands 13 and 14 (corresponding to the yeast Sec13 loop spanning amino acids Trp^{153} to Lys^{172}) is missing from the structure, similarly to yeast Sec13. Compared to the yeast Sec13 loop 13-14, the loop in human Sec13 is five amino acids longer. Initial four amino acids of the loops are fully conserved between human and yeast.

Overlay of yeast Sec13 and Seh1 structures reveals significant structural conservation. Overall features of a six-bladed β-propeller are conserved in Seh1. The region corresponding to β-strand 1 in Sec13 does not assume a defined secondary structure in Seh1. Compared to yeast Sec13, β-strands of the third blade of Seh1 are shorter. The loop connecting β-strands 12 and 13 (corresponding to strands 13 and 14 in Sec13) is partially absent from the structure of Seh1.

Amino acid contacts in the Seh1-Nup85 interaction sites do not appear to be conserved between yeast Sec13 and Seh1. The area corresponding to the first β-strand of the Nup145C insertion blade does not assume a defined secondary structure in Nup85.

Amplification of Sec31 constructs and site-directed *in vitro* mutagenesis of Sec13

Based on the comparison of the crystal structure obtained in this project (ySec13 in complex with yNup145C) to known structures of the ySec13-ySec31 complex, a Sec13 ($AA12^{I\rightarrow F}$) mutant was designed. The Sec13 ($AA12^{I\rightarrow F}$) mutation was hypothesized to disrupt ySec13-yNup145C binding without affecting ySec13-ySec31 binding. This mutant was produced through site-directed in vitro mutagenesis. The gene coding for ySec13 was part of a bicistronic pETDuet-1 plasmid vector. The other available expression cassette contained a gene coding for either Nup145C or a fragment of Sec31. Due to difficulties in PCR amplification of full length ySec31, the $Sec31^{1-748}$ and $Sec31^{372-411}$ fragments were PCR-amplified from yeast genomic DNA and cloned into the plasmid vector. Results of PCR amplifications were judged from agarose gel electrophoresis. Construct sequences were verified by DNA sequencing.

Coexpression and copurification of protein Sec13 with Nup145C, Sec31$^{1-748}$ and Sec31$^{372-411}$

To confirm the hypothesis that the Sec13 (AA12$^{I\rightarrow F}$) mutation selectively disrupts ySec13-yNup145C binding, either wild type Sec13 or Sec13 (AA12$^{I\rightarrow F}$) were coexpressed with full length Nup145C, Sec31$^{372-411}$ (Sec31 amino acids 372-411) or with Sec31$^{1-748}$ (Sec31 amino acids 1-748). Binding efficiency was judged from protein copurification on Ni-NTA agarose and from gel filtration chromatography.

For protein copurification on Ni-NTA agarose, the His_6-tag was attached to the Sec13 amino terminus in Sec13-Sec31$^{372-411}$ copurifications and to the Sec31 amino terminus in Sec13-Sec31$^{1-748}$ copurifications. In Sec13-Nup145C copurifications, two different arrangements were used, with the His_6-tag attached to either the Sec13 or Nup145C amino termini.

Induction

Induction efficiency was judged from SDS PAGE analysis of culture samples taken before and after induction with 200 μmol dm^{-3} of IPTG. Both wild type Sec13 and Sec13 (AA12$^{I\rightarrow F}$) had high expression regardless of the coexpression partner. Expression of Nup145C was high when coexpressed with wild type Sec13 both when the His_6-tag was attached to Sec13 and to Nup145C. When Nup145C was coexpressed with His_6-tagged Sec13 (AA12$^{I\rightarrow F}$) expression of Nup145C was detected. The quality of SDS PAGE gels of samples containing other protein combinations was too low for direct detection of Sec31$^{1-748}$ or Nup145C expression from cell lysate. Sec13 expression was clearly detectable in these samples. Expression of Sec31$^{372-411}$ could not be tracked by SDS PAGE of cell lysates due to the small size of the peptide.

Solubility

Both wild type Sec13 and the Sec13 ($AA12^{I\rightarrow F}$) mutant were present predominantly in the soluble fraction of the cell lysate. When coexpressed with wild type Sec13, Nup145C was also soluble. Nup145C was found predominantly in the insoluble fraction of the cell lysate when coexpressed with Sec13 ($AA12^{I\rightarrow F}$) (Figure 17). Solubility of $Sec31^{1-748}$ and $Sec31^{372-411}$ could not be monitored because of the low quality of SDS PAGE results.

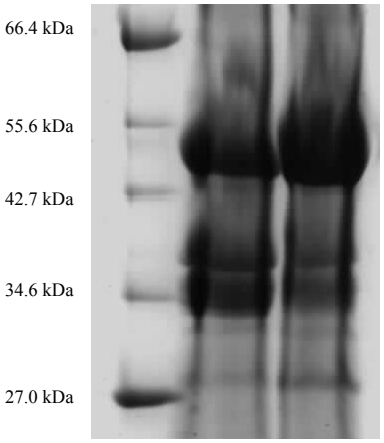

Figure 17. Insoluble fractions of wild type Sec13 copurified with His_6-tagged Nup145C (lane 2) and of Sec13 ($AA12^{I\rightarrow F}$) copurified with His_6-tagged Nup145C

Copurification on Ni-NTA agarose

Nup145C copurified readily with His_6-tagged wild type Sec13. Copurification of Nup145C with His_6-tagged Sec13 ($AA12^{I\rightarrow F}$) was also observed, but with a far lower ratio of the Nup145C amount to the amount of Sec13 (Figure 18).

Wild type Sec13 copurified with His_6-tagged Nup145C. A marked decrease in size was seen in His_6-tagged Nup145C coexpressed with wild type Sec13 when compared to other SDS PAGE gels of Nup145C purifications. The stained band of Nup145C shifted from a position corresponding to ~81 kDa to a new position corresponding to ~42 kDa in size. This can be explained by prolonged exposure to protease digestion, due to the particular conditions during the experiment. Sec13 ($AA12^{I\rightarrow F}$) copurified with His_6-tagged Nup145C to a much lesser extent (Figure 19).

Expression of $Sec31^{1-748}$ was too low for direct analysis of copurification. Similarly $Sec31^{372-411}$ was too small to be visualized on standard 10% SDS PAGE gels.

