Biophysical and Structural Characterization of Components From The Nuclear Pore Complex and the Ubiquitin Pathway

by

James R. Partridge
B.S. University of California at Davis
Davis, CA 2004

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

DOCTOR OF PHILOSOPHY

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAY 2010

© 2010 Massachusetts Institute of Technology.
All rights reserved.

Signature of Author........................................ James R. Partridge
Department of Biology
May 4, 2010

Certified by........................................................ Thomas U. Schwartz
Associate Professor of Biology
Thesis Supervisor

Accepted by.................................................. Stephen P. Bell
Professor of Biology
Chair, Biology Graduate Committee
Biophysical and Structural Characterization of Components From The Nuclear Pore Complex and the Ubiquitin Pathway

by

James R. Partridge

Abstract

Formation of an endomembrane system in the eukaryotic cell is a hallmark of biological evolution. One such system is the nuclear envelope (NE), composed of an inner and outer membrane, used to form a nucleus and enclose the cell’s genome. Access to the nucleus from the cytoplasm is mediated by a massive macromolecular machine called the nuclear pore complex (NPC). The NPC resides as a circular opening embedded in the NE and is composed of only ~30 proteins that assemble with octagonal symmetry as biochemically defined subcomplexes to form the NPC. One such subcomplex is the Nsp1 / Nup62 complex, composed of three proteins and stabilized by coiled-coil interactions. Here we reconstitute a tetrameric assembly between the Nsp1-complex and a fourth nucleoporin (Nup) Nic96. Nic96 harbors a 20 kDa coiled-coil domain at the N-terminus followed by a 65 kDa stacked helical domain. The coiled-coil domain of the Nsp1-complex and the N-terminus of Nic96 combine to form a tetrameric assembly, integrated into the NPC lattice scaffold via the stacked helical domain of Nic96. We characterized the coiled-coil assembly with size exclusion chromatography and analytical ultracentrifugation. Deletion experiments and point mutations, directed by hydrophobic cluster analysis, were used to map connecting helices between members of the protein assembly.

Although the core of the NPC is a rigid scaffold built for structural integrity, the NPC as a whole is a dynamic macromolecular machine. Protein transport is regulated by the small G protein Ran. Ran interacts with the NPC of metazoa via two asymmetrically localized components, Nup153 at the nuclear face and Nup358 at the cytoplasmic face. Both Nups contain distinct RANBP2 type zinc finger (ZnF) domains. We present crystallographic data detailing the interaction between Nup153-ZnFs and RanGDP. A crystal-engineering approach led to well-diffracting crystals so that all ZnF-Ran complex structures are refined to high resolution. Each of the four zinc finger modules of Nup153 binds one Ran molecule in largely independent fashion. Nup153-ZnFs bind RanGDP with higher affinity than RanGTP, however the modest difference suggests that this may not be physiologically meaningful. ZnFs may be used to concentrate Ran at the NPC to facilitate nucleocytoplasmic transport.

In a separate study we present a structural analysis of the HECT domain from the E3 ubiquitin ligase HUWE1 and with biophysical data we show that an N-terminal helix stabilizes the HECT domain. This element modulates activity, as measured by self-ubiquitination induced in the absence of this helix, distinct from its effects on Ub conjugation of substrate Mcl-1. Such subtle structural elements in this domain potentially regulate the variable substrate specificity displayed by all HECT domain type, E3 ubiquitin ligases.

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

Abstract ........................................................................................................................................... 2

Figure and Table Legend ................................................................................................................ 5

## Chapter 1: Introduction .................................................................................................................. 8

### Introduction to the Nuclear Pore Complex ................................................................................. 9
- Overall structure ................................................................................................................................. 9
- Modularity ........................................................................................................................................... 11
- Protein composition ............................................................................................................................. 12
- Dynamics ........................................................................................................................................... 12
- Domain architecture ........................................................................................................................... 13
- FG-repeats ......................................................................................................................................... 13
- Coiled-coils ........................................................................................................................................ 14
- β-Propellers ...................................................................................................................................... 15
- α-Helical domains .............................................................................................................................. 16
- ACE1 domains .................................................................................................................................... 17

### Structural characterization of NPC subcomplexes ................................................................. 18
- NPC subcomplexes ............................................................................................................................ 18
- In vitro nucleoporin subcomplex reassembly ................................................................................... 20
- Purification of NPC subcomplexes for crystallographic and biophysical studies ......................... 20
- The Nsp1 / Nup62 subcomplex ......................................................................................................... 21

### The interaction between Nup153 and the small GTPase Ran .................................................. 23
- Nup153 is a versatile nucleoporin ..................................................................................................... 23
- RanBP2-type zinc fingers in the eukaryotic cell .................................................................................. 24
- Structural and biophysical analysis of Nup153 zinc fingers and RanGDP ...................................... 26

## Figures - Chapter 1 ....................................................................................................................... 28

## Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain .................................................................................................................................. 34

### Introduction .................................................................................................................................. 35

### Results and Discussion ............................................................................................................... 37
- Crystallographic analysis of Nup153-ZnF•RanGDP complexes ......................................................... 37
- Overall structure of the ZnF•RanGDP complex ............................................................................... 38
- Engineering improved crystal contacts ............................................................................................. 40
- Comparison of the four Nup153 zinc fingers bound to RanGDP ....................................................... 41
- Isothermal titration calorimetry ......................................................................................................... 43
- Conservation of RanBP2-type ZnF cassettes in nucleoporins ........................................................... 45
- Comparison of binding interactions among RanBP2-type zinc fingers .......................................... 46

### Discussion .................................................................................................................................... 47

### Materials and Methods .............................................................................................................. 47
- Protein purification .............................................................................................................................. 48
- Crystallization .................................................................................................................................... 49
- Structure determination ....................................................................................................................... 50
- Isothermal titration calorimetry ......................................................................................................... 50
- Sequence analysis .............................................................................................................................. 51
- Protein Data Bank accession numbers ............................................................................................. 51

Acknowledgements .......................................................................................................................... 51

## Figures - Chapter 2 ....................................................................................................................... 52
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex ................................................................. 69

Introduction .................................................................................................................. 70

Results .......................................................................................................................... 75

- Reassembly of the trimeric yNsp1/rNup62 subcomplex ........................................... 75
- Mapping interactions within the Nsp1 complex ....................................................... 77
- Biophysical characterization of the Nsp1 complex .................................................. 79
- Assembly of a tetrameric assembly between the Nsp1 complex and Nic96 ............... 81

Discussion ...................................................................................................................... 82

Methods ....................................................................................................................... 83

- Protein purification .................................................................................................... 84
- Hydrophobic cluster analysis .................................................................................... 85
- Mapping helical interactions ..................................................................................... 85
- Analytical ultracentrifugation .................................................................................... 86

Figures - Chapter 3 ........................................................................................................ 87

Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities ................................................................. 95

Introduction .................................................................................................................. 96

- Mechanism of ubiquitination mediated by E3 HECT domains ................................ 97
- Structural characteristics of the E3 HECT domain .................................................. 98

Results .......................................................................................................................... 100

- Structure of the HUWE1 HECT domain ................................................................. 100
- Catalytic activity of the HECT domain ...................................................................... 103
- Thioester formation in the HECT domain ................................................................. 105
- Substrate ubiquitination catalyzed by the HECT domain ......................................... 106
- Catalytic activity of the C4341A mutants ................................................................. 107

Discussion ...................................................................................................................... 107

Methods ....................................................................................................................... 110

- Plasmids ..................................................................................................................... 110
- Bacterial protein expression and purification ............................................................ 111
- Circular dichroism ...................................................................................................... 111
- Biochemical assays .................................................................................................... 111
- Mcl-1 ubiquitination assay ......................................................................................... 112
- Thioester assay .......................................................................................................... 112
- Single-turnover assay ............................................................................................... 112
- Crystallization of HUWE1 HECT domain ................................................................. 113
- Data collection and processing .................................................................................. 114
- Acknowledgements .................................................................................................... 114

Figures - Chapter 4 ........................................................................................................ 116

Chapter 5: Conclusion ................................................................................................... 127

Summary ....................................................................................................................... 128

Future Directions ...................................................................................................... 129

- A role for Nup153 zinc fingers in nucleocytoplasmic transport ................................. 129
- The tetrameric assembly between Nic96 and the Nsp1 subcomplex ......................... 131
- The HUWE1 HECT domain ...................................................................................... 132
- A balance between conformational flexibility and enzyme promiscuity .................. 132

References ..................................................................................................................... 135
**Figure and Table Legend**

<table>
<thead>
<tr>
<th>Figure/Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Overall structure of the Nuclear Pore Complex</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Cryo-electron tomography of the NPC at ~60 Å resolution</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Schematic representation of the modular NPC assembly</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Inventory of the NPC</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Structures of nucleoporins</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Ran regulates nucleocytoplasmic transport</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Domain architecture of Nup153</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Alignment of the four Ran-binding zinc fingers of Nup153 from R. norvegicus</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Ribbon diagram representation of the Nup153-ZnF2•RanGDP complex</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Crystallographic lattice formed in the crystallization of a Nup153-ZnF•RanGDP complex</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>Comparison of the four individual zinc fingers with RanGDP</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>Conserved water network and non-conserved hydrogen bond</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>Surface representation of RanGDP / RanGTP with or without Nup153-ZnF</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>Isothermal titration calorimetry (ITC) for each individual ZnF binding RanGDP</td>
</tr>
<tr>
<td>Figure 2.9</td>
<td>ITC illustrating low micro-molar affinity between RanGDP and two tandem ZnF pairs from the Nup153- and Nup358-ZnF domain</td>
</tr>
<tr>
<td>Figure 2.10</td>
<td>Weblogs from members of the RanBP2 class of zinc fingers known to have unique binding partners</td>
</tr>
<tr>
<td>Figure 2.11</td>
<td>Sequence alignment of the human Nup358/RanBP2 and Nup153 ZnFs, including the respective C-terminal linker regions</td>
</tr>
<tr>
<td>Figure 2.12</td>
<td>Nup153ZnF12•Ran crystal contains Ran and intact Nup153ZnF12</td>
</tr>
<tr>
<td>Table 2.1</td>
<td>Crystallographic Data and Refinement Statistics</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>Thermodynamic parameters for ZnF-Ran binding in ITC experiments</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>Experimental constructs, abbreviations, and descriptions</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Domain architecture of both the Nsp1 and Nup62 trimeric complex</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Hydrophobic cluster analysis of the Nup62 trimeric assembly</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Systematic deletion of helices from Nup57 map an interaction between Nup57 and Nsp1</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Analytical ultracentrifugation of Nsp1 / Nup62 assemblies</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Purification of a tetrameric assembly between full-length Nic96 and the trimeric Nsp1 complex</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>AUC data analysis</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>The α1 helix stabilizes the HUWE1 HECT domain</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Structure of the HUWE1 HECT domain</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>E3 Ubiquitin ligase activity of HUWE1 HECT domain</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Detection of Ub~thioesters in the HUWE1 HECT domains</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Substrate ubiquitination activity of the HUWE1 HECT domains</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Ubiquitination activity of the HUWE1 HECT domains at 37°C</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Both HUWE1 Δα1 and +α1 HECT domains maintain the β conformation</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Mutation of the catalytic cysteine (C4341) to alanine abolishes activity of the HECT domain</td>
</tr>
</tbody>
</table>

| Table 4.1    | Crystallographic data and refinement statistics |

---

**Tables:**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Crystallographic Data and Refinement Statistics</td>
</tr>
<tr>
<td>2.2</td>
<td>Thermodynamic parameters for ZnF-Ran binding in ITC experiments</td>
</tr>
<tr>
<td>2.3</td>
<td>Experimental constructs, abbreviations, and descriptions</td>
</tr>
<tr>
<td>3.1</td>
<td>AUC data analysis</td>
</tr>
<tr>
<td>4.1</td>
<td>The α1 helix stabilizes the HUWE1 HECT domain</td>
</tr>
<tr>
<td>4.2</td>
<td>Structure of the HUWE1 HECT domain</td>
</tr>
<tr>
<td>4.3</td>
<td>E3 Ubiquitin ligase activity of HUWE1 HECT domain</td>
</tr>
<tr>
<td>4.4</td>
<td>Detection of Ub~thioesters in the HUWE1 HECT domains</td>
</tr>
<tr>
<td>4.5</td>
<td>Substrate ubiquitination activity of the HUWE1 HECT domains</td>
</tr>
<tr>
<td>4.6</td>
<td>Ubiquitination activity of the HUWE1 HECT domains at 37°C</td>
</tr>
<tr>
<td>4.7</td>
<td>Both HUWE1 Δα1 and +α1 HECT domains maintain the β conformation</td>
</tr>
<tr>
<td>4.8</td>
<td>Mutation of the catalytic cysteine (C4341) to alanine abolishes activity of the HECT domain</td>
</tr>
</tbody>
</table>

---

**Figures:**

- **Figure 1.1** Overall structure of the Nuclear Pore Complex
- **Figure 1.2** Cryo-electron tomography of the NPC at ~60 Å resolution
- **Figure 1.3** Schematic representation of the modular NPC assembly
- **Figure 1.4** Inventory of the NPC
- **Figure 1.5** Structures of nucleoporins
- **Figure 1.6** Ran regulates nucleocytoplasmic transport
- **Figure 2.1** Domain architecture of Nup153
- **Figure 2.2** Alignment of the four Ran-binding zinc fingers of Nup153 from *R. norvegicus*
- **Figure 2.3** Ribbon diagram representation of the Nup153-ZnF2•RanGDP complex
- **Figure 2.4** Crystallographic lattice formed in the crystallization of a Nup153-ZnF•RanGDP complex
- **Figure 2.5** Comparison of the four individual zinc fingers with RanGDP
- **Figure 2.6** Conserved water network and non-conserved hydrogen bond
- **Figure 2.7** Surface representation of RanGDP / RanGTP with or without Nup153-ZnF
- **Figure 2.8** Isothermal titration calorimetry (ITC) for each individual ZnF binding RanGDP
- **Figure 2.9** ITC illustrating low micro-molar affinity between RanGDP and two tandem ZnF pairs from the Nup153- and Nup358-ZnF domain
- **Figure 2.10** Weblogs from members of the RanBP2 class of zinc fingers known to have unique binding partners
- **Figure 2.11** Sequence alignment of the human Nup358/RanBP2 and Nup153 ZnFs, including the respective C-terminal linker regions
- **Figure 2.12** Nup153ZnF12•Ran crystal contains Ran and intact Nup153ZnF12
- **Table 2.1** Crystallographic Data and Refinement Statistics
- **Table 2.2** Thermodynamic parameters for ZnF-Ran binding in ITC experiments
- **Table 2.3** Experimental constructs, abbreviations, and descriptions
- **Figure 3.1** Domain architecture of both the Nsp1 and Nup62 trimeric complex
- **Figure 3.2** Hydrophobic cluster analysis of the Nup62 trimeric assembly
- **Figure 3.3** Systematic deletion of helices from Nup57 map an interaction between Nup57 and Nsp1
- **Figure 3.4** Analytical ultracentrifugation of Nsp1 / Nup62 assemblies
- **Figure 3.5** Purification of a tetrameric assembly between full-length Nic96 and the trimeric Nsp1 complex
- **Table 3.1** AUC data analysis
- **Figure 4.1** The α1 helix stabilizes the HUWE1 HECT domain
- **Figure 4.2** Structure of the HUWE1 HECT domain
- **Figure 4.3** E3 Ubiquitin ligase activity of HUWE1 HECT domain
- **Figure 4.4** Detection of Ub~thioesters in the HUWE1 HECT domains
- **Figure 4.5** Substrate ubiquitination activity of the HUWE1 HECT domains
- **Figure 4.6** Ubiquitination activity of the HUWE1 HECT domains at 37°C
- **Figure 4.7** Both HUWE1 Δα1 and +α1 HECT domains maintain the β conformation
- **Figure 4.8** Mutation of the catalytic cysteine (C4341) to alanine abolishes activity of the HECT domain
- **Table 4.1** Crystallographic data and refinement statistics
Acknowledgements