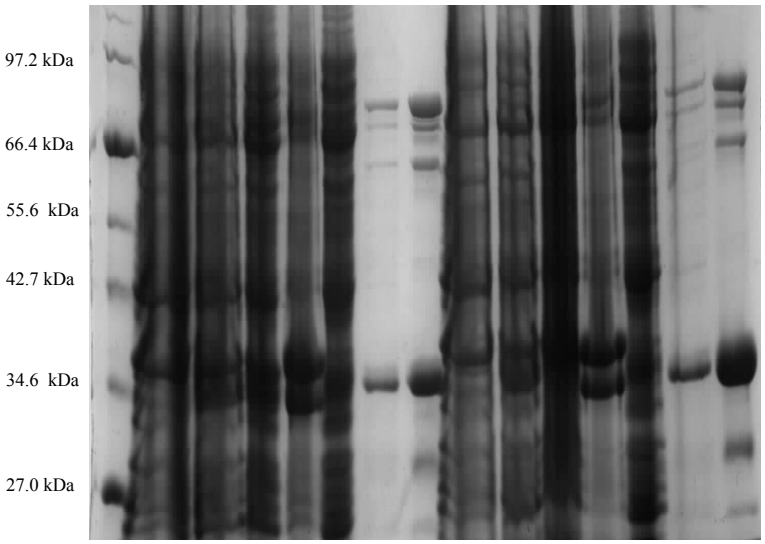

Figure 18. Coexpression and copurification of Nup145C with His_6-tagged wild type Sec13 (lanes 2-8) and Sec13 ($AA12^{I\rightarrow F}$) (lanes 9-15). Lane 1: marker; lanes 2, 9: uninduced bacterial culture; lanes 3,10: induced bacterial culture; lanes 4, 11: soluble fraction of the cell lysate; lanes 5, 12: insoluble fraction of the cell lysate; lanes 6, 13: Ni-NTA column flow trough; lanes 7, 14: Ni-NTA column wash fractions; lanes 8, 15: Ni-NTA column eluates

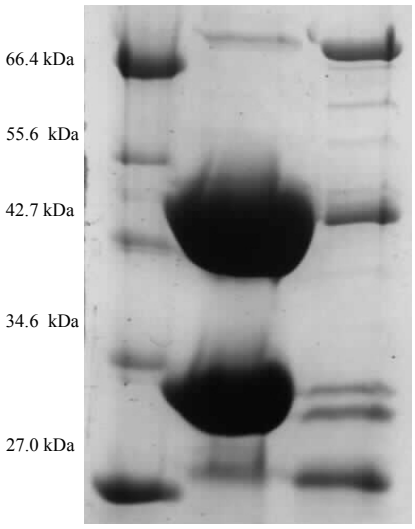

Figure 19. Ni-NTA column eluates of wild type Sec13 copurified with His_6-tagged Nup145C (lane 2) and of Sec13 ($AA12^{I\rightarrow F}$) copurified with His_6-tagged Nup145C

Gel filtration

Gel filtration chromatography provided further data on protein binding and expression levels. In the chromatogram of Ni-NTA agarose eluate containing His_6-tagged wild type Sec13 copurified with Nup145C, three peaks were visible. A major peak, appearing at ~220 ml of elution volume was shown by SDS PAGE to contain one protein, corresponding in size to Sec13. A minor peak appearing at ~160 ml of elution volume was shown by SDS PAGE to contain two proteins. One protein corresponded in size to Sec13 while the other corresponded to Nup145C. A third peak, partially merged with the minor peak, appeared at ~175 ml of elution

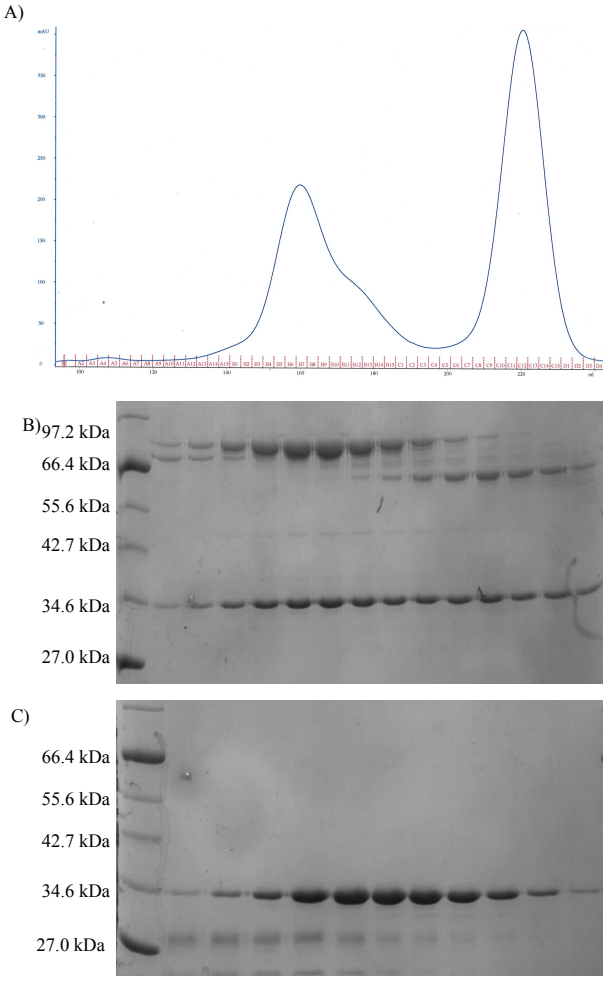

Figure 20. A) Gel filtration chromatogram of of Ni-NTA agarose eluate containing His_6-tagged wild type Sec13 copurified with Nup145C B) SDS PAGE of fractions B2-B15 C) SDS PAGE of fractions C7-D2; Lane 1 is the protein marker in both B and C

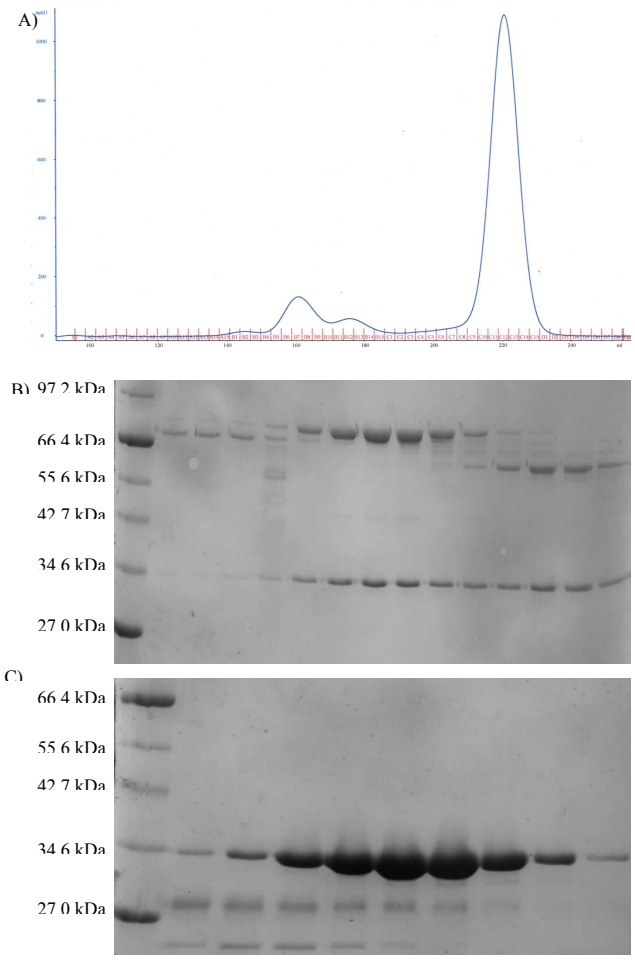

Figure 21. A) Gel filtration chromatogram of chromatogram of Ni-NTA agarose eluate containing His_6-tagged Sec13 ($AA12^{I\rightarrow F}$) copurified with Nup145C B) SDS PAGE of fractions B1-B14 C) SDS PAGE of fractions C8-D1; Lane 1 is the protein marker in both B and C

volume. Fractions belonging to this peak were shown by SDS PAGE to contain two proteins. One protein, again, corresponded in size to Sec13. The other protein was estimated to have a molecular mass of approximately 55 kDa. This protein is believed to be a degradation product of Nup145C, due to the numerous bands of intermediate sizes gradually appearing in the fractions corresponding to the transition between the 160 ml and the 175 ml peaks (Figure 20).

The three peaks that were observed in the chromatogram of Ni-NTA agarose eluate containing His_6-tagged wild type Sec13 copurified with Nup145C could also be detected in the chromatogram of Ni-NTA agarose eluate containing His_6-tagged Sec13 ($AA12^{I\rightarrow F}$) copurified with Nup145C. The relative intensities of the peaks were, however, different (Figure 21). The peak appearing at ~175 ml of eluate was better resolved from the ~160 ml. The compositions of the three peaks, as examined by SDS PAGE, were the same as in the peaks observed from the chromatogram of Ni-NTA agarose eluate containing His_6-tagged wild type Sec13 copurified with Nup145C.

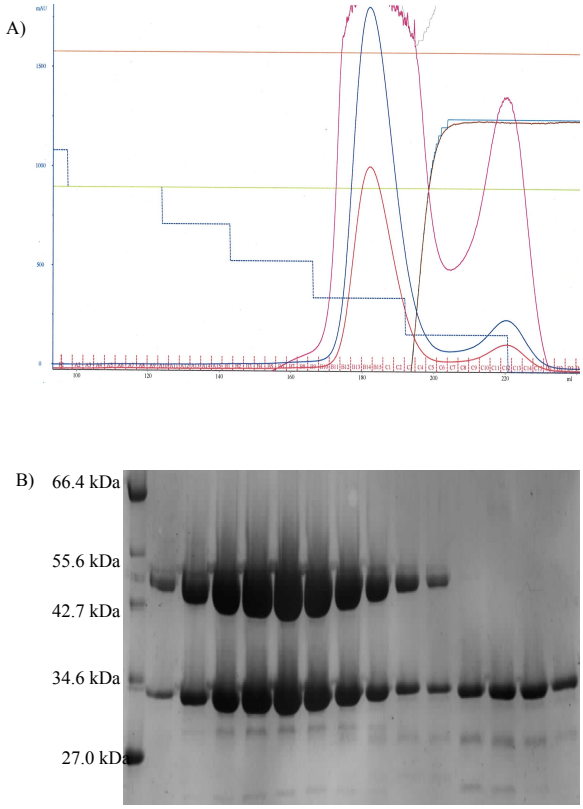

Figure 22. A) Gel filtration chromatogram of Ni-NTA agarose eluate containing His_6-tagged Nup145C copurified with wild type Sec13 B) SDS PAGE of fractions B11-C5, C11-C14

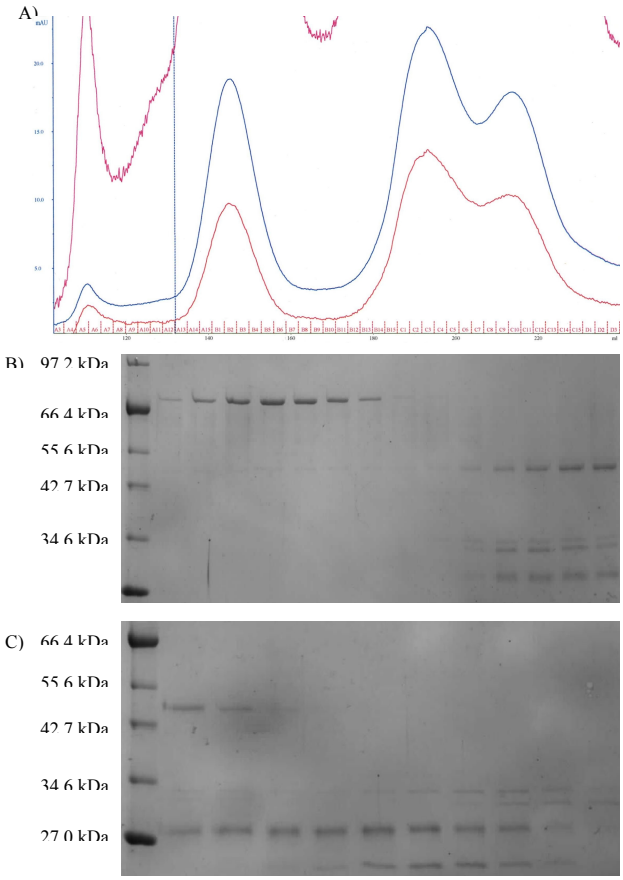

Figure 23. A) Gel filtration chromatogram of Ni-NTA agarose eluate containing His_6-tagged Nup145C copurified with Sec13 ($AA12^{I\rightarrow F}$) B) SDS PAGE of fractions A14-B5, B13-C4, C) SDS PAGE of fractions C5-C14; Lane 1 is the protein marker in both B and C

In the chromatogram of Ni-NTA agarose eluate containing His_6-tagged Nup145C copurified with wild type Sec13, two peaks could be observed. A major peak appeared at ~180 ml of elution volume. This peak was shown by SDS PAGE analysis to contain two proteins: one corresponding in size to Sec13 and a ~45 kDa protein. The ~45 kDa protein is believed to be a Nup145C degradation product, as the same protein was observed during purification on Ni-NTA agarose (Figure 19). Fractions taken from a second, minor peak that could be observed at ~220 ml of elution volume contained only one protein that corresponded to Sec13 in molecular mass as estimated from SDS PAGE analysis (Figure 22).