The work presented in this dissertation would not be possible without the efforts and dedication of several people and institutions. First and foremost, I want to thank my advisor, Thomas Schwartz. Thomas has provided a seemingly limitless amount of guidance and encouragement during the past five years. With clarity and passion, Thomas has trained me as a crystallographer, and I am thankful for his direction. Prior to entering the graduate program at MIT there were several other scientists that guided my research and instructed me over the years. As an undergraduate at UC Davis, lacking any research experience, Clark Lagarias was brave enough to take me into his lab and give me an independent research project. The experience in his lab intensified my interest in biological research, and for this I am grateful. Also at UC Davis, my work with Jodi Nunnari finally convinced me to pursue a PhD, and I am thankful for her sustained guidance over the years. My short three months of work with Frank Slack at Yale University convinced me to leave the great state of California and come to the East Coast. Frank’s patience in training me how to work with C. elegans and think about developmental biology was remarkable. He is a fantastic teacher and research scientist, and I am grateful for his instruction.

Four months after Thomas started his lab at MIT he met three naive graduate students, Steve Brohawn, James Whittle, and myself. Together the three of us have formed the core of Thomas’ lab and worked day after day as both friends and colleagues. In the presence of some very gifted undergraduates, fellow grad students, and postdocs we have remained close friends. I want to thank Steve and James for being great friends, as well as a source of sanity, insanity, and fresh ideas. Thank you to Brian, Nina, and Silvija for being great work colleagues. Big thanks to Sandra and T.B. for continuing to be great friends and helping during my formative years in the Schwartz lab.

The members of my thesis committee, Amy Keating and Mike Yaffe, have been a constant source of guidance and encouragement throughout the years. I appreciate their work, not only as thesis advisors, but also as teachers inside the classroom. I had the pleasure of having both Amy and Mike as instructors and I want to thank them for their work and dedication.

None of this work would be possible without my family. I want to say thank you to my brothers and sister, Tom, Vince, and Roz. Living so far away it is sometimes hard to express how often I think about you guys. Your visits throughout the years have made me so much happier than you can imagine. To my dad, I love you more than you will ever know. To my mom, I want to say thank you for all of your love and support throughout the years. You are always a source of inspiration for me and I would not have been able to finish this journey without you. To Greg, I appreciate all that you have done in our lives and especially for showing me how to live life to the fullest.

To my best friend and wife, Teresa. There is really no way I can adequately thank you. You have remained my biggest supporter during our time here at MIT. You have always believed in me, but most importantly you believed in me when I did not. Your unconditional love has been a constant source of happiness for me and I am forever grateful.
for César and Jack
A portion of the material presented in this chapter was adapted, with permission, from the following publication:

Introduction to the Nuclear Pore Complex

The hallmark of eukaryotic cells is an elaborate endomembrane system that creates membrane-enclosed organelles. The nucleus is a prominent organelle, as it harbors the genetic material of the cell. The Nuclear Pore Complex (NPC) is the sole gateway into the nucleus and perforates the nuclear envelope where the inner nuclear membrane (INM) and outer nuclear membrane (ONM) of the nuclear envelope (NE) are fused. NPCs are among the largest multiprotein assemblies in the quiescent cell and were first described 50 years ago by electron microscopy (Watson, 1959). For general reviews the reader is also referred to (D'Angelo and Hetzer, 2008; Lim et al., 2008b; Tran and Wente, 2006) and for the mechanism of NPC assembly to (Antonin et al., 2008) and for nucleocytoplasmic transport of proteins and RNA-based molecules to (Carmody and Wente, 2009; Cook et al., 2007; Pemberton and Paschal, 2005; Stewart, 2007; Weis, 2003). The emerging role of the NPC in gene regulation and nuclear organization is addressed in (Akhtar and Gasser, 2007; Heessen and Fornerod, 2007).

Overall structure

The first electron micrographs of the NPC showed that it forms an octagonal ring whose central channel is less electron dense than the eight lobes that surround it. Considering overall shape, scanning electron microscopy experiments (SEM) have recorded some of the most stunning NPC images (Fig. 1.1). While the architectural core is grossly symmetric about the plane of the membrane, the peripheral components on the nuclear and cytoplasmic faces are distinct. These peripheral components recapitulate the eightfold symmetry about the transport axis exhibited by the architectural core. On the cytoplasmic side, eight knobs, thought to be attachment sides for fibrous extensions, are visible in NPCs from multicellular species (Kiseleva et al., 2000). In yeast, these features are less pronounced, but are likely still present (Kiseleva et al., 2004). On the nucleoplasmic side, a ring termed the nuclear basket is suspended from eight filaments that join the NPC. Concerning size, the diameter of the NPC appears to be
similar in all eukaryotes, about 90-120nm (Akey and Radermacher, 1993; Beck et al., 2007; Fahrenkrog, 2000; Hinshaw et al., 1992; Stoffler et al., 2003). However, there is still considerable uncertainty about the height, determined to be ~30-50nm (Alber et al., 2007b; Elad et al., 2009).

Further work has led to progressively more detailed reconstructions of the NPC. Cryo-electron microscopy that relies on averaging images from many NPCs has been employed to study the core NPC structure, the part that spans the distance between the faces of INM and ONM (Akey and Radermacher, 1993; Hinshaw et al., 1992). These studies have shown that the scaffold ring structure, the electron dense material near the nuclear membrane, has alternating thicker and thinner regions, hence it is often called the spoke ring (Akey and Radermacher, 1993; Hinshaw et al., 1992). The scaffold structure appears to penetrate the pore membrane to form a perinuclear ring structure. Using cryo-electron tomography, the best pictures of complete NPCs have been achieved, extending even to a resolution of ~ 6 nm (Beck et al., 2007; Elad et al., 2009). With this technique, details of the ring structures become apparent. The scaffold can be divided into three main ring elements: a central spoke ring is sandwiched between a cytoplasmic ring and a nucleoplasmic ring. The rings appear to float on top of one another, indicating that material connecting them is less electron-dense than the rings themselves. Alternatively, this may be due to technical difficulties, such as the 'missing cone' problem or poor resolution in the Z-direction.

The central transport cavity of the nuclear pore complex shows no distinct structural features, consistent with the perception that it is filled by an aqueous meshwork formed by natively unfolded FG-domains, which are long polypeptide sequences found in several nucleoporins that contain phenylalanine-glycine (FG) repeats but are otherwise hydrophilic. These extensions are thought to form a distinct, semi-permeable environment that prevents the
diffusion of large molecules, unless they are bound to nuclear transport factors that facilitate entry into this central cavity.

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

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

**Modularity**

A characteristic of the NPC is its high degree of modularity, which manifests itself at several levels. First, the NPC is organized around a central eightfold rotational symmetry. Second, only ~30 nucleoporins, composed of a limited set of domain topologies, come together to construct the NPC. Third, nucleoporins have various dwell times at the NPC, with only a fraction being stably attached at all times. Finally, the stably attached nucleoporins are arranged into subcomplexes, each of which assembles in multiple copies to build the entire NPC (Fig. 1.2 and Fig. 1.3). This modularity is the basis for approaching structural determination of the assembly at atomic resolution (Schwartz, 2005).
Protein composition

Two studies, using *S. cerevisiae* (Rout et al., 2000) and rat hepatocytes (Cronshaw et al., 2002) as starting material, determined an inventory of nucleoporins. In both studies, cell extracts were enriched for NPCs by fractionation and analyzed by mass-spectrometry to identify the proteins purified thereby. The set of proteins found in both organisms is largely identical.

The nucleoporins can be broadly classified into three categories (Fig. 1.4). ~10 contain disordered N- and/or C-terminal regions that are rich in phenylalanine-glycine (FG) repeats. These FG-repeat regions emanate into and form the transport barrier in the channel of the NPC.

~15 nucleoporins have distinct architectural functions and form the NPC scaffold structure.

Three nucleoporins have transmembrane domains and anchor the NPC in the circular openings in the NE. Immunogold-labeling of all nucleoporins shows that the majority of the Nups, notably scaffold nucleoporins, are symmetrically localized around a two-fold symmetry axis in the plane of the NE, perpendicular to the eightfold rotational symmetry about the main transport channel (Rout et al., 2000). Based on simple hydrodynamic and volumetric calculations the size of the NPC was estimated to range from 66 MDa in *S. cerevisiae* (Rout and Blobel, 1993) to 125 MDa in vertebrates (Reichelt et al., 1990). Calculations based on the stoichiometry of nucleoporins obtained in the proteomic studies, however, indicate that the NPC size is only 44 MDa in *S. cerevisiae* and ~ 60 MDa in rat. The discrepancy supports the conclusion that the NPC is a porous, lattice-like assembly, rather than a solid entity (Brohawn et al., 2008; Hinshaw et al., 1992), which accounts for the overestimate of mass based on volumetric analysis.

Dynamics

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

**Domain architecture**

Until about five years ago, very little high-resolution structural information on nucleoporins was available. This was largely due to the technical difficulties of obtaining nucleoporins of sufficient quantity and quality for structural studies, a challenge particularly severe in the case of scaffold nucleoporins. Despite the scarcity of experimental evidence, structural predictions grouped nucleoporins into a small set of fold classes (Berke et al., 2004; Devos et al., 2004; Devos et al., 2006; Schwartz, 2005). First, FG-domains, the primary transport factor interaction sites, are present in about one third of all nucleoporins. Second, coiled-coil domains are present in a number of nucleoporins. Third, scaffold nucleoporins are largely composed of β-propellers, α-helical domains, or a tandem combination of both. Using this simple classification, about 76 % of the mass of the yeast NPC was accounted for.

**FG-repeats**

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

**Coiled-coils**

Coiled-coils in the NPC fulfill significant structural roles. The nuclear basket of the NPC is mainly constructed from the large coiled-coil proteins Mlp-1/2 in yeast and Tpr in vertebrates. Coiled-coils are often used for protein-protein interactions, thus the nuclear basket may serve as a general recruitment platform to bring accessory factors close to the NPC. The desumoylating enzyme Ulpl, for example, is stably associated with the nuclear basket (Li and Hochstrasser, 2000). Lining the central NPC channel are six nucleoporins containing coiled-coil regions. The FG-Nup Nsp1 is part of two distinct entities, the Nsp1-Nup57-Nup49 complex (Grandi et al., 1993) and the Nsp1-Nup82-Nup159 complex (Bailer et al., 2001). In both, the proteins are held together by coiled-coil interactions (Bailer et al., 2001) and the Nsp1-Nup57-Nup49 complex is, in addition, tethered to the NPC scaffold via the N-terminal coiled-coil region of Nic96 (Grandi et al., 1995). So far, only a homodimerized 10 kDa fragment of Nup58 (the vertebrate orthologue to Nup57) has been structurally characterized (Melcák et al.). Biochemical analysis suggests that the network involves specific rather than promiscuous interactions, arguing for a specific tethering function for the coiled-coil segments. It will be interesting to see these coiled-coil
interactions in atomic detail in order to manipulate them and potentially swap the attached FG-domains within the NPC. Such experiments could provide important insight into the organization of the FG-network.

**β-Propellers**

A large portion of the NPC scaffold is built from β-propellers, one of the most abundant classes of proteins, especially in eukaryotes, and with diverse functions (Chaudhuri et al., 2008; Paoli, 2001). Sets of Nups were initially identified as β-propellers based on sequence analysis. In yeast, only Sec13 and Seh1 contain the signature WD-40 repeat motif and were among the very first β-propellers to be recognized (Pryer et al., 1993). More Nups have since been recognized as β-propellers despite the lack of signature sequence motifs. The N-terminal domain of Nup133 was the first experimentally determined β-propeller of the NPC and after this structure was solved, the additional non-canonical β-propeller domains in the NPC were identified (Berke et al., 2004). To date, five of the eight universally conserved β-propellers in the NPC are structurally characterized (Fig. 1.5). In Nup133, Nup120, and Nup159 (hNup214) the β-propellers are N-terminal and seven-bladed. While forming a distinct entity in Nup133 and Nup159 (Weirich et al., 2004), physically tethered but otherwise not interacting strongly with the C-terminal part of the protein, the β-propeller in Nup120 is fully integrated with an adjacent helical domain to build one continuous oblong domain (Leaks et al., 2009). Seh1 and Sec13 are so far unique variations of β-propellers in that they are open and 6-bladed (Brohawn et al., 2008; Debler et al., 2008; Fath et al., 2007; Hsia et al., 2007). Their partner proteins insert a seventh blade into the β-propeller to complete the domain in trans. The function of the β-propellers is architectural and it is widely assumed that they serve as protein-protein interaction sites. Peripheral β-propellers can recruit accessory proteins, like the mRNA export factor Dbp5
(von Moeller et al., 2009), whereas those more centrally located likely are used to connect subcomplexes.