The chromatogram of Ni-NTA agarose eluate containing His_6-tagged Nup145C copurified with Sec13 ($AA12^{I\rightarrow F}$) had three peaks. All of these were of low intensity and contained proteins with several different molecular masses, presumed to be impurities(Figure 23).

Four peaks could, be observed in the chromatogram of Ni-NTA agarose eluate containing His_6-tagged $Sec31^{1-748}$ copurified with wild type Sec13. A major peak appeared at ~110 ml of elution volume. Fractions corresponding to this peak were shown by SDS PAGE to contain two proteins with estimated molecular masses of ~85 kDa and ~33 kDa. These agree well with molecular masses of 83.5 kDa for Sec31 and 33 kDa for Sec13. The three minor peaks were shown to contain proteins of several different molecular masses, presumed to be contaminants from copurification on Ni-NTA agarose. A peak that appeared at 220 ml of elution volume contained a protein with a molecular mass of ~33 kDa, believed to be Sec13 (Figure 24).

The chromatogram of Ni-NTA agarose eluate containing His_6-tagged $Sec31^{1-748}$ copurified with Sec13 ($AA12^{I\rightarrow F}$) was quite similar to that of His_6-tagged $Sec31^{1-748}$ copurified with wild type Sec13. The only notable difference was a fifth peak that appeared at ~195 ml of elution volume and was partially merged with the peaks appearing at ~180 ml and ~220 ml of elution volume (Figure 25).

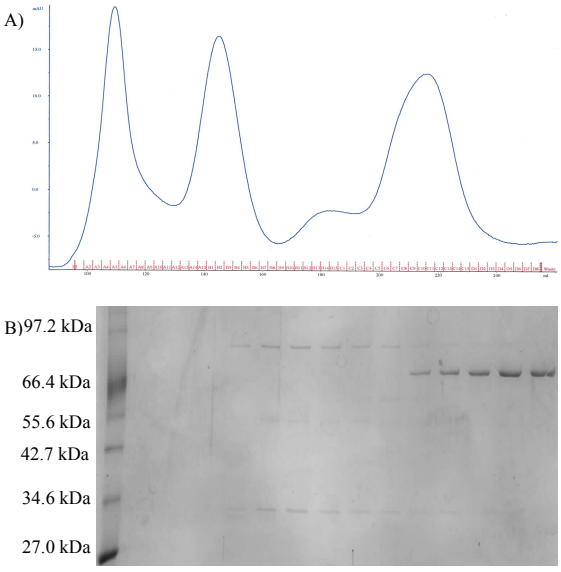

Figure 24. A) Gel filtration chromatogram of Ni-NTA agarose eluate containing His_6-tagged $Sec31^{1-748}$ copurified with wild type Sec13 B) SDS PAGE of fractions A1-A9,

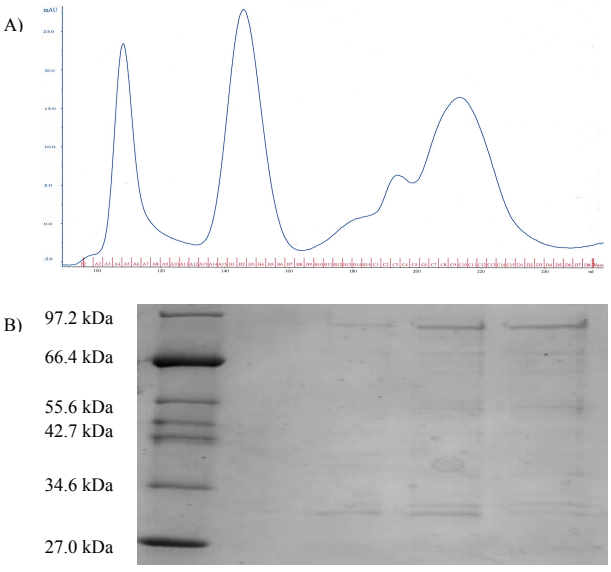

Figure 25. A) Gel filtration chromatogram of Ni-NTA agarose eluate containing His_6-tagged $Sec31^{1-748}$ copurified with Sec13 ($AA12^{I\rightarrow F}$) B) SDS PAGE of fractions A3-A6. Lane 1 is the protein marker

The gel filtrations of Ni-NTA agarose eluates containing His_6-tagged $Sec31^{372-411}$ copurified with wild type Sec13 or Sec13 ($AA12^{I\rightarrow F}$) served for assessment of the ability $Sec31^{372-411}$ of to bind Sec13. The chromatogram of Ni-NTA agarose eluate containing His_6-tagged $Sec31^{1-748}$ copurified with wild type Sec13 displayed a single peak at ~160 ml of elution volume. SDS PAGE analysis on 16.5% polyacrylamide gels showed that the fractions corresponding to this peak contained

three protein species, appearing in two peaks of protein concentration (as estimated from SDS PAGE). A protein with an estimated molecular mass of ~26 kDa formed a minor peak that appeared at ~150 ml of elution volume. A second protein species with an estimated molecular mass of ~33 kDa (close to the molecular mass of Sec13), was present in fractions collected from ~145 ml to ~180 ml of elution volume and formed a peak at 160 ml of elution volume. A third protein species was detectable in fraction collected from ~155 ml to ~165 ml of elution volume, reaching a peak concentration at ~160 ml of elution volume. This protein species had a molecular mass bellow 6.5 kDa, in frame with the size of the Sec31$^{372\text{-}411}$ fragment (4.3 kDa) (Figure 26).

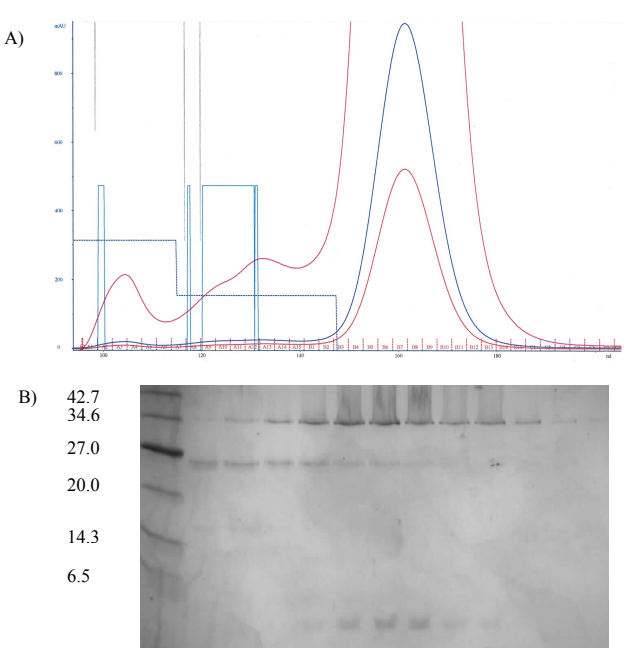

Figure 26. A) Gel filtration chromatogram of Ni-NTA agarose eluate containing His$_6$-tagged wild type Sec13 copurified with Sec31$^{372\text{-}411}$ B) SDS PAGE of fractions B2-B13. Lane 1 is the protein marker

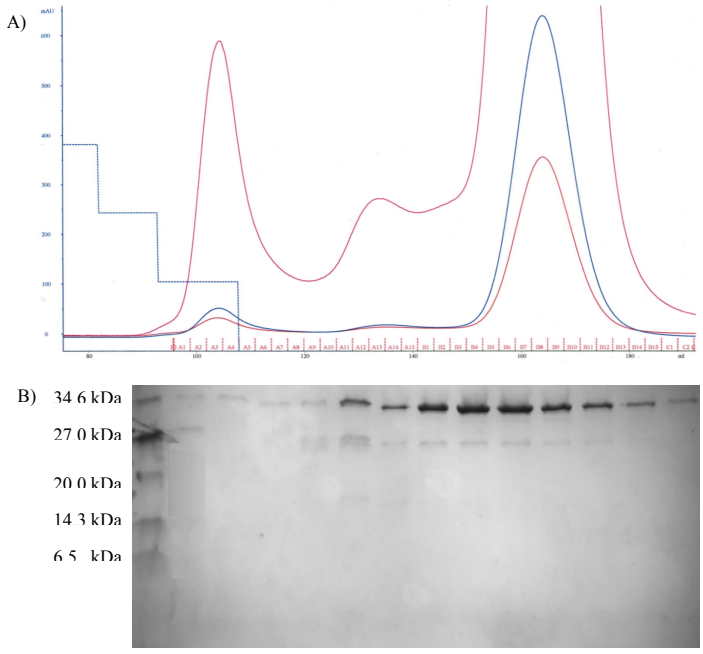

Figure 27. A) Gel filtration chromatogram of Ni-NTA agarose eluate containing His$_6$-tagged Sec13 (AA12$^{I\rightarrow F}$) copurified with Sec31$^{372\text{-}411}$ B) SDS PAGE of fractions A2-A4, B4-B13. Lane 1 is the protein marker

In the chromatogram of Ni-NTA agarose eluate containing His$_6$-tagged Sec31$^{372\text{-}411}$ copurified with Sec13 (AA12$^{I\rightarrow F}$), two peaks could be observed. A major peak appeared at ~165 ml of elution volume. Eluate fractions belonging to this peak were shown by SDS PAGE to contain a ~33 kDa protein, corresponding to Sec13. A second protein that produced much weaker bands at ~26 kDa is believed to be a contaminant from purification on Ni-NTA agarose. Some of the samples forming the peak displayed additional bands in SDS PAGE gels. These bands did not appear in neighboring samples and were therefore disregarded from analysis (Figure 27).

Discussion

Crystal structure determination

Expression of Sec13/Nup145C(109-179) fusion protein was sufficient for performing crystallization experiments. A combination of His_6-tagged protein affinity purification on Ni-NTA agarose combined with gel filtration chromatography provided Sec13/Nup145C(109-179) fusion protein samples of high purity for crystallization experiments. During cell lysis and subsequent protein affinity purification, a buffer containing 20 mM imidazole was used ,as protein binding to Ni-NTA agarose was significantly less efficient in buffers containing higher concentrations of imidazole.

Seeding experiments proved useful for improvement of crystal morphology and increase in size. Full-grown crystals were usually observed one day after setting of the experiment. Typically, a membranous layer enveloping the drop containing a mix of protein and precipitant could be seen. Protein crystals grew attached to this layer. During removal of crystals from drops, this layer tore into peaces that rapidly consolidated to form additional protein crystals.

A dataset with diffraction maxima up to a resolution of 2.45 Å was obtained. This was deemed sufficient for comparative analysis of the Sec13-Nup145C interface. Since two structures of yeast Sec13 have already been solved, the molecular replacement method was used for phase determination. Modeling of protein sequence into the electron density map was successful for most parts of the protein. Exceptions were two regions presumed to be highly mobile loops, the region linking the carboxy terminal part of Sec13 to the amino terminus of Nup145C and an amino terminal segment of the construct. Several lysine side chains could not be straightforwardly modeled into the electron density map.

The molecular structure revealed several inter-protein contacts. These were indicated by PISA to be crystal contacts, probably without physiological significance. These results are supported by the fact that the Sec13/Nup145C(109-179) fusion protein behaved as a monomer during gel filtration. Protein contacts might be one of the reasons for rapid formation of protein crystals.

The structure of the construct is a seven bladed β-propeller, having a donut-like shape, with a central cavity extending from the side of the protein facing the amino-terminus of Nup145C, roughly half-way along the axes of the structure. Further down the central axes, an isolated, solvent filled cavern is found. Six blades of the propeller, composed of four β-strands each, are parts of Sec13. The seventh blade contains three β-strands and is provided by the carboxy terminal region of the Nup145C insertion fragment. The only α-helical part of the structure is the amino terminal α-helix of the Nup145C fragment. The α-helix and insertion blade of Nup145C are linked by a seven amino acid long region lacking a defined secondary structure.