**α-Helical domains**

α-Helical domains make up more than half of the mass of the NPC scaffold. Structural predictions have classified the non-coiled-coil α-helical domains into a strongly related group of α-helical solenoids (Devos et al., 2006). α-solenoids are characterized by a two or three helix unit that is repeatedly stacked to form an elongated, often superhelical domain with N and C terminus at opposite ends of the molecule (Kobe and Kajava, 2000). Such regular, α-helical repeat structures are, often in combination with β-propellers, common scaffolds in large protein assemblies such as the clathrin vesicle coat (Edeling et al., 2006), the protein phosphatase 2A holoenzyme (Xu et al., 2006) and the anaphase promoting complex (Herzog et al., 2009), to name a few. Surprisingly, structural characterization of α-helical domain containing Nups has revealed three different α-helical folds, each distinct from a regular α-solenoid arrangement (Boehmer et al., 2008; Jeudy and Schwartz, 2007; Leksa et al., 2009; Schrader et al., 2008b). Nic96 was the first experimentally determined α-helical structure of a scaffold nucleoporin and it showed an unexpected, atypical α-helical topology (Jeudy and Schwartz, 2007; Schrader et al., 2008). The 30 helices of the ~65 kDa domain, excluding the ~200 N-terminal coiled-coil domain, are arranged in a J-like topology, forming an oblong domain. The chain starts in the middle of the elongated domain, traverses up on one side of the molecule, folds back over a stretch of 7 helices and then continues past the N terminus to the other end of the molecule (Fig. 1.5). Three other α-helical scaffold nucleoporins (Nup84, Nup85 and Nup145C) have since been structurally characterized and shown to adopt the same fold as Nic96, pointing to a common ancestor (Brohawn et al., 2008; see below). A second, distinct α-helical fold has been identified in structures of Nup133 and Nup170, which are more distantly related, but share an extended
and stretched α-helical stack (Boehmer et al., 2008; Whittle and Schwartz, 2009), substantially different from the first group. The third was revealed in the structure of Nup120, which forms a domain that fully integrates a β-propeller with an α-helical domain (Leksa et al., 2009). The α-helical segment is built around a central stalk of two long helices wrapped with 9 additional helices in an unprecedented fashion. In summary, the α-helical domains that occur in the NPC fall in unique classes that provide a significant challenge for structure prediction methods. One obvious challenge is the exceedingly low sequence conservation, even between orthologs, apparent in the inconsistent nucleoporin nomenclature. Poor sequence conservation is likely due to some degree of malleability of the scaffold structure and the construction from common sequence elements (Aravind et al., 2006). Whether poor sequence conservation is further the result of adaptive evolution, linking several architectural nucleoporins to speciation, is an intriguing possibility that should be explored in more detail (Presgraves et al., 2003; Tang and Presgraves, 2009).

**ACE1 domains**

As mentioned above, the four α-helical scaffold nucleoporins Nic96, Nup145C, Nup85, and Nup84 are constructed around a common ~65 kDa domain composed of 28 helices (there are some non-canonical additions in each of these nucleoporins). Notably, this domain has to date only been identified outside the NPC in Sec31, one of the main building blocks of the COPII vesicle coat. The commonality was surprising. Sequence conservation between the five members is so low that no specific structural relationship was inferred previously (Alber et al., 2007, Hsia et al., 2007). This domain, which we termed Ancestral Coatomer Element 1 (ACE1), is a structural manifestation of the likely common origin of the NPC and the COPII vesicle coat (Devos et al., 2004). ACE1 is constructed from three modules, crown, trunk and tail, that together form an elongated molecule of ~140 Å x 45 Å x 45 Å. Structural superposition of ACE1 proteins shows that individual modules are closely aligned, while differences in linkers between
modules results in significant differences in their relative orientations. These differences, as well as proteolytic susceptibility data, suggest at least modestly flexible hinges connect the modules, especially the trunk and the tail. This likely explains why all crystal constructs except for Nic96 contain either the trunk and crown (Brohawn and Schwartz, 2009; Brohawn et al., 2008; Debler et al., 2008; Hsia et al., 2007) or the tail (Boehmer et al., 2008). Even with the structural information in hand, it is difficult to find additional ACE1 proteins. Beyond a few residues conserved between orthologs, ACE1 is not characterized by a distinct sequence motif. The reason for this amazing degeneracy of sequence is that for folding the ACE1 domain only some general sequence profiles need to be satisfied. For example, helices $\alpha_5$, $\alpha_7$, $\alpha_{15}$ and $\alpha_{17}$ are typically hydrophobic, because they are incased by surrounding helices and are largely buried and solvent inaccessible. Thus, a combination of sequence profile evaluation, $\alpha$-helical prediction, and overall length are currently the only indicators for the ACE1 domain. The two remaining $\alpha$-helical scaffold nucleoporins without any crystallographic structural information are Nup188 and Nup192. Whether they also belong to the ACE1 class, remains to be determined, but it appears unlikely.

**Structural characterization of NPC subcomplexes**

**NPC subcomplexes**

Most nucleoporins are organized into discrete subcomplexes each present in multiple copies that arrange according to the symmetry elements of the NPC to form the complete structure. The subcomplexes are biochemically defined and reflect the stable interaction of subsets of nucleoporins. Interestingly, these subcomplexes are also found as entities in mitotic extracts of higher eukaryotes, when the nuclear envelope breaks down during open mitosis (Matsuoka et al., 1999). At the end of mitosis, NPCs reassemble from these subcomplexes in a defined order (Dultz et al., 2008). The eight spokes are arranged around the central rotational axis and composed of 5 subcomplexes (Fig. 1.3). Nup82/Nup159/Nsp1 form a subcomplex
localized at the cytoplasmic side of the NPC (Fig. 1.3) (Belgareh et al., 1998). A second pool of Nsp1 complexes with Nup57 and Nup49 and resides in the center of the NPC, forming the bulk of the central transport barrier (Grandi et al., 1993). The scaffold ring is constructed from two major subcomplexes: the heptameric Y- or Nup84-complex and the heteromeric Nic96 complex. The Y-complex is the best-characterized subcomplex of the NPC and is essential for its assembly, as shown in several organisms (Boehmer et al., 2003; Fabre and Hurt, 1997; Galy et al., 2003; Harel et al., 2003; Walther et al., 2003a). It has 7 universally conserved components – Nup84, Nup85, Nup120, Nup133, Nup145C, Sec13 and Seh1 – that assemble stoichiometrically and exhibit the eponymous Y-shape in electron micrographs (Kampmann and Blobel, 2009; Lutzmann et al., 2002; Siniosoglou et al., 2000). In many eukaryotes, notably excluding S. cerevisiae, three additional proteins, Nup37, Nup43, and ELYS/MEL-28, are considered members of the Y-complex, but their architectural role is unclear (Cronshaw et al., 2002; Franz et al., 2007; Rasala et al., 2006). In most models, the Y-complex is thought to symmetrically localize to the cytoplasmic and the nucleoplasmic face of the NPC sandwiching the Nic96 complex. The Nic96 complex is not as well defined as the Y-complex, likely reflecting the fact that it associates less stably. However, Nic96 interacts directly with Nup53/59 (Hawryluk-Gara et al., 2005), and co-immunoprecipitation with Nup188 (Nehrbass et al., 1996) and with Nup192 have been reported (Kosova et al., 1999). Further, the Nic96 complex is the tether to the Nsp1 complex in the center of the NPC. The newest defined subcomplex contains the transmembrane Nup Ndc1, considered an anchor for the NPC in the pore membrane. This complex contains Nup157/170 and Nup53/59, which connect the Ndc1 complex to the Nic96 complex (Makio et al., 2009; Onischenko et al., 2009). The other two transmembrane Nups, Pom34 and Pom152, are reported to interact with Ndc1 as well, albeit less strongly. Mlp1/2 are attached to the NPC ring via Nup60 (Feuerbach et al., 2002) and likely form the nuclear basket structure (Strambio-de-Castillia et al., 1999).
Chapter 1: Introduction

In vitro nucleoporin subcomplex reassembly

The NPC exhibits an extraordinarily high level of symmetry, such that the study of modular and biochemically distinct NPC subcomplexes offers a feasible inroad for studying the entire NPC (Schwartz, 2005). The research described in this dissertation offers a glimpse at our progress towards reassembly of large protein subcomplexes in vitro, followed by biophysical and crystallographic analysis that allows for the description of binary nucleoporin interactions at atomic resolution. Assembly of large NPC subcomplexes in vitro requires significant knowledge of the domain architecture and secondary structural elements for every nucleoporin in question. This information allows for the identification of minimal structural elements responsible for binary protein interactions between nucleoporins. Our research has benefited greatly from increasingly powerful secondary structure prediction methods that allow us to identify and methodically probe minimal binding domains prior to recombinant protein expression and purification studies.

Purification of NPC subcomplexes for crystallographic and biophysical studies

Nucleoporins (Nups) are characteristically large proteins that are typically insoluble when recombinantly expressed in vitro. Although large and difficult to work with on an individual basis, Nups interact via minimal binding domains that can be isolated and expressed recombinantly in E. coli to yield milligram amounts of protein from 1 liter of liquid bacteria culture. The ability to produce milligram amounts of soluble protein is critical for many of the biophysical and crystallographic studies presented in this dissertation and solubility is typically improved when working with smaller minimal binding domains rather then the entire nucleoporin. In addition, many of these binding domains are largely insoluble unless coexpressed from a polycistronic vector, or coexpressed from two individual plasmids. Polycistronic vectors have become a critical tool for expression studies of large complexes and are especially useful when components are insoluble if expressed alone (Selleck and Tan, 2008; Tan, 2001). We reason
that many of these large nucleoporin complexes are unstable or improperly folded in the absence of appropriate subcomplex binding partners. For instance, the Nup62 / Nsp1 complex described herein, solubility of the trimeric complex Nup62 / Nsp1 complex is significantly improved with coexpression of all three components. Following the expression and subsequent purification of a stable trimeric complex, biophysical and crystallographic studies have provided a powerful means for further characterization of NPC subcomplex assembly.

**The Nsp1 / Nup62 subcomplex**

As with several other Nups, the Nsp1 complex was originally identified in extracts of fractionated rat liver nuclei as a biochemically stable subcomplex composed primarily of three proteins (Fig. 1.3) (Davis and Blobel, 1986; Rout and Blobel, 1993; Snow et al., 1987). The nomenclature of Nups varies between species and the Nsp1 subcomplex is not spared from this confusion. The yeast nomenclature for the three members of the Nsp1 complex is as follows: Nsp1, Nup57, and Nup49. In mammals these same proteins are named: Nup62, Nup54, and Nup58 respectively. In several instances we will compare the two trimeric complexes and we will invoke the yeast nomenclature for generalizations made between the two complexes. The size and sequence of Nsp1, Nup57, and Nup49, differ among eukaryotes, however the overall structural characteristics remain concordant. Nsp1, Nup57, and Nup49 are symmetrically localized within the NPC as judged by immunogold electron microscopy (Rout et al., 2000). Each member of the Nsp1 complex serves as an essential component for both function and structure of the NPC, as determined with deletions and point-mutants (Grandi et al., 1995; Hurt, 1988; Wente et al., 1992), as well as being essential for the formation of a functional minimal NPC structure (Strawn et al., 2004). Additional *in vitro* experiments, such as the immunodepletion of the Nsp1 complex during reconstitution of the NPC with rat nuclei and *Xenopus* egg extracts, disrupts transport of NLS tagged cargo and even the formation of NPCs
as revealed by scanning EM (Finlay et al., 1991; Mutvei et al., 1992). In addition to the three
major components of the Nsp1 complex a fourth protein, Nic96 (yeast) / Nup93 (metazoan), is
present as a partially stable component of the Nsp1 complex (Grandi et al., 1993; Grandi et al.,
1995; Guan et al., 1995). It is believed that Nic96 serves as a bridging molecule, linking
together the heterotrimeric Nsp1 complex with the more structurally significant five-membered
Nic96 subcomplex.

The Nsp1 complex occupies the central channel of the NPC and plays a key functional
role in the transport of cargo across the nuclear membrane (Fahrenkrog et al., 1998). Nsp1 is
also known to interact with Nup82 and Nup159 on the cytoplasmic face of the NPC as a
separate subcomplex (Belgareh et al., 1998), an interaction that was not part of this thesis.
Domain architecture of each component from the Nsp1 complex maintains a conserved
arrangement of secondary structure consisting of a coiled-coiled domain, approximately 200
residues in length, flanked by fiber-like extensions of unstructured amino acid sequence
containing Phe-Gly-rich (FG) domains (Hu et al., 1996; Schwarz-Herion et al., 2007). FG-repeat
domains are found in a total of ~11 different Nups and remarkably account for almost 13% of
NPC's total mass to fill the central channel of the NPC (Brohawn et al., 2009). It is well
established that FG-repeats are the primary interaction sites between the NPC and nuclear
transport receptors (NTRs) that shuttle cargo during nucleocytoplasmic transport (Denning et
al., 2003; Isgro and Schulten, 2007a, b; Macara, 2001; Peters, 2005). The coiled-coil domains
from each member of the Nsp1 complex interact to form a trimeric assembly that lines the inner-
channel of the NPC, while the FG-repeat regions are largely unstructured and as such are
nonessential for formation of the Nsp1 complex (Fig. 1.4) (Bailer et al., 2001; Finlay and Forbes,
1990; Hu et al., 1996; Strawn et al., 2004). This large trimeric coiled-coil assembly serves as a
scaffold to support the FG-repeat domains and fuse these functional elements to the core NPC
scaffold.
Although the core of the NPC is a rigid scaffold built for structural integrity, the NPC is at least in part also a dynamic macromolecular machine. Evidence is emerging to suggest that the cytoplasmic ring, the nuclear basket, and the luminal spoke ring of NPC undergo significant conformational rearrangements (Beck et al., 2007). Aside from itself being a dynamic and flexible machine, some nucleoporins are dynamic and shuttle on and off the NPC (Rabut et al., 2004). Moreover, following breakdown of the nuclear envelope during cell division in metazoa, the NPC disassembles into the discrete subcomplexes mentioned above, and is believed to reassemble in a step-wise manner at the completion of mitosis (Dultz et al., 2008; Matsuoka et al., 1999). Little is known of the binary interactions between nucleoporins that facilitate assembly between subcomplexes. Nic96 is a large helical nucleoporin that is believed to be the “linker” nucleoporin that connects the Nic96 complex with the Nsp1 complex (Fig. 1.3) (Grandi et al., 1993). Nic96 contains a coiled-coil domain at the N-terminus that facilitates a binary interaction with Nsp1 to bridge together the Nic96 and Nsp1 complexes to form a tetrameric coiled-coil assembly. The study of such a large coiled-coil assembly not only offers promise towards a better understanding of NPC subcomplex assembly, but offers a great deal of insight towards understanding how large coiled-coil interactions maintain specificity.

The interaction between Nup153 and the small GTPase Ran

_Nup153 is a versatile nucleoporin_

Although many nucleoporins share common structural domains, Nup153 is one protein that defies many of these structural classifications. Nup153 is an essential protein (Galy et al., 2003; Harborth et al., 2001) that is found in higher order eukaryotes. Its domain topology is roughly conserved between species, but distinct differences are observed (Dimaano et al., 2001; Shah and Forbes, 1998; Sukegawa and Blobel, 1993). Curiously, Nup153 is absent in single cell eukaryotes, including _Saccharomyces cerevisiae_ and _Saccharomyces pombe_., arguing for a specific role in metazoa. Aside from a direct role in regulating transport Nup153 is
important in coordinating disassembly and reformation of the NPC (Favreau et al., 1996; Walther et al., 2003b). Current data suggests that Nup153 is primarily localized to the nucleoplasmic side of the NPC and that it does interact directly with the NPC scaffold (Boehmer et al., 2003; Fahrenkrog et al., 2002; Krull et al., 2004; Pante et al., 2000; Su`kegawa and Blobel, 1993).

The unstructured N-terminal domain of Nup153 can be divided into 3 smaller sub-domains based on function and known interaction partners (Fig. 2.1). The extreme N-terminus of human Nup153, residues 1-144, is known as the nuclear envelope targeting domain (NETD) and contains a predicted amphipathic helix that directs Nup153 to the nuclear envelope (Enarson et al., 1998). The nuclear pore association region, located between residues 39-339, is responsible for targeting Nup153 to the NPC structural scaffold via the Nup107 complex (Vasu et al., 2001; Walther et al., 2003a; Walther et al., 2001). Overlapping partially with the nuclear pore association region is the RNA binding domain between residues 250-400 (Ball et al., 2007; Bastos et al., 1996; Dimaano et al., 2001). Similar to roughly a third of all nups,, Nup153 contains an FG-repeat domain at the C-terminus (residues 881-1475) and thus plays a direct role in the transport of cargo through the NPC.

**RanBP2-type zinc fingers in the eukaryotic cell**

Our primary focus for this study was the protein binding zinc finger domain located in the middle of Nup153 between residues 650-880,. The domain of the human protein consists of four C2-C2 zinc fingers that bind the small GTPase Ran (Su`kegawa and Blobel, 1993). This protein binding zinc finger domain is found in only two Nups, Nup153 and Nup358, although Nup358 harbors eight zinc fingers instead of four (Wu et al., 1995; Yokoyama et al., 1995). The Ran binding zinc fingers found in Nup153 and Nup358 (Nup358 is also known as RanBP2) have previously been referred to as the RanBP2-type of protein binding zinc fingers, and more generally fall within the family of Npl4 type zinc fingers (NZF) (Meyer et al., 2002; Plambeck et
RanBP2-type/NZF zinc fingers conform to a consensus sequence pattern: W-X-C-X(2,4)-C-X(3)-N-X(6)-C-X(2)-C (Fig. 2.2). Structural studies have described the overall fold of the RanBP2/NZF class of zinc fingers as two orthogonal β-hairpin loops with two cysteines each to coordinate a single zinc ion in the central zinc finger core. In addition to four highly conserved cysteine residues, an absolutely conserved tryptophan residue stabilizes a hydrophobic core while an absolutely conserved asparagine residue bridges the two hairpins together (Higa et al., 2007; Plambeck et al., 2003; Yu et al., 2006). Additional studies provide structural evidence for a diverse set of binding partners associated with the RanBP2 class of zinc fingers, detailing interactions made with ubiquitin, RanGDP, and RNA (Alam et al., 2004; Partridge and Schwartz, 2009; Schrader et al., 2008a). Specifically, the mammalian nuclear protein localization 4 (Npl4) protein contains a single RanBP2-type zinc finger that binds ubiquitin (Meyer et al., 2002), and with ubiquitin fusion degradation 1 (Ufd1), Npl4 forms an adaptor complex for the AAA ATPase p97 (known as Cdc48 in yeast). Together the complex between Npl4 and Ufd1 performs several processes, including nuclear envelope closure, regulated ubiquitin dependent processing, endoplasmic reticulum-associated degradation, and mitotic spindle disassembly (Cao et al., 2003; Hetzer et al., 2001). Similar to Nup153, the yeast homologue of Npl4 does not contain a ubiquitin binding zinc finger (Meyer et al., 2002; Ye et al., 2003). Aside from Npl4 other NZF containing proteins bind ubiquitin, including: Vps36 (Alam et al., 2004), and Tab2/Tab3 (Kanayama et al., 2004). Crystallographic data demonstrates that ubiquitin and RanGDP share the same binding interface on NZF/RanBP2-type zinc finger molecules (PDB codes: 1Q5W, 3GJ3, 3CH5) from Npl4 and Nup153 respectively (Alam et al., 2004; Schrader et al., 2008a), and it is clear that ubiquitin does not bind with Nup153/Nup358 zinc fingers (Higa et al., 2007). These studies clearly demonstrate that only two or three residues from each zinc finger module, a domain that is only twenty residues long, mediate ligand specificity. Another protein with a RanBP2-type zinc finger is ZRANB2 and was recently
shown to bind single-stranded RNA in a manner unique from zinc finger interactions described in structures with Npl4 and Nup153 (Loughlin et al., 2009). Binding between ZRANB2 and single-stranded RNA (ssRNA) occurs on the side of the zinc finger module, away from the "knuckle" where both ubiquitin and RanGDP bind. A tryptophan stacking interaction exists whereby a highly conserved tryptophan side chain from ZRANB2 is stacked between two bases of target ssRNA sequence. The authors speculate that this particular zinc finger module may play some role in splicing, as the recognition motif strongly resembles a 5' splice site (Loughlin et al., 2009), although the details of ZRANB2's role in the splicing mechanism are yet to be determined. Together this data highlights the significant diversity of the RanBP2-type zinc finger motif and confirms the relevance of small, modular protein binding zinc fingers in multiple processes inside the eukaryotic cell.

**Structural and biophysical analysis of Nup153 zinc fingers and RanGDP**

The master regulator of nucleocytoplasmic transport through the NPC is the small GTPase Ran. Ran, a member of the Ras superfamily, is a small protein, only 216 amino acids long and 24.4 kDa in weight, but is one-hundred percent conserved throughout eukaryotes and essential for regulation of mitosis and nucleocytoplasmic transport (Bischoff and Ponstingl, 1991a, b; Drivas et al., 1990; Melchior et al., 1993; Moore and Blobel, 1993). Transport between the nucleus and the cytoplasm occurs through an energy dependent process and the NPC acts as a physical barrier to block cytoplasmic proteins from entering the nucleus. However, following translation in the cytoplasm nuclear proteins are actively imported through the NPC and into the nucleus. This dynamic process is regulated by a concentration gradient of Ran between the nucleus and the cytoplasm, with a high concentration of RanGTP inside the nucleus and a high concentration of RanGDP in the cytoplasm. This gradient is regulated by the asymmetric localization of two regulator proteins, RanGAP and RCC1 (Fig. 1.6).
RanBP2-type zinc fingers from both Nup153 and Nup358 interact directly with Ran, however the details are uncertain. As a canonical small GTPase, Ran has two switch regions that adopt a significant conformational change when bound to either GDP or GTP (Stewart et al., 1998). RanGTP hydrolysis at the cytoplasmic face of the NPC releases export cargo from the ternary cargo-exportin-RanGTP complex, while RanGTP at the nuclear face of the NPC releases import cargo from the binary cargo-importin complex (Gorlich and Kutay, 1999). This well established gradient of RanGTP versus RanGDP across the nuclear envelope provides the energy for nucleocytoplasmic protein transport. Since both Nup153 and Nup358 are each preferentially positioned at either side of the NPC, we asked whether they might influence the Ran gradient by selectively binding to Ran in one or the other nucleotide-bound state. The existing literature on the selectivity of RanBP2-type zinc fingers is controversial (Higa et al., 2007; Nakielny et al., 1999; Yaseen and Blobel, 1999). In this dissertation we present a series of crystallographic and biophysical experiments detailing the interaction between Nup153, Nup358, and Ran.
Figure 1.1 – Overall structure of the Nuclear Pore Complex
(a) Representative micrographs of NPCs from diverse eukaryotes and obtained by scanning electron microscopy. The distinct surface features that define cytoplasmic and nucleoplasmic face of the NPC are conserved, so are the overall dimensions in the plane of the nuclear envelope. Scale bar indicates 100nm. (b) Cryo-electron tomographic (cryo-ET) reconstruction of the human NPC. The nuclear basket structure and the cytoplasmic extensions are omitted for clarity. The central eightfold rotational symmetry is clearly visible. A comparison between cryo-ET reconstructions of NPCs from diverse species (not shown) reveals substantial differences in the overall height (Elad et al., 2009).
Figure 1.2 - Cryo-electron tomography of the NPC at ~60 Å resolution
Cut away view of the NPC (colored in brown/beige) embedded in the nuclear envelope (gray). The three major domains of the NPC, the cytoplasmic filaments, the structural scaffold, and the nuclear basket, are each clearly distinguishable. This view highlights the eight-fold rational symmetry exhibited by the NPC and four distinct sections are shown in this orientation. To better highlight this, one of the repeating sections has been marked with a black dotted line (Beck et al., 2007).
Figure 1.3 - Schematic representation of the modular NPC assembly

The NPC is built from ~30 nucleoporins, organized in a small set of defined subcomplexes. This cartoon shows the major subcomplexes that make up the lattice-like scaffold (blue), the membrane-attachment (green), and the FG-network (gray) of the NPC. *S. cerevisiae* components on the left, metazoan with specific additional components on the right. A few peripheral Nups are left out for clarity. Simplified representation and connections are not to be taken literally.
### Chapter 1: Introduction

**Nucleopon Metazoan Domain Architecture Abundance Mass per NPC**

<table>
<thead>
<tr>
<th>Nucleoporin</th>
<th>Metazoan homolog</th>
<th>Domain Architecture</th>
<th>Abundance</th>
<th>Mass per NPC</th>
<th>Fraction of total NPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nup192</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35%</td>
</tr>
<tr>
<td>Nup188</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20%</td>
</tr>
<tr>
<td>Nic96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12%</td>
</tr>
<tr>
<td>Nup157/170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9%</td>
</tr>
<tr>
<td>Nup93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10%</td>
</tr>
<tr>
<td>Nup133</td>
<td>Nup133</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup120</td>
<td>Nup120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup85</td>
<td>Nup85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup145C</td>
<td>Nup96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seh1</td>
<td>Seh1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sec13</td>
<td>Sec13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mlp1</td>
<td>Tpr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mlp2</td>
<td>Tpr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup2</td>
<td>Nup50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pom152</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pom34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pom121</td>
<td>gp210</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nsp1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup159</td>
<td>Nup214/CAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gle1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup42</td>
<td>NLP1/hCG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup358</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALADIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rae1/Gle2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup145N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup116</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1.4 - Inventory of the NPC**

Summary of the nucleoporins that make up the NPC. Domain architecture of nucleoporins from *S. cerevisiae* as determined by x-ray crystallography or prediction (where structural information is still lacking). Abundance and derived mass calculations are based on published Nup/NPC stoichiometries (Cronshaw et al., 2002; Rout et al., 2000). Nucleoporins specific to metazoa are italicized.
Chapter 1: Introduction

Figure 1.5 – Structures of nucleoporins
Comprehensive list of all representative nucleoporin structures published up until July 2009. PDB accession codes are indicated. Structures are gradient-colored red- or blue-to-white from N to C terminus. Residue information for each crystallized fragment is given below the structure. Structures are shown in the assembly state that is supported by crystallographic and biochemical evidence. Structures are from *S. cerevisiae* unless noted otherwise (h, human; m, mouse; r, rat). 2QX5 (Jeudy and Schwartz, 2007);
Chapter 1: Introduction

2RFO(Schrader et al., 2008b); 3EWE(Brohawn et al., 2008); 3F3F(Debler et al., 2008); 3CQC(Boehmer et al., 2008); 314R, 315P, 315Q (Whittle and Schwartz, 2009), 3BG1(Hsia et al., 2007); 3HXR(Leksa et al., 2009); 1XKS(Berke et al., 2004); 1XIP(Weirich et al., 2004); 20IT(Napetschnig et al., 2007); 20SZ(Melcdk et al.); 1WWH(Handa et al., 2006); 1KO6(Hodel et al., 2002); 2Q5X/Y(Sun and Guo, 2008); 2BPT(Liu and Stewart, 2005); 3CH5(Schrader et al., 2008a); 3GJ3-8(Partridge and Schwartz, 2009); 1RRP(Vetter et al., 1999).

Figure 1.6 - Ran regulates nucleocytoplasmic transport
Ran is the master regulator of nucleocytoplasmic transport. RanGDP is recycled from the cytoplasm and into the nucleus by NTF2. In the nucleus RCC1 acts as a nucleotide exchange factor to reload Ran with GTP. In the nucleus RanGTP causes disruption of imported complexes between importins and cargo containing an NLS. RanGTP interacts directly with importin-β to stimulate this process. Binding of RanGTP with the exportin CRM1 promotes the assembly of an export complex formed between CRM1 and cargo containing a NES. In the cytoplasm, RanGTP interacts with the Ran GTPase activating protein (RanGAP) and RanBP2 or RanBP1 to stimulate hydrolysis of GTP and dissociate the exportin complex. RanGDP, importins, and exportins are all recycled through the NPC and the cycle begins again. Image taken from (Clarke and Zhang, 2008).
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

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


Experimental contributions: James R. Partridge conducted all experiments.
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

Introduction

Nucleocytoplasmic transport is controlled and facilitated by protein assemblies termed nuclear pore complexes (NPCs). NPCs reside in circular openings of the nuclear envelope where inner and outer nuclear membranes are fused. The NPC is a large macromolecular assembly with a calculated mass of ~50 MDa (Alber et al., 2007b). Based on electron-microscopy (EM) studies from X. laevis oocytes and S. cerevisiae, the general shape and structure of the pore is conserved across eukaryotes (Kiseleva et al., 2004; Kiseleva et al., 2001). These EM studies define the pore as a ring embedded in the nuclear envelope, exhibiting 8-fold rotational symmetry around a central axis and imperfect 2-fold symmetry between the cytoplasmic and nucleoplasmic faces. Despite its size, the NPC is only made up of ~30 proteins, or nucleoporins (Nups), arranged in a few biochemically defined subcomplexes that assemble the entire structure in a modular fashion (Schwartz, 2005). As the single constitutive barrier to regulate permeability, the NPC transports a wide range of substrates across the double membrane of the nuclear envelope (D'Angelo and Hetzer, 2008; Tran and Wente, 2006; Weis, 2003). Active transport through the NPC is mediated by nuclear transport receptors (NTRs), also called karyopherins or importins/exportins (Chook and Blobel, 2001; Cook et al., 2007).