Exceptionally conserved regions of the protein tend to be those that are structurally or functionally significant (Gruber et al., 2007). Analysis of sequence conservation reveals that solvent-exposed segments of the β-propeller blades tend to be less conserved, resulting in a surface mostly lacking large conserved regions. The area surrounding the Nup145C insertion blade, the surface facing the amino-terminal α-helix of the Nup145C domain invasion motif and the surface lining the entrance into the central cavity form the only continuous conserved areas.

Large, charged patches of protein surface may indicate functionally important sites (Voet and Voet, 2004). Surface charge distribution for the Sec13/Nup145C(109-179) fusion protein is roughly uniform, with large neutral regions interrupted by a roughly equal number of positively and negatively charged patches.

The temperature factor (B-factor) is proportional to structural flexibility and is thus an important indicator of dynamic regions within protein structures (Drenth, 2007). An examination of the B-factor distribution reveals that the interior regions of the protein construct are more rigid than the segments on the periphery of the structure. Several parts of the structure, including the Nup145C amino terminal α-helix and several loops, are exceptionally mobile. Observation of the B-factors for loops that could not be fully modeled into the structure (i.e. loops linking β-strands 11-12 and 13-14) revealed them to be highly mobile, suggesting that the lack of electron density for these loops was indeed a consequence of high mobility.

Nup145C is bound to Sec13 through heterotypic interactions, including two salt bridges, numerous hydrogen bonds and hydrophobic contacts. Several structurally bound water molecules are visible in the interface.

Comparison of the Sec13-Nup145C interface to the Sec13-Sec31 interface revealed differences in binding of Sec13 to its two interacting partners. While the hydrogen bonds and electrostatic interactions binding Sec13 to Sec31 are fairly similar to those in the Sec13-Nup145C

interface (e.g. Sec13 amino acids Asp205 and Asp14 form salt bridge pairs with similarly positioned positively charged amino acids in both complexes), positions and types of amino acids involved in hydrophobic contacts do not appear to be conserved.

Sec13 mutant design and binding analysis

Lack of conservation between the Sec13-Sec31 and the Sec13-Nup145C hydrophobic contacts was used for design of a mutant that would specifically disrupt the Sec13-Nup145C interactions, without having an effect on Sec13-Sec31 binding. A mutation that changed the Sec13 amino acid Ile^{12} into a phenylalanine, Sec13 ($AA12^{I\rightarrow F}$), was hypothesized to sterically hinder Sec13-Nup145C binding. Since Ile^{12} is part of a hydrophobic pocket binding Nup145C amino acids Phe^{144}, Tyr^{148} and Phe^{150}, this mutation was believed to lead to the disruption of Sec13-Nup145C interaction. Since the analogous sequence in Sec31 (amino acids 378-384) lacks amino acids with aromatic side chains, the corresponding region in the Sec13-Sec31 complex should contain enough space to accommodate a phenylalanine side chain.

This hypothesis was tested by copurification experiments. Wild type Sec13 and Sec13 ($AA12^{I\rightarrow F}$) were each coexpressed with either Nup145C, $Sec31^{372-411}$ (containing only the Sec31 insertion blade) or $Sec31^{1-748}$ (this construct contains all the domains involved in binding Sec13). The proteins were copurified on Ni-NTA agarose and further run through gel filtration columns. All fractions from the Ni-NTA copurification and gel filtration chromatography were analyzed on SDS PAGE gels.

Results demonstrated that the Sec13 ($AA12^{I\rightarrow F}$) had similar expression levels as wild type Sec13. Further, both proteins were equally soluble and displayed the same properties during gel filtration chromatography, demonstrating that the $AA12^{I\rightarrow F}$ mutation did not radically effect the structure of Sec13.

Expression of $Sec31^{1-748}$ was too low for determination of binding from the results of Ni-NTA purification. Gel filtration chromatography revealed that both wild type Sec13 and Sec13 ($AA12^{I\rightarrow F}$) copurified with $Sec31^{1-748}$ in similar amounts. Both types of Sec13 appeared in the same gel filtration chromatography fractions as $Sec31^{1-748}$, confirming that the Sec13 ($AA12^{I\rightarrow F}$) mutation is not disruptive to Sec13-Sec31 binding.

Copurification of wild type Sec13 and Sec13 ($AA12^{I\rightarrow F}$) with $Sec31^{372-411}$, again, could not be assessed directly from the results of purification on Ni-NTA agarose. Due to the small size of $Sec31^{372-411}$ (~4.3 kDa) samples had to be analyzed on high density (16.5%) SDS PAGE gels. Since analysis on 16.5% SDS PAGE gels requires high sample purity, relatively impure samples purified on Ni-NTA agarose did not return clear results. Gel filtration chromatography revealed

that wild type Sec13 copurified with Sec31$^{372-411}$ on Ni-NTA agarose. Both proteins appeared in the same elution fractions, indicating that they traveled through the gel filtration column as a complex. The same results could not be repeated for Sec31$^{372-411}$ coexpressed with Sec13 (AA12$^{I\rightarrow F}$).

The fact that the Sec13 (AA12$^{I\rightarrow F}$) mutant binds Sec31$^{1-748}$, but does not detectably bind Sec31$^{372-411}$, indicates that the Sec31 insertion fragment alone cannot account for the Sec13-Sec31 interaction. Since no interactions between Sec13 and the amino and carboxy terminal domains of Sec31 were observed in the crystal structures, it is improbable (although not impossible) that these two domains of Sec31 enhance Sec13-Sec31$^{1-748}$ binding. Changed orientations of amino acid side chains in the two Sec13-Sec31 structures might provide an explanation for this effect. In the structure of Sec13 bound to the Sec31 amino terminal β-propeller domain, a salt bridge is formed between the Sec13 amino acid Arg^{259} and the Sec31 amino acid Glu^{379}. In the structure of Sec13 bound to the carboxy terminal, α-helical domain of Sec31, Glu^{379} is part of the amino terminal region of the molecule and its side chain is oriented towards the bulk solvent, severing a salt bridge. Since Glu^{379} would also be part of the amino terminus in Sec31$^{372-411}$, but not in Sec31$^{1-748}$, it is possible that a similar situation occurs in these two constructs. Additional experiments with mutants of Sec31 Glu^{379} are necessary to investigate this hypothesis.