The small G protein Ran is the master regulator of NTR-mediated, nucleocytoplasmic protein transport (Gorlich and Kutay, 1999). Ran selectively promotes binding or release of import or export cargos to NTRs by means of a chemical gradient. Ran binds mostly GDP in the cytoplasm, and mostly GTP in the nucleus. The GTPase-activating protein RanGAP, localized to the cytoplasmic face of the NPC, and the chromatin-bound GTP exchange factor, RCC1, together promote this asymmetry by modulating nucleotide hydrolysis and exchange respectively. The established gradient provides directionality to protein transport. In the nucleus, RanGTP releases import-cargo from NTRs by competitive binding. RanGTP is recycled back to
the cytoplasm via a trimeric complex formed with NTRs and export-cargo. At the cytoplasmic face of the NPC, RanGTP interacts with RanGAP to hydrolyze GTP and disrupt the trimeric NTR-mediated export complex. NTF2 (nuclear transport factor 2) recycles RanGDP back into the nucleus (Stewart, 1998).

Nup153 and Nup358 (RanBP2) are large, metazoan-specific nucleoporins with multiple roles (Ball and Ullman, 2005; Shah and Forbes, 1998; Sukegawa and Blobel, 1993; Wu et al., 1995; Yokoyama et al., 1995). Both interact with Ran through a zinc finger cassette composed of several individual zinc finger (ZnF) motifs (Fig. 2.1) (Nakielny et al., 1999; Yaseen and Blobel, 1999). Nup153 is predominantly localized to the nuclear face of the NPC, although recent studies suggest the three major domains of Nup153 are localized to different regions of the NPC (Fig. 2.1) (Fahrenkrog et al., 2002; Krull et al., 2004; Pante et al., 1994; Walther et al., 2001). The N-terminal portion contains a pore-targeting region and an RNA binding domain (Bastos et al., 1996; Boehmer et al., 2003; Dimaano et al., 2001; Ullman et al., 1999; Walther et al., 2001). The ZnF cassette harbors multiple zinc fingers and defines the center of Nup153. The C-terminal region of Nup153 harbors ~30 phenylalanine-glycine (FG-) repeats, unstructured motifs found in several Nups lining the inner channel of the NPC that are responsible for NTR interaction (Denning et al., 2003; Ribbeck and Gorlich, 2001, 2002; Shah et al., 1998). Nup358 has several characterized domains including a cyclophilin homology domain, a SUMO ligase domain, a structural leucine-rich region, previously characterized Ran binding domains (RanBDs), and a cassette containing multiple zinc fingers (Joseph et al., 2004; Pichler et al., 2002; Salina et al., 2003; Wu et al., 1995; Yokoyama et al., 1995). The ZnFs of Nup153 and Nup358 are representative of the “RanBP2-type” ZnF family, recognized by the conserved sequence pattern W-X-C-X(2,4)-C-X(3)-N-X(6)-C-X(2)-C (Fig. 2.2)(Hulo et al., 2006). RanBP2-type zinc fingers fold into a structure composed of two β-hairpin strands that sandwich a Zn$^{2+}$ ion coordinated with four cysteine residues (Gamsjaeger et al., 2007). The RanBP2-type zinc
finger structure is distinct from other zinc fingers, however it only defines a common scaffold, not a common function. Ran binding has been reported for Nup153 and Nup358 zinc fingers and not for the other structural homologs (Alam et al., 2004; Higa et al., 2007). It is unclear what role these Ran-binding zinc fingers play at the NPC, nor has it been conclusively analyzed whether they bind Ran in a nucleotide dependent or independent manner (Higa et al., 2007; Schrader et al., 2008a).

We have determined the crystal structures of all four ZnFs of Nup153 in complex with RanGDP. Our structural data suggests that all ZnF modules preferentially bind to RanGDP rather than RanGTP, supported by mutational and microcalorimetric data. While the primary sequence of the ZnF in the cassette is conserved, the number of ZnFs varies among species. Our data supports a largely uncooperative model for binding of Ran to the individual ZnF modules within Nup153 or Nup358, explaining why the exact number of consecutive ZnFs is not conserved. Although we detect differences between ZnF binding with RanGDP versus with RanGTP, they are moderate and may not be of functional consequence. We propose that the ZnFs within Nup153 and Nup358 are primarily used to create a ‘Ran sink’ and thereby increase the local concentration of Ran at both the nucleoplasmic and cytoplasmic face of the NPC.

Results and Discussion

Crystallographic analysis of Nup153-ZnF•RanGDP complexes

All protein constructs used in this study were from Rattus norvegicus, except Nup358 from Homo sapiens. The central region of Nup153 (residues 658 – 885) contains four zinc fingers; ZnF1 (residues 658 – 686), ZnF2 (residues 723 – 750), ZnF3 (residues 790 – 817), and ZnF4 (residues 848 – 885). The individual ZnF domains were cloned and recombinantly expressed as glutathione-S-transferase fusion proteins in Escherichia coli. In addition to individual ZnF domains, the tandem pairs of ZnF1 and ZnF2 (ZnF12, residues 658 – 750) as
well as ZnF3 and ZnF4 (ZnF34, residues 790 – 885) were examined (see Table 2.3 for all protein constructs used in this study). Full-length Ran was cloned and expressed as a His-tag fusion protein in *E. coli*. RanGDP was separated from RanGTP using ion-exchange chromatography and the nucleotide-loaded state was validated with HPLC. Our RanGDP structure is solved at 1.48 Å resolution with two molecules in the asymmetric unit. Molecule A is well ordered with all residues in both switch regions defined. However, in molecule B a portion of the switch II region (residues 69 – 74) is disordered. In molecule A, a van der Waals interaction between Phe77 and the neighboring molecule B orders switch II. Without this packing interaction and in the absence of the gamma-phosphate from GTP switch II is flexible.

Diffraction quality crystals of a Nup153-ZnF in complex with RanGDP were initially obtained for ZnF2, ZnF4, ZnF12, and ZnF34, but only the ZnF2 complex crystals diffracted satisfactorily. Well-diffracting crystals of all other complexes were obtained after introducing a structure-based, surface point mutation in Ran, F35S, to stabilize a crystal contact as described below. Despite significant effort, no crystals were obtained for any ZnF construct in complex with RanGTP (in the form of the GTPase deficient mutant, RanQ69L).

**Overall structure of the ZnF•RanGDP complex**

Each Nup153-ZnF and RanGDP bind with a 1:1 stoichiometry. ZnF-bound RanGDP is nearly identical to unbound RanGDP with an rmsd of 0.90 Å and 0.40 Å when compared to two RanGDP structures (Protein Data Bank (PDB) codes 1BYU (Stewart et al., 1998) and 3GJO (this study) respectively). The bound ZnF also maintains a structure similar to the unbound ZnF (rmsd 0.77 Å compared to PDB code 2GQE (Higa et al., 2007)). Four short β-strands form two orthogonal hairpins flanking the hydrophobic core containing the strictly conserved Trp7ZnF residue (for ZnF residue numbering scheme, see Fig. 2.2). A single Zn²⁺ ion is sandwiched between the two β-hairpins, and coordinated by four cysteine residues (Cys9, Cys12, Cys23, and Cys26). The side chain amide of Asn16 is highly conserved in all RanBP2-type zinc
fingers, suggesting the hydrogen bond formed between Asn16 and the backbone amide group of residue 24 helps maintain the RanBP2-type ZnF fold. Nup153-ZnF binds to RanGDP at a region neighboring and partially encompassing switch I, residues 28 to 48 (residues 65 to 84 define the switch II region, Fig. 2.3) (Vetter et al., 1999). The binding interface measures 469 Å², with major hydrophobic interactions made with Ran structural elements β1, β3, and α4, as well as electrostatic interactions with the N-terminus and switch I region of RanGDP. Hydrophobic interactions mediated by the “LVA” motif of Nup153 ZnFs, have been highlighted in previous biochemical and structural studies (Higa et al., 2007; Schrader et al., 2008a). Leu13Nup153, Val14Nup153, and Ala25Nup153 from Nup153-ZnF2 interact with RanGDP between switch I and strand β1 of RanGDP, with some interaction at the switch II region. Leu13Nup153 is situated in a hydrophobic pocket formed by the aliphatic carbon chain of Lys38Ran, Val47Ran, and Pro49Ran. Val14Nup153 interacts primarily with Trp64Ran, Lys12Ran, and with the carbon chain of Gln82Ran. Ala25Nup153 forms hydrophobic interactions with Ile81Ran, and Trp64Ran, and Leu43Ran. Critical hydrogen bonds include the backbone carbonyl group of residue 8Nup153 binding with the side chain Lys38Ran, and the backbone carbonyl of Cys26Nup153, binding with the side chain of Thr42Ran. Both Ran residues are found in the switch I region, and are known to undergo significant rearrangement in the RanGTP conformation. To summarize, the Nup153-ZnF2•RanGDP structure is principally stabilized by hydrophobic interactions with a constant region of Ran, and two hydrogen bonds with the switch I region of Ran.

Our 1.8 Å Nup153-ZnF2•RanGDP structure superimposes well with the 2.1 Å crystal structure from Schrader et al., solved in an unrelated space group (rmsd 0.55 Å compared to PDB code 3CH5 (Schrader et al., 2008a)). However, the alternative crystal-packing observed in this study results in a significantly different interpretation of parts of the ZnF•RanGDP interface. The contact between Phe714Nup153-ZnF2 and a hydrophobic pocket of Ran near β1 and β4, has been suggested to be a significant contact between Nup153-ZnF2 and RanGDP (Schrader et
al., 2008a). In our Nup153-ZnF2•RanGDP structure, Phe72\textsuperscript{Ran} from a symmetry-related Ran molecule occupies the hydrophobic pocket in question. The ZnF2 construct used in our study does not include residues 703-722, N-terminal to ZnF2 as used in Schrader et al., but is representative of the ZnF2 construct used to determine the unbound Nup153-ZnF2 NMR structure (Higa et al., 2007). In addition, the constructs and solved structures of ZnF12 and ZnF34 do include this region, yet do not bind to the hydrophobic pocket as observed in the study by Schrader et al. Our structural data, in accordance with our binding data, suggests the linker region between the ZnF domains has a rather small contribution in ZnF binding with RanGDP.

**Engineering improved crystal contacts**

The structures of RanGDP in complex with ZnF1, ZnF3, ZnF4, ZnF12 and ZnF34, respectively, were solved at high resolution only after re-engineering a crystal contact involving Ran, to improve crystal packing. In the initial orthorhombic Nup153-ZnF2•RanGDP crystals, the packing interaction was very weak along one crystallographic axis, where a single contact point was observed. We reasoned that this sub-optimal packing might be the cause for weak diffraction, high temperature factors, and high mosaicity initially observed (data not shown). The weak contact is formed by a Ran•Ran interaction, involving a single hydrogen-bond and a small, strained vdW interface between Phe35\textsuperscript{Ran} and Pro58\textsuperscript{Ran'} (Ran' denotes the symmetry mate) (Fig. 2.4). Several studies have shown that site-directed surface mutagenesis can be used to improve crystal quality (Banatao et al., 2006; Cooper et al., 2007; Mizutani et al., 2008). In our case, we reasoned that changing Phe35\textsuperscript{Ran} to serine would reduce the energetic penalty of a clash between Phe35\textsuperscript{Ran} and Pro58\textsuperscript{Ran'}, and be compatible with the hydrophilic character of the neighboring His53\textsuperscript{Ran'}. In similar fashion Phe35\textsuperscript{Ran} was mutated to aspartic acid in hopes of forming an additional H-bond with His53\textsuperscript{Ran'}. Although the RanF35D mutation proved unable to
crystallize, the RanF35S point mutation crystallized and improved the diffraction limit of ZnF4, ZnF12, and ZnF34 crystals by 1 Å, and enabled crystallization of ZnF1 and ZnF3.

The modified interface between neighboring RanF35S molecules generates a stronger network of hydrogen bonds to stabilize the Nup153-ZnF•RanGDP crystal lattice, and although Ser35\textsuperscript{Ran} is just out of range for making an H-bond with His53\textsuperscript{Ran'}, the predominant contact remains between Thr32\textsuperscript{Ran} and His53\textsuperscript{Ran'}. By removing Phe35\textsuperscript{Ran} we have alleviated the clash with neighboring Pro58\textsuperscript{Ran'} and stabilized the critical hydrogen bond between Thr32\textsuperscript{Ran} and His53\textsuperscript{Ran'}. The overall structure of Nup153-ZnF•RanGDP is unchanged by the F35S mutation, but the packing of complexes with RanF35S results in slightly different cell dimensions and orientations of symmetry axes (reducing in some crystal forms the orthorhombic to a monoclinic space group with a non-crystallographic symmetry axis).

**Comparison of the four Nup153 zinc fingers bound to RanGDP**

The RanF35S mutant enabled us to individually crystallize all four ZnF modules of Nup153 with Ran, allowing for a comprehensive analysis of the interaction. Data collection and refinement statistics for all constructs are listed in Table 1. Crystals of RanGDP alone were obtained and the structure refined to 1.48 Å (Rwork/Rfree = 18.0%/20.3%). Each of the ZnF structures in complex with RanGDP contains one Nup153-ZnF molecule bound to one molecule of RanGDP (Fig. 2.5a). The orientations of each Nup153-ZnF with respect to RanGDP are similar in the various structures (Table 2.4).