When Nup145C was coexpressed with wild type Sec13, expression of Sec13 was much higher than that of Nup145C, although both proteins had high expression levels. Nup145C copurified on Ni-NTA agarose with wild type Sec13. Both Nup145C and Sec13 appeared in the same fractions of gel filtration eluates, indicating that they traveled through the column as a single complex. A large peak of pure Sec13 was also observed. This could be explained by the higher expression levels of Sec13. To verify this hypothesis, His_6-tagged Nup145C was coexpressed with tag-free Sec13 and the lysate was purified on Ni-NTA agarose. The gel filtration chromatogram of the Ni-NTA elution contained a large peak that was shown by SDS-PAGE to contain Sec13 and Nup145C. A small peak corresponding to non-bound Sec13 was observed, signifying that some Sec13 had dissociated from Nup145C or that Sec13 could non-specifically bind Ni-NTA agarose.

Nup145C coexpressed with Sec13 (AA12$^{I\rightarrow F}$) was found mostly in the insoluble fraction of the cell lysate. Since Nup145C alone is not soluble *in vitro* this was an indication that the Sec13 (AA12$^{I\rightarrow F}$) mutant disrupted Nup145C binding. However, a small fraction of Nup145C did copurify with Sec13 (AA12$^{I\rightarrow F}$) on Ni-NTA agarose. Further, a small peak containing Nup145C

bound to Sec13 (AA12$^{I\rightarrow F}$) could be detected in the gel filtration chromatograms of the the Ni-NTA eluates. This peak was significantly smaller than the corresponding peak in chromatogram of the wild type Sec13-Nup145C complex. A large peak of pure Sec13 was also observed.

A gel filtration chromatogram of His_6-tagged Nup145C coexpressed with tag-free Sec13 (AA12$^{I\rightarrow F}$) contained peaks of very low intensity, which were shown by SDS PAGE to contain several proteins of different molecular masses, most probably impurities.

Together, the copurification and gel filtration data indicate that, while the Sec13 (AA12$^{I\rightarrow F}$) mutation does not completely impair Sec13-Nup145C binding it significantly reduces the affinity of Sec13 for Nup145C. The physiological significance of the Sec13 (AA12$^{I\rightarrow F}$) mutation remains to be determined. Green-fluorescent protein (GFP)-tagged protein fluorescence localization an yeast growth assays are possible methods for assessing the affinity of the Sec13 (AA12$^{I\rightarrow F}$) mutant for Nup145C *in vivo*.

Conclusions

- A crystal structure of yeast Sec13 bound to the yeast Nup145C insertion blade was successfully determined.
- Comparison with known structures of Sec13 confirmed general structural conservation of the seven bladed β-propeller fold.
- Comparison of the yeast Sec13- yeast Sec31 and yeast Sec13-yeast Nup145C interaction interfaces revealed several conserved polar contacts, but also found a general lack of conservation in the arrangement of hydrophobic contacts.
- Based on the known crystal structures of Sec13-containing complexes, a Sec13 ($AA12^{I\rightarrow F}$) mutant, that was proposed to selectively disrupt the Sec13-Nup145C interaction without affecting Sec13 binding to Sec31, was designed.
- Effect of the Sec13 ($AA12^{I\rightarrow F}$) mutant on Sec13 affinities for Nup145C and Sec31 was examined by copurification on Ni-NTA agarose followed by gel filtration chromatography. The mutation was confirmed to significantly affect Sec13 binding to Nup145C without disrupting the Sec13-Sec31 interaction.

Bibliography

BOOKS

1. Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P. (2002): Molecular biology of the cell. Garland Science, New York.
2. Berg J.M., Tymoczko J.L., Stryer L. (2002): Biochemistry. W.H. Freeman and Company, New York, page 112.
3. Campbell N.A., Reece J.B. (2002): Biology. Benjamin Cummings, San Francisco.
4. Drenth J. (2007): Principles of Protein X-Ray Crystallography. Springer Science+Business Media, New York
5. Roe S. (2001): Protein Purification Techniques. Oxford University Press, Oxford.
6. Voet D., Voet J.G. (2004): Biochemistry. John Wiley and Sons, Hoboken.

PAPERS

1. Akey C.W., Radermacher M. (1993): Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. J. Cell. Biol.**122(1)**:1-19.
2. Alber F., Dokudovskaya S., Veenhoff L.M., Zhang W., Kipper J., Devos D., Suprapto A., Karni-Schmidt O., Williams R., Chait B.T., Sali A., Rout M.P. (2007): The molecular architecture of the nuclear pore complex. Nature. **450(7170)**:695-701.
3. Bayliss R., Littlewood T., Stewart M. (2000): Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell. **102(1)**:99-108.
4. Beck M., Förster F., Ecke M., Plitzko J.M., Melchior F., Gerisch G., Baumeister W., Medalia O. (2004): Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. Science. **306(5700)**:1387-90.
5. Beck M., Lucić V., Förster F., Baumeister W., Medalia O. (2007): Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. Nature. **449(7162)**:611-5.
6. Belgareh N., Rabut G., Baï S.W., van Overbeek M., Beaudouin J., Daigle N., Zatsepina O.V., Pasteau F., Labas V., Fromont-Racine M., Ellenberg J., Doye V. (2001): An

evolutionarily conserved NPC subcomplex, which redistributes in part to kinetochores in mammalian cells. J. Cell. Biol. **154(6)**:1147-60.

7. Berke I.C., Boehmer T., Blobel G., Schwartz T.U. (2004): Structural and functional analysis of Nup133 domains reveals modular building blocks of the nuclear pore complex. J. Cell. Biol. **167(4)**:591-7.
8. Bi X., Mancias J.D., Goldberg J. (2007): Insights into COPII coat nucleation from the structure of Sec23.Sar1 complexed with the active fragment of Sec31. Dev. Cell. **13(5)**:635-45.
9. Boehmer T., Jeudy S., Berke I.C., Schwartz T.U. (2008): Structural and functional studies of Nup107/Nup133 interaction and its implications for the architecture of the nuclear pore complex. Mol. Cell. **30(6)**:721-31.
10. Brohawn S.G., Leksa N.C., Spear E.D., Rajashankar K.R., Schwartz T.U. (2008): Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. Science. **322(5906)**:1369-73.
11. Cronshaw J.M., Krutchinsky A.N., Zhang W., Chait B.T., Matunis M.J. (2002): Proteomic analysis of the mammalian nuclear pore complex. J. Cell. Biol. **158(5)**:915-27
12. D'Angelo M.A., Hetzer M.W. (2008): Structure, dynamics and function of nuclear pore complexes. Trends Cell Biol. **18(10)**:456-66.
13. Devos D., Dokudovskaya S., Alber F., Williams R., Chait B.T., Sali A., Rout M.P. (2004): Components of coated vesicles and nuclear pore complexes share a common molecular architecture. PLoS Biol. **2(12)**:e380.
14. Devos D., Dokudovskaya S., Williams R., Alber F., Eswar N., Chait B.T., Rout M.P., Sali A (2006): Simple fold composition and modular architecture of the nuclear pore complex. Proc. Natl. Acad. Sci. USA. **103(7)**:2172-7.
15. Elliott D.J., Grellscheid S.N. (2006): Alternative RNA splicing regulation in the testis. Reproduction. **132(6)**:811-9
16. Fath S., Mancias J.D., Bi X., Goldberg J. (2007): Structure and organization of coat proteins in the COPII cage. Cell. **129(7)**:1325-36.
17. Fribourg S., Braun I.C., Izaurralde E., Conti E. (2001): Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. Mol. Cell. **8(3)**:645-56.
18. Gruber J., Zawaira A., Saunders R., Barrett C.P., Noble M.E. (2007): Computational

analyses of the surface properties of protein-protein interfaces. Acta. Crystallogr. D. Biol. Crystallogr. **63(Pt 1)**:50-7.

19. Hsia K.C., Stavropoulos P., Blobel G., Hoelz A. (2007): Architecture of a coat for the nuclear pore membrane. Cell. **131(7)**:1313-26.
20. Jeudy S., Schwartz T.U. (2007): Crystal structure of nucleoporin Nic96 reveals a novel, intricate helical domain architecture. J. Biol. Chem. **282(48)**:34904-12.
21. Kiseleva E., Allen T.D., Rutherford S., Bucci M., Wente S.R., Goldberg M.W. (2004): Yeast nuclear pore complexes have a cytoplasmic ring and internal filaments. J. Struct. Biol. **145(3)**:272-88.
22. Loïodice I., Alves A., Rabut G., Van Overbeek M., Ellenberg J., Sibarita J.B., Doye V. (2004): The entire Nup107-160 complex, including three new members, is targeted as one entity to kinetochores in mitosis. Mol. Biol. Cell. **15(7)**:3333-44.
23. Lutzmann M., Kunze R., Buerer A., Aebi U., Hurt E. (2002): Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins. EMBO J. **21(3)**:387-97.
24. Matzat L.H., Berberoglu S., Lévesque L. (2008): Formation of a Tap/NXF1 homotypic complex is mediated through the amino-terminal domain of Tap and enhances interaction with nucleoporins. Mol. Biol. Cell. **19(1)**:327-38.
25. Menon B.B., Sarma N.J., Pasula S., Deminoff S.J., Willis K.A., Barbara K.E., Andrews B., Santangelo G.M. (2005): Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. Proc. Natl. Acad. Sci. USA. **102(16)**:5749-54.
26. Moreno F., Ahuatzi D., Riera A., Palomino C.A., Herrero P. (2005): Glucose sensing through the Hxk2-dependent signalling pathway. Biochem. Soc. Trans. **33(1)**:265-8.
27. Nagai S, Dubrana K, Tsai-Pflugfelder M, Davidson MB, Roberts TM, Brown GW, Varela E, Hediger F, Gasser SM, Krogan NJ. (2008): Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. Science. **322(5901)**:597-602.
28. Orjalo A.V., Arnaoutov A., Shen Z., Boyarchuk Y., Zeitlin S.G., Fontoura B., Briggs S., Dasso M., Forbes D.J. (2006): The Nup107-160 nucleoporin complex is required for correct bipolar spindle assembly. Mol. Biol. Cell. **17(9)**:3806-18.
29. Osmani A.H., Davies J., Liu H.L., Nile A., Osmani S.A. (2006): Systematic deletion and mitotic localization of the nuclear pore complex proteins of Aspergillus nidulans. Mol. Biol. Cell. **17(12)**:4946-61.

30. Pemberton L.F., Paschal B.M. (2005): Mechanisms of receptor-mediated nuclear import and nuclear export. Traffic. **6(3)**:187-98.
31. Peters R. (2005): Translocation through the nuclear pore complex: selectivity and speed by reduction-of-dimensionality. Traffic. **6(5)**:421-7.
32. Pryer N.K., Salama N.R., Schekman R., Kaiser C.A. (1993): Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum in vitro. J. Cell. Biol. **120(4)**:865-75.
33. Ratner G.A., Hodel A.E., Powers M.A. (2007): Molecular determinants of binding between Gly-Leu-Phe-Gly nucleoporins and the nuclear pore complex. J. Biol. Chem. **282(47)**:33968-76.
34. Rout M.P., Aitchison J.D., Suprapto A., Hjertaas K., Zhao Y., Chait B.T. (2000): The yeast nuclear pore complex: composition, architecture, and transport mechanism. J. Cell. Biol. **148(4)**:635-51.
35. Schwartz T.U. (2005): Modularity within the architecture of the nuclear pore complex. Curr. Opin. Struct. Biol. **15(2)**:221-6.
36. Siniossoglou S., Wimmer C., Rieger M., Doye V., Tekotte H., Weise C., Emig S., Segref A., Hurt E.C. (1996): A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. Cell. **84(2)**:265-75.
37. Suntharalingam M., Wente S.R. (2003): Peering through the pore: nuclear pore complex structure, assembly, and function. Dev. Cell. **4(6)**:775-89.
38. Tran E.J., Wente S.R. (2006): Dynamic nuclear pore complexes: life on the edge. Cell. **125(6)**:1041-53.
39. Vasu S., Shah S., Orjalo A., Park M., Fischer W.H., Forbes D.J. (2001): Novel vertebrate nucleoporins Nup133 and Nup160 play a role in mRNA export. J. Cell. Biol. **155(3)**:339-54.
40. Yang Q., Rout M.P., Akey C.W. (1998): Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. Mol. Cell . **1(2)**:223-34.