We originally reported that ZnF1 has the highest deviation (mean rmsd value of 1.07 Å) from the other ZnFs, concomitant with a very high B factor of ~200 Å\(^2\) (Partridge and Schwartz, 2009). In the meantime we found a better crystal, judged by higher resolution limit (2.4 vs. 2.7 Å) and better R\textsubscript{sym} (6.5 v. 12.3). Importantly, the new crystal is also better ordered (judged by lower mosaicity (0.32. vs. 0.85)and lower Wilson B-factor (53). The ZnF domain is better resolved, but no major novelties are observed.
In addition to conserved hydrophobic interactions, the backbone carbonyl of residue $^{10}\text{ZnF}$ and the side chain of Lys$^{38\text{Ran}}$, as well as the backbone carbonyl of Cys$^{26\text{ZnF}}$ and the hydroxyl group of Thr$^{42\text{Ran}}$ form two conserved H-bonds (Fig. 2.5a-c). A water network, identically observed in each of the four individual ZnF•RanGDP structures, further stabilizes the ZnF•Ran interaction by mediating hydrogen bonds (Fig. 2.6).

Nucleotide-dependent changes in the conformation of Ran occur at the switch I and switch II regions (Vetter et al., 1999). When bound to GDP, Ran is in the open conformation with switch I swung out, away from the nucleotide and closer to $\beta$1 (Fig. 2.3). Upon binding of GTP, the two switch regions close to accommodate the $\gamma$-phosphate. As highlighted in Fig. 2.7, residues of Ran critical for hydrogen bonding with Nup153-ZnF shift away from the ZnF binding site when Ran is in the GTP-bound conformation. Superimposing RanGTP (PDB code 1WA5) and our RanGDP structure bound to ZnFs, shows that Lys$^{38\text{Ran}}$ is shifted 26 Å away from the ZnF binding site when Ran is bound to GTP. Thr$^{42\text{Ran}}$ is buried in the RanGTP structure to interact with the magnesium ion and thus the H-bond with ZnF cannot be maintained. To highlight the influence of the conformational shift between RanGDP/RanGTP and binding of ZnFs, we have modeled a putative Nup153-ZnF2•RanGTP complex, replacing RanGDP in our structure with RanGTP (PDB code 1WA5) (Fig. 2.7) (Matsuura and Stewart, 2004). In this modeled complex, the calculated binding interface area is reduced by 27% to 345 Å$^2$. H-bonds to switch I residues are not only lost, but in the case of Thr$^{42\text{Ran}}$, are mutually exclusive between the RanGDP and RanGTP conformations, suggesting a mechanism for preferential binding of ZnF to RanGDP over RanGTP.

Apart from these interactions in the switch I region, some interactions remain unique to the individual ZnFs. The principle distinction between the binding modes of the Nup153-ZnFs is the ability or inability to form a H-bond with Gln$^{10\text{Ran}}$. Both Gln$^{8\text{ZnF1}}$ and Glu$^{8\text{ZnF3}}$ H-bond via the side chain carbonyl, while Asp$^{2\text{ZnF2}}$ and Asp$^{2\text{ZnF4}}$ do not, because the shorter aspartic acid side
Protein constructs containing tandem ZnF pairs (ZnF12, ZnF34) have been crystallized in the same condition as the individual zinc finger domains. The electron density map for each crystal shows only one ZnF bound to RanGDP in agreement with the structural interactions described above for single ZnF constructs. In these structures, predominantly the second ZnF (ZnF2 and ZnF4) is seen to bind with RanF35SGDP, even though, considering the crystal packing, ZnF1 and ZnF3 could be accommodated as well. Since ZnF2 and ZnF4 tentatively bind weaker to RanGDP than ZnF1 and ZnF3 (see below), we would expect that binding of the N-terminal ZnF should be preferred. We conclude that the crystal packing favors binding of the C-terminal ZnF in the tandem constructs. Since a second zinc finger is not visible in the electron density we confirmed that the tandem Nup153-ZnF12 polypeptide is intact based on SDS-PAGE analysis of the Nup153-ZnF12-RanGDP crystals (Fig. 2.12). We also analyzed the anomalous signal from Zn at the measured (high-energy remote) wavelength of 0.98 Å (Zn K-edge 1.2837 Å), but could only make out the signal of the Zn atom in the bound ZnF (data not shown). To distinguish which ZnF was bound, both possibilities were modeled and refined. R factors were consistently lower for the C-terminal ZnF and the difference density showed reduced noise (not shown). Interestingly, and in contrast to the study by Schrader et al., we do not observe any interaction of the linker residues between ZnF1 and ZnF2 with Ran.

**Isothermal titration calorimetry**

To expand upon our crystallographic analysis of the Nup153-ZnF interaction with RanGDP, we performed binding assays using isothermal titration calorimetry (ITC). These experiments demonstrate that individual Nup153-ZnF domains bind to RanGDP with modestly varying affinity (Table 2.2). ZnF1 / ZnF3 have comparable binding characteristics that differ from those measured for ZnF2 / ZnF4. ZnF1 is shown to bind RanGDP with a measured $K_d$ of
6.5 μM and ZnF3 binds with a $K_D$ of 6.6 μM, in an exothermic reaction. ZnF2 / ZnF4 have a lower binding affinity for RanGDP, with 49 μM and 47 μM respectively, in an endothermic reaction. The enthalpic differences in RanGDP binding of these ZnFs is likely due to conformational contributions, thus not readily explainable based on our structures. The variation in enthalpy between sites is also measured with the ZnF pairs and the full Nup153 ZnF domain as described below and in Table 2.2.

Structural data suggests that the missing hydrogen bond between Gln10$^{\text{RanGDP}}$ and residue 8$^{\text{ZnF}}$ may be responsible for this differences. In ZnF2 / ZnF4 Asp8 is too short to H-bond with Gln10$^{\text{RanGDP}}$, in contrast to Glu/Gln8 in ZnF1 / ZnF3 which forms this bond (Fig. 2.6). When residue 8$^{\text{ZnF}}$ was mutated to a glutamine in ZnF2 and ZnF4, the binding affinity for RanGDP increased to values very close to those measured for ZnF1 and ZnF3 (Fig. 2.8a, 6b, 6c, 6d). Nup153 was also tested to compare the differences in affinity for RanGDP vs. RanGTP (in the form of the GTPase deficient mutant, RanQ69L) (Fig. 2.8e). We next asked whether individual ZnF domains exhibit an allosteric effect on their neighbors. Pairs of ZnFs from Nup153 and Nup358 (Nup153-ZnF12, Nup153-ZnF34, and Nup358-ZnF12) were assayed for RanGDP binding (Fig. 2.9). A two-site model best fit the data in each experiment with Nup153. The two binding sites in the tandem constructs exhibit affinities that are similar to the values measured for the individual ZnFs, indicating no significant allosteric effect. The data also suggests that the flexible linker between ZnF domains does not significantly contribute to binding affinity for Nup153-ZnF2, as our data measured with and without the linker agree with previously measured affinity (Schrader et al., 2008a). Additionally, the binding of Nup358-ZnF12 to RanGDP was measured to characterize ZnFs of Nup358 in comparison with those of Nup153. Based on the high level of sequence conservation between Nup153 and Nup358 (Fig. 2.10), it is not surprising to find that ZnFs from Nup358 behave similarly. The ITC data for Nup358-ZnF12 is best fit to a single site model, with two molecules of RanGDP binding per molecule of
Nup358-ZnF12. This corresponds to each ZnF binding independently, similar to Nup153, however differs from Nup153 because both ZnFs of Nup358 lack the Glu/Gln8 residue needed to H-bond with Gln10$^{\text{RanGDP}}$. To summarize the data from our calorimetric experiments, we compared binding of the entire ZnF regions of Nup153 to RanGDP or RanQ69LGTP (Fig 2.8). Nup153 binds RanGDP in a 2-site model with affinities of 1 µM and 8 µM, with two molecules of RanGDP binding per site. This accounts for each ZnF binding with RanGDP independently for a total of four molecules of RanGDP binding with the ZnF domain of Nup153, in agreement with published data (Higa et al., 2007). RanQ69LGTP binds with lower affinity to Nup153, with one site showing negligible binding and the other exhibiting a $K_D$ of 20 µM in an independent 2-site model. The Nup358-ZnF cassette binds both GDP- and GTP-bound Ran, and again, the interaction with RanGDP is measurably stronger (Fig. 2.9).

**Conservation of RanBP2-type ZnF cassettes in nucleoporins**

We have performed a phylogenetic analysis to put our structural and biochemical data in an evolutionary context. The RanBP2-type ZnF domain can be readily recognized by its characteristic signature sequence motif (Fig. 2.2). The ZnF domain is found in all Nup153 and Nup358 homologs, which are exclusively present in animals, but absent in plants and fungi. RanBP2-type ZnF cassettes usually contain 4 repeats in Nup153, and 8 repeats in Nup358. Exceptions are however not infrequent. The number in sequenced Nup153 homologs varies between 2 and 6, while in Nup358 we found between 2 and 8 ZnFs (data not shown). These variations match our observation that the ZnF modules function largely independently and not in a cooperative fashion. Similarly, we fail to find a specific signature that would define an order in which the ZnF modules are arranged within the ZnF region. We also asked whether there is a distinction between ZnFs present in Nup153 compared to those in Nup358. Given the asymmetric distribution of Nup153 and Nup358 at the nucleoplasmic and cytoplasmic side of the NPC, respectively, one could hypothesize that the ZnFs exhibit locus-specific tasks.
However, we do not find such differences. The sequence conservation of the Ran-binding ZnF matches well among structurally and functionally important residues. The structurally important residues Trp7, Cys9, Cys12, Asn16, Cys23 and Cys26 are strictly conserved (Fig. 2.10). The residues involved in hydrophobic interactions with Ran Leu13, Val14, and Ala25, are also quite well conserved across all animals and the functional significance has been recognized even before structural data was available (Higa et al., 2007). Comparing just the linker between individual ZnFs shows no sequence conservation, consistent with it not having a significant role in protein-protein interactions.

**Comparison of binding interactions among RanBP2-type zinc fingers**

RanBP2-type zinc fingers are recognized by the conserved sequence W-X(2)-C-X(3)-N-X(6)-C-X(2)-C as shown in Fig. 2.2. This sequence signature only defines the structural scaffold, but excludes the residues important for binding interactions. Currently three different binding partners are known. The Nup153/Nup358 class binds Ran, the Npl4 class binds ubiquitin and the ZRANB2 class binds single-stranded RNA (ssRNA) (Schrader et al., 2008a; Wang et al., 2003). Phylogenetic analysis of these RanBP2-ZnF classes readily reveals that each type of ZnF has additional conserved residues (Fig. 2.10). For Nup153/Nup358, these are the discussed residues Leu13, Val14 and Ala25 (Higa et al., 2007). In the Npl4-type, the same ZnF surface binds ubiquitin involving the conserved residues Thr13, Phe14 and Met25. By switching these residues the Nup153 ZnF can be converted into a ubiquitin-binding moiety (Higa et al., 2007). The ssRNA-binding splicing factor ZRANB2 is the best conserved of the three known RanBP2-ZnF classes. The publication of the ssRNA-bound ZRANB2 structure (PDB code 3G9Y) details ssRNA bound at a different surface of the scaffold via the highly conserved residues between residues 14 and 20 (Loughlin et al., 2009). Considering the sequence signature of the Nup153/Nup358-ZnF class it is unlikely that it can bind nucleic acids, as was reported when Nup153 was first described (Sukegawa and Blobel, 1993). In summary, the
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

RanBP2-type zinc finger emerges as a scaffold for various binding partners and it will be interesting to learn about the interacting partners of the additional, less characterized classes, including Sharpin, Mdm2, and Mdm4 (Lim et al., 2001; Yu et al., 2006).

Discussion

Here, we present a comprehensive analysis of the interaction of the Nup153/Nup358 ZnF class with Ran. We identify individual residues within the four ZnFs of mammalian Nup153 that modulate the binding affinity to Ran. The G protein is preferentially bound in the GDP-bound form, because switch I residues are involved in critical hydrogen bond interactions with ZnF and because switch I contributes about one-third of the total interaction surface. Nonetheless, binding to RanGTP is also observed and the local RanGTP to RanGDP ratio likely determines the nucleotide-binding state of ZnF-bound Ran. We cannot detect a significant difference between the binding behavior of Nup358 and that of Nup153 fingers, arguing for a common function for both proteins. Neither Nup358 nor Nup153 are universally conserved nucleoporins, but are specific to animals. We suggest that the ZnF moieties are used to increase the local concentration of Ran at both the nucleoplasmic and cytoplasmic face of the NPC. This may accelerate nucleocytoplasmic protein transport by keeping Ran close to the NPC, a sophistication that may be dispensable in unicellular eukaryotes, or potentially replaced by a separate mechanism, as proposed for plants (Meier et al., 2008). Our phylogenetic analysis suggests that other RanBP2-type ZnFs do not bind Ran, since they share only the structurally important residues and use distinct binding interfaces for their respective protein and nucleic acid interactions.

Materials and Methods
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

**Protein purification**

Bacterial expression constructs for Nup153 domains from *Rattus norvegicus* and Nup358 from *Homo sapiens* were cloned as glutathione-S-transferase (GST) fusion proteins in the pGEX-6P1 vector (GE Healthcare) (Table 2.1). Ran was expressed as a 6xHis-tag fusion protein from a pET-28a vector, engineered to contain a protease 3C site after the N-terminal affinity tag. All proteins were expressed in *Escherichia coli* strain BL21(DE3) RIL (Stratagene).

Bacterial cell pellets harboring GST fusion proteins were suspended in 20 mM potassium phosphate pH 7.0, 150 mM NaCl, 2 mM DTT and lysed using a french press. The crude lysate was supplemented with 200 μM phenylmethanesulfonyl fluoride (PMSF) and centrifuged at 15 000 g for 15 minutes. Soluble protein was mixed with 0.5ml of glutathione Sepharose beads (4 Fast Flow, GE Healthcare) per 1000 OD of cells for 2 hrs at 277 K. After three batch washes in resuspension buffer, resin was washed in ZnCl₂ buffer (10 mM Tris/HCl pH 8.0, 100 mM NaCl, 10 μM ZnCl₂, 2 mM DTT), and Nup153 proteins were eluted directly from the resin by incubating with protease overnight at 277 K. The eluted protein was purified by anion exchange chromatography on a HiTrapQ column (GE Healthcare) via a linear NaCl gradient and size exclusion chromatography using a Superdex S75 26/60 column (GE Healthcare) run in 15 mM Tris/HCl pH 8.0, 150 mM NaCl 1 mM MgCl₂, 5 μM ZnCl₂, 1.5 mM DTT. All truncations of the full-length Nup153-ZnF domain, as well as point mutations, were generated with PCR mutagenesis and purified as described above.

Bacterial cell pellets harboring His-tagged Ran were suspended in 20 mM Tris/HCl pH 8.0, 200 mM NaCl, 5 mM imidazole, 3 mM β-mercaptoethanol (β-ME), and lysed using a french press. Crude lysate was supplemented with 200 μM phenylmethanesulfonyl fluoride and clarified by centrifugation at 15 000 g for 15 minutes. The soluble fraction was then incubated with 1ml Ni-NTA per 1000 ODs for 1 hour at 277 K and loaded onto a disposable column (Pierce). The column was washed with 4 bed volumes of 10 mM Tris/HCl pH 8.0, 400 mM NaCl, 10 mM imidazole, 3 mM β-ME, and eluted with 6 bed volumes of 10 mM Tris/HCl pH 8.0,
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

50 mM NaCl, 150 mM imidazole, 3 mM β-ME. Eluted protein was dialyzed against 10 mM Tris/HCl pH 8.0, 100 mM NaCl, 1 mM DTT, for 1 hour before the 6xHis-tag was cleaved. RanGDP was separated from RanGTP using anion exchange chromatography on a HiTrapQ column (GE Healthcare) via a linear NaCl gradient, followed by purification with size exclusion chromatography using a Superdex S200 26/60 column (GE Healthcare) in 15 mM Tris/HCl pH 8.0, 150 mM NaCl 1 mM MgCl₂, 5 μM ZnCl₂, 1.5 mM DTT. RanQ69L and RanF35S were generated with PCR mutagenesis and purified as described above. The nucleotide-bound state of Ran was confirmed by HPLC, on a analytical C₁₈ column. The nucleotide was released from Ran by heat denaturation at 369 K and centrifugation to separate coagulated protein from soluble nucleotide. The soluble fraction was loaded on a HPLC and GTP was separated from GDP with a linear acetonitrile gradient in 50 mM triethanolamine/HCl pH 7.5, 10 mM tetrabutylammonium-sulfamate (TBA).

Crystallization

Ran was concentrated to 15 mg/ml using Vivaspin 20 concentrators (Sartorius) and mixed with Nup153-ZnF protein at equimolar concentrations. All Nup153-ZnF fragments in complex with RanGDP were crystallized in the same condition (0.1 M BisTris/HCl pH 6.5, 18-20% PEG3350) using the hanging-drop method and mixing 1 μL protein with 1 μL of reservoir solution at 291 K. Crystals of wt-RanGDP in complex with ZnF2 grew in 5-7 days forming birefringent plates of 400 x 200 x 30 μm. RanGDP crystals formed on the edges of Nup153-RanGDP complex crystals after two weeks, as small 150 x 150 x 150 μm bipyramidal crystals. The crystallization of tandem zinc finger constructs Nup153ZnF12 and Nup153ZnF34 was confirmed by fishing crystals from crystallization drops, washing the crystals with mother liquor, dissolving the crystals in SDS loading buffer, and analyzing protein composition of the crystal with SDS-PAGE. Crystals of Nup153 ZnF constructs in complex with RanF35S-GDP grew in 1-2 days. All crystals were cryoprotected by adding 12% (v/v) glycerol to the crystallization solution.
before flash freezing in liquid nitrogen. All datasets were collected at the NE-CAT beamlines 24ID-C and 24ID-E at Argonne National Laboratory. Crystal parameters and data collection statistics are listed in Table 2.1.

**Structure determination**

Data reduction was carried out using HKL2000 (Minor, 1997). The initial Nup153-ZnF2-RanGDP structure was solved by molecular replacement, using RanGDP as the search model (PDB code 1BYU). All other structures were solved by molecular replacement using our 1.48 Å RanGDP structure as the search model. Model building was carried out using COOT (Emsley and Cowtan, 2004). PHENIX (Adams et al., 2002) was used for all refinement steps. For the tandem ZnF constructs only one zinc finger domain was visible. The second ZnF is present in each tandem ZnF construct as determined by analysis of the difference maps and refinement statistics with the respective ZnF sequences. Proper assignment of the correct ZnF correlates with better refinement statistics, judged by lower $R$ and $R_{free}$ values. Data collection and refinement statistics are summarized in Table 2.1. All figures were made using PyMOL (DeLano, 2002).

**Isothermal titration calorimetry**

RanGDP and all ZnF constructs were purified and dialyzed into the same buffer containing 15 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM MgCl₂, 5 μM ZnCl₂, 1.5 mM DTT, and either 2.5 μM GDP or 2.5 μM GTP as indicated, prior to ITC. Protein concentrations were determined spectrophotometrically at 280 nm immediately before the experiment. ITC was performed using a VP-ITC microcalorimeter (MicroCal, Northhampton, MA). Titrations were performed at 293 K, or 278 K, by injecting 14 μL aliquots of RanGDP, or RanQ69LGTP (Table 2.2), into the ITC cell containing 1.43 ml of Nup153 or Nup358 ZnF proteins (Table 2.2). Binding stoichiometry, enthalpy and entropy, as well as the equilibrium binding dissociation
constant was determined using a "one site model" for individual zinc fingers and Nup358 ZnF12, or a "two sets of independent sites" model for all other experiments, to define molecular association in the software suite, MicroCal Origin 2.9 (MicroCal).

**Sequence analysis**

Sequences were aligned using MUSCLE (Edgar, 2004) and edited in JaIVIEW (Waterhouse et al., 2009). Sequence logos were made using WebLogo (Crooks et al., 2004; Schneider and Stephens, 1990).

**Protein Data Bank accession numbers**

Coordinates and structure factors for all crystal structures have been deposited in the PDB (IDs 3GJ0, 3GJ3, 3GJ4, 3GJ5, 3GJ6, 3GJ7, 3GJ8).

**Acknowledgements**

We thank the staff at beamlines 24-ID-C/-E at Argonne National Laboratory and X6A at the National Synchrotron Light Source for assistance and data collection, past and present members of the Schwartz lab for discussions and editing, especially T. Boehmer for early experimental advice and constructs. We would like to thank Katie Ullman for the kind gift of Nup358. The Biophysical Instrumentation Facility for the Study of Complex Macromolecular Systems (NSF-0070319 and NIH-GM68762) provided instrumentation. Supported by a Pew Scholarship (T.U.S.) and the Herman Eisen Fellowship in Biology (J.R.P.).
Figure 2.1 - Domain architecture of Nup153
Nup153 is a large nucleoporin with a diverse set of functions. The N-terminal domain between residues 1 and 650 contains three functional regions: the nuclear envelope targeting domain (NETD), the nuclear pore associating region and the RNA binding domain. At the C-terminus is a large unstructured FG-repeat domain between residues 880 and 1475. Between residues 650 and 880 is a Ran-binding zinc finger domain consisting of four individual RanBP2-type zinc finger modules. Each zinc finger module contains 30 residues and between each module is an unstructured stretch of ~35 residues.

Figure 2.2 - Alignment of the four Ran-binding zinc fingers of Nup153 from R. norvegicus
The numbering used for individual zinc fingers is listed above. Identical residues are highlighted in dark blue, with decreasing levels of conservation highlighted in lighter shades of blue. Below is the consensus sequence for the "RanBP2" family of zinc finger proteins.
Figure 2.3 - Ribbon diagram representation of the Nup153-ZnF2-RanGDP complex
RanGDP β-sheets are colored dark blue with α-helices colored light blue. The Nup153-ZnF molecule is colored orange. Both the switch I and switch II regions of Ran are highlighted. Representative of the “RanBP2” family of protein-binding zinc fingers, the Nup153 ZnF contains two orthogonal β-hairpins with four cysteine residues, colored yellow, to coordinate a single Zn$^{2+}$ ion colored grey. The zinc finger binds near the switch I region of RanGDP, indicated here by the secondary structural element β2. The GDP nucleotide is shown in the center of Ran with a bound Mg$^{2+}$ ion.
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

Figure 2.4 - Crystallographic lattice formed in the crystallization of a Nup153-ZnF-RanGDP complex

Residues highlighted in red indicate contacts made between neighboring Ran molecules down a single plane of the P21 lattice. This plane represents the smallest crystallographic interface and was chosen for crystal engineering to stabilize packing, in hopes of improving diffraction. The enlarged 3D image on the right depicts residues binding at the interface. Thr32' and His53 make a single hydrogen bond, mirrored across a 2-fold symmetrical face as indicated. A clash between Phe35' and Pro58, though tolerated, was theorized to partially counteract the strong interaction between Thr32' and His53. A point-mutation of Phe35' to Serine stabilizes the crystal contact and increases the resolution of our crystallographic experiments by 1 Å.
Figure 2.5 - Comparison of the four individual zinc fingers with RanGDP

Zinc fingers (ZnF), colored orange, in complex with RanGDP, colored blue. The four individual ZnFs are overlaid to highlight differences and similarities in binding with Ran. Residues that make intermolecular hydrogen bonds are colored in green, while residues that facilitate hydrophobic interactions are colored in red.
interactions are colored violet. (a) A overview of the Nup153-ZnF●RanGDP complex showing that each ZnF binds independently and at the same location with RanGDP. Hydrogen bonds conserved among the four ZnFs are labeled. The variable hydrogen bond at residue $8^{\text{ZnF}}$ with Gln$10^{\text{Ran}}$ is only formed between Ran and ZnF1 or ZnF3. (b) An enlarged view of all residues responsible for the hydrophobic interaction at the Nup153-ZnF●RanGDP interface. (c) An enlarged view of hydrogen bonds, in green, that facilitate the Nup153-ZnF●RanGDP interaction. Conserved H-bonds are made between the carbonyl of Cys$26^{\text{ZnF}}$ with the side chain hydroxyl group of Thr$42^{\text{Ran}}$, and Lys$38^{\text{Ran}}$ with the carbonyl at residue $10^{\text{ZnF}}$. A variable H-bond is present at residue $8^{\text{ZnF}}$ binding with Gln$10^{\text{Ran}}$. A stable bond is formed between Gln$8^{\text{ZnF1}}$ and Gln$10^{\text{Ran}}$, as well as between Glu$8^{\text{ZnF3}}$ and Gln$10^{\text{Ran}}$. Residue 8 of ZnF2 and ZnF4, colored dark-blue, contain an Asp residue, too short to make the H-bond with Gln$10^{\text{Ran}}$. 
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

Figure 2.6 - Conserved water network and non-conserved hydrogen bond
(a) Conserved water network at the interface between Nup153-ZnF and RanGDP. Residues facilitating these contacts are labeled and highlighted in green, with water molecules colored violet. (b) A more detailed view of the non-conserved hydrogen bond made with Gln10\textsuperscript{Ran} by Gln8\textsuperscript{ZnF1} and Glu8\textsuperscript{ZnF3}, colored green. Asp8\textsuperscript{ZnF4} and Asp8\textsuperscript{ZnF4}, colored blue, are unable to bridge the distance necessary to make the H-bond with Gln10\textsuperscript{Ran}, a calculated distance of \textasciitilde 5.0 Å.
Figure 2.7 - Surface representation of RanGDP / RanGTP with or without Nup153-ZnF
RanGDP or RanGTP, grey, with or without Nup153-ZnF, blue ribbon-diagram. Residues involved in
intermolecular contact are colored orange. (a) ZnF•RanGDP complex. Lys38Ran and Thr42Ran, located in the switch I region, form conserved H-bonds with the ZnF, and are, known to undergo significant conformational changes upon binding of GTP. Both residues are positioned to interact with ZnF in this conformation. (b) RanGDP interaction surface in the absence of ZnF. (c) RanGTP in absence of modeled ZnF. Lys38Ran and Thr42Ran are displaced up to 30 Å, and are no longer able to interact with the Nup153-ZnF in this conformation.
Point-mutations are used to probe the variable H-bond interaction with Gln10^{Ran}, determined to be the interaction responsible for discrepancies in affinities measured between ZnFs and RanGDP. Data for independent ZnFs is fit to a single site model with $N=1$. (a) WT-ZnF1 is shown in red compared with mutated ZnF1 in black. Mutating Gln8 to Asp is shown to have a minor effect on binding with Gln10^{Ran}. (b) WT-ZnF2 is shown in red with the mutant protein shown in black. WT-ZnF2 has low affinity binding measured at 47 μM, however by mutating Asp8 to Gln high affinity binding is measured at 4 μM. Note that the two titrations were performed at different temperatures, 278K (red, WT) 293K (black, mutant). No signal was recorded for WT at room temperature (data not shown). (c) WT-ZnF3 is shown in red and mutant ZnF3 is shown in black. Mutating Glu8 to Asp is enough to decreased the measured affinity by disrupting a single H-bond with Gln10^{Ran}. (d) WT-ZnF4 is shown in red and has a measured binding constant of 47 μM. However, mutating Asp6 to Gln introduces a new H-bond with Gln10^{Ran} and we subsequently measure higher affinity binding. Note that the two titrations were performed at different temperatures, 278K (red, WT) 293K (black, mutant). No signal was recorded for WT at room temperature (data not shown). (e) Binding of the WT Nup153-ZnF domain to GDP (in red) is compared to binding...
with RanGTP (in black). In agreement with our structural observations, RanGDP binds ZnFs with strong and weak sites, while RanGTP binds with the same stoichiometry, but only with weak affinity. Experimental values for N, $K_d$, enthalpy, $\Delta H$, and entropy, $TAS$, are listed in Table 2.2. Constructs are described in Table 2.3.
Figure 2.9 - ITC illustrating low micro-molar affinity between RanGDP and two tandem ZnF pairs from the Nup153- and Nup358-ZnF domain

(a) Interaction between RanGDP and ZnF12. The data is best fit to a two-site model suggesting each ZnF binds independently, but with different affinities. (b) Interaction between ZnF34 and RanGDP, again suggesting two independent sites with different affinities. (c) Interaction between Nup358-ZnF12 and RanGDP. As with Nup153, the ZnFs of Nup358 bind with low μM affinity to RanGDP. The data is best fit to a single site model in accordance with both ZnFs from Nup358 lacking the residue necessary at pos8 to H-bond with Gln10Ran. Experimental values for N, Kd, enthalpy, ΔH, and entropy, TΔS, are listed in Table 2.2. Constructs are described in Table 2.3.
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

Figure 2.10 - Weblogos from members of the RanBP2 class of zinc fingers known to have unique binding partners

The logos highlight similarities and between the three groups of zinc fingers: the Ran binding Nup153/Nup358 group, the ubiquitin binding Npl4 group, and the single-stranded RNA binding ZRANB2 group. The logos demonstrate an overall conservation of the canonical RanBP2 type zinc finger sequence W-X-C-X(2,4)-C-X(3)-N-X(6)-C-X(2)-C, while highlighting regions important for facilitating binding with a unique binding partner. (a) ZnFs from Nup153 and Nup358 that bind with Ran. Only the first finger from each protein has been used in the alignment. The previously described “LVA,” in positions 13, 14, and 25, have low information content, although position 25 has been shown influence Ran binding (Higa et al., 2007). (b) The logo for Npl4 again demonstrates the elevated conservation of positions 13, 14, and 15, previously shown to modulate interaction with ubiquitin. (c) The ZRANB2 logo highlights the strong conservation this sequence in eukaryotes. A structure of ZRANB2 (PDB codes 3G9Y and 2K1P), demonstrates the strict conservation and the preponderance of positive charges at positions 14-20 that regulate binding of single-stranded RNA (Loughlin et al., 2009). Each logo is based on alignments of sequences from a diverse group of eukaryotes.
Figure 2.11 - Sequence alignment of the human Nup358/RanBP2 and Nup153 ZnFs, including the respective C-terminal linker regions

(a) Conservation of the "RanBP2-like" ZnF motif (residues 1343-1833) and C-terminal linker (residues 650-913) between Nup358 and Nup153 from human. Strict conservation is shown in dark-blue, while lighter shades of blue depict decreasing levels of conservation. The linker regions between these ZnFs is neither conserved in sequence or length. (b) Weblogo of the same alignment showing conservation in the ZnF motif, and lack thereof in the linker region.
Figure 2.12 - Nup153ZnF12•Ran crystal contains Ran and intact Nup153ZnF12
In the electron density map for Nup153ZnF12•Ran we can only detect one visible zinc finger. To ensure that this is not the result of protein degradation the crystals used for data collection were analyzed by SDS-PAGE to confirm that the Nup153ZnF12 polypeptide has not been degraded. Here we show three different crystals that have been fished from a crystallization drop and run on a 17% poly-acrylamide gel, Crystal 1, Crystal 2, and Crystal 3. In addition, a gradient of the Nup153ZnF12•Ran protein solution used for crystallization has been included for reference. This data shows that the crystals with tandem Nup153ZnF12 polypeptide (13.5 kDa) contain an intact Nup153 polypeptide in addition to Ran.
Table 2.1 - X-ray crystallographic data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data Set</th>
<th>RanGDP</th>
<th>Nup153ZnF12</th>
<th>Nup153ZnF34</th>
<th>Nup153ZnF1</th>
<th>Nup153ZnF2</th>
<th>Nup153ZnF3</th>
<th>Nup153ZnF4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB Code</td>
<td>3GJO</td>
<td>3GJ7</td>
<td>3GJ8</td>
<td>3GJ6</td>
<td>3GJ3</td>
<td>3GJ4</td>
<td>3GJ5</td>
</tr>
<tr>
<td>Data Collection</td>
<td>C</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P4,2,2</td>
<td>P2,1</td>
<td>P2,1</td>
<td>P2,2,2</td>
<td>P2,2,2</td>
<td>P2,1</td>
<td>P2,1</td>
</tr>
<tr>
<td>Cell dimensions:</td>
<td>a,b,c (Å)</td>
<td>81.5, 81.5, 130.5</td>
<td>58.5, 61.1, 80.2</td>
<td>74.3, 61.7, 70.6</td>
<td>60.1, 80.3, 54.6</td>
<td>59.7, 80.2, 58.0</td>
<td>68.4, 61.6, 72.1</td>
</tr>
<tr>
<td>β (°)</td>
<td>90</td>
<td>93.7</td>
<td>112.3</td>
<td>90</td>
<td>90</td>
<td>110.0</td>
<td>112.8</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>71,145</td>
<td>39,366</td>
<td>52,824</td>
<td>18,561</td>
<td>26,894</td>
<td>29,511</td>
<td>53,894</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50-1.48</td>
<td>50-1.95</td>
<td>50-1.82</td>
<td>50-2.40</td>
<td>50-1.80</td>
<td>50-2.15</td>
<td>50-1.79</td>
</tr>
<tr>
<td>(1.53-1.48)</td>
<td>(2.02-1.95)</td>
<td>(1.89-1.82)</td>
<td>(2.46-2.40)</td>
<td>(1.86-1.80)</td>
<td>(2.23-2.15)</td>
<td>(1.85-1.79)</td>
<td></td>
</tr>
<tr>
<td>Rsym</td>
<td>7.5 (99.9)</td>
<td>9.6 (82.7)</td>
<td>6.9 (51.6)</td>
<td>6.5 (44.7)</td>
<td>7.2 (73.3)</td>
<td>9.4 (45.8)</td>
<td>6.3 (33.3)</td>
</tr>
<tr>
<td>Completeness</td>
<td>95.6 (93.0)</td>
<td>93.3 (90.1)</td>
<td>99.3 (99.3)</td>
<td>94.7 (76.2)</td>
<td>99.6 (99.4)</td>
<td>95.3 (92.4)</td>
<td>98.1 (97.2)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>9.7 (7.7)</td>
<td>2.3 (1.7)</td>
<td>3.2 (3.0)</td>
<td>2.0 (2.0)</td>
<td>7.8 (7.7)</td>
<td>2.7 (2.8)</td>
<td>2.7 (2.7)</td>
</tr>
<tr>
<td>I/σ</td>
<td>32.3 (1.8)</td>
<td>12.9 (1.2)</td>
<td>19.2 (2.0)</td>
<td>22.9 (2.3)</td>
<td>30.2 (2.9)</td>
<td>15.2 (2.5)</td>
<td>18.6 (3.0)</td>
</tr>
<tr>
<td>Wilson B factor</td>
<td>19.6</td>
<td>37.0</td>
<td>24.3</td>
<td>53.3</td>
<td>26.4</td>
<td>34.9</td>
<td>23.2</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>30 - 1.48</td>
<td>30 - 1.93</td>
<td>30 - 1.82</td>
<td>30 - 2.40</td>
<td>30 - 1.78</td>
<td>30 - 2.15</td>
<td>30 - 1.79</td>
</tr>
<tr>
<td>Nonhydrogen Atoms</td>
<td>3264</td>
<td>3689</td>
<td>3714</td>
<td>1851</td>
<td>1854</td>
<td>3652</td>
<td>3654</td>
</tr>
<tr>
<td>Water molecules</td>
<td>478</td>
<td>294</td>
<td>548</td>
<td>46</td>
<td>250</td>
<td>206</td>
<td>458</td>
</tr>
<tr>
<td>Rwork/Rfree (%)</td>
<td>18.0 /20.3</td>
<td>18.6 /22.9</td>
<td>17.0 /20.3</td>
<td>19.0 /26.6</td>
<td>16.1 /19.4</td>
<td>20.5 /25.7</td>
<td>19.9 /23.3</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td>Bond lengths (Å)</td>
<td>0.008</td>
<td>0.007</td>
<td>0.005</td>
<td>0.008</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>2.4</td>
<td>1.06</td>
<td>0.943</td>
<td>1.12</td>
<td>1.11</td>
<td>1.02</td>
<td>1.07</td>
</tr>
<tr>
<td>B factors (Å²)</td>
<td>ZnF</td>
<td>-</td>
<td>86.0</td>
<td>58.2</td>
<td>132.5</td>
<td>50.7</td>
<td>71.5</td>
</tr>
<tr>
<td>Ran</td>
<td>24.8</td>
<td>51.1</td>
<td>32.9</td>
<td>54.7</td>
<td>33.4</td>
<td>48.5</td>
<td>40.5</td>
</tr>
<tr>
<td>GDP</td>
<td>18.9</td>
<td>40.1</td>
<td>24.1</td>
<td>46.6</td>
<td>24.4</td>
<td>39.1</td>
<td>34.0</td>
</tr>
<tr>
<td>Water</td>
<td>37.2</td>
<td>54.8</td>
<td>45.7</td>
<td>48.5</td>
<td>43.8</td>
<td>46.8</td>
<td>53.1</td>
</tr>
<tr>
<td>Ramachandran Plot (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most Favored</td>
<td>93.3</td>
<td>89.0</td>
<td>90.1</td>
<td>83.6</td>
<td>88.9</td>
<td>88.7</td>
<td>90.3</td>
</tr>
<tr>
<td>Allowed</td>
<td>6.7</td>
<td>10.5</td>
<td>9.1</td>
<td>16.4</td>
<td>11.1</td>
<td>10.8</td>
<td>9.3</td>
</tr>
<tr>
<td>Generously Allowed</td>
<td>0</td>
<td>0.3</td>
<td>0.7</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Chapter 2: Crystallographic and Biochemical Analysis of the Ran-Binding Zinc Finger Domain

\[ R_{sym} = \frac{\sum |I_i| - \langle |I_i| \rangle}{\sum |I_i|}, \] where \( I_i \) is the intensity of the \( i \)th observation and 
\[ \langle |I_i| \rangle \] is the mean intensity of the reflection.

\[ R_{work} = \frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{calc}|}, \]

\[ R_{free} = R \text{ value for a randomly selected subset (5\%) of the data that were not used for minimization of the crystallographic residual.} \]

Table 2.2 - Thermodynamic parameters for ZnF-Ran binding in ITC experiments

<table>
<thead>
<tr>
<th>Titrand/Titrant</th>
<th>[Titrand] (( \mu \text{M} ))</th>
<th>[Titrant] (mM)</th>
<th>T (K)</th>
<th>( K_{d1} ) (( \mu \text{M} ))</th>
<th>( K_{d2} ) (( \mu \text{M} ))</th>
<th>( N )</th>
<th>( \Delta H_1 / \Delta H_2 ) (kJ/mol)</th>
<th>( T \Delta S_1 / T \Delta S_2 ) (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nup153-ZnF1234</td>
<td>12   RenGDP   0.31</td>
<td>293  1  2  8  2</td>
<td>-40.8/3.9</td>
<td>-7.38/32.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12   RenQ89L  0.40</td>
<td>293  20 2 - -</td>
<td>-39.3/-</td>
<td>-12.9/-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF12</td>
<td>15   RenGDP   0.29</td>
<td>293  4  1  57 1</td>
<td>-41.0/20.4</td>
<td>-10.3/44.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF34</td>
<td>23   RenGDP   0.35</td>
<td>293  3  1  59 1</td>
<td>-17.2/-2.2</td>
<td>14.0/21.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF1</td>
<td>13   RenGDP   0.30</td>
<td>293  6  1 - -</td>
<td>-7.8 21.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF1Q8D</td>
<td>20  RenGDP   0.36</td>
<td>293  8  1 - -</td>
<td>-11.2 17.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF2</td>
<td>32   RenQ89L  1.25</td>
<td>278  49 1 - -</td>
<td>16.8 39.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF2D8Q</td>
<td>18   RenQ89L  0.34</td>
<td>293  4  1 - -</td>
<td>-5.5 24.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF3</td>
<td>29   RenQ89L  0.30</td>
<td>293  7  1 - -</td>
<td>-3.1 26.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF3E8D</td>
<td>17   RenQ89L  0.34</td>
<td>293 - - - -</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF4</td>
<td>50   RenQ89L  0.95</td>
<td>278  47 - -</td>
<td>12.6 35.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup153-ZnF4D8Q</td>
<td>19   RenQ89L  0.36</td>
<td>293  3  1 - -</td>
<td>-6.6 24.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup358-ZnF12</td>
<td>25   RenGDP   1.25</td>
<td>278  15 2 - -</td>
<td>-3.4 22.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2.3 - Experimental constructs, abbreviations, and descriptions

<table>
<thead>
<tr>
<th>Experimental Constructs</th>
<th>Protein</th>
<th>Residues</th>
<th># ZnFs</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ran</td>
<td>RAN</td>
<td>1:216</td>
<td>0</td>
<td>Full-length WT Ran</td>
</tr>
<tr>
<td>RanQ69L</td>
<td>RAN</td>
<td>1:216</td>
<td>0</td>
<td>Full-length Ran with Gln69 mutated to Leu to lock Ran in a GTP nucleotide bound state</td>
</tr>
<tr>
<td>RanF35S</td>
<td>RAN</td>
<td>1:216</td>
<td>0</td>
<td>Full-length WT Ran with Phe35 mutated to Ser to increase the strength of the Ran-ZnF crystal lattice</td>
</tr>
<tr>
<td>Nup153-ZnF1234</td>
<td>Nup153</td>
<td>658:876</td>
<td>4</td>
<td>Entire zinc finger domain (ZnF) of Nup153. The WT linker between each ZnF remains intact.</td>
</tr>
<tr>
<td>ZnF12</td>
<td>Nup153</td>
<td>658:750</td>
<td>2</td>
<td>Zinc finger domain (ZnF) of Nup153 shortened to include only the first two ZnFs of the domain, ZnF1 and ZnF2. The WT linker between the two ZnFs remains.</td>
</tr>
<tr>
<td>ZnF34</td>
<td>Nup153</td>
<td>790:876</td>
<td>2</td>
<td>Zinc finger domain (ZnF) of Nup153 shortened to include only the first two ZnFs of the domain, ZnF3 and ZnF4. The WT linker between the two ZnFs remains.</td>
</tr>
<tr>
<td>ZnF1</td>
<td>Nup153</td>
<td>658:686</td>
<td>1</td>
<td>Zinc finger domain of Nup153 shortened to include only ZnF1 with no linker</td>
</tr>
<tr>
<td>ZnF2</td>
<td>Nup153</td>
<td>723:750</td>
<td>1</td>
<td>Zinc finger domain of Nup153 shortened to include only ZnF2 with no linker</td>
</tr>
<tr>
<td>ZnF3</td>
<td>Nup153</td>
<td>790:817</td>
<td>1</td>
<td>Zinc finger domain of Nup153 shortened to include only ZnF3 with no linker</td>
</tr>
<tr>
<td>ZnF4</td>
<td>Nup153</td>
<td>848:876</td>
<td>1</td>
<td>Zinc finger domain of Nup153 shortened to include only ZnF4 with no linker</td>
</tr>
<tr>
<td>Nup358-ZnF12</td>
<td>Nup358</td>
<td>1390:1489</td>
<td>2</td>
<td>Zinc finger domain of Nup358. Includes 2 ZnFs with WT linkers between the ZnFs.</td>
</tr>
</tbody>
</table>
Experimental contributions: James Partridge and Thomas Schwartz have conducted all experiments.
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

Introduction

The nuclear pore complex (NPC) is one of the largest protein assemblies inside the eukaryotic cell and is the sole conduit by which all nucleocytoplasmic transport proceeds. Cryo-electron microscopy/tomography studies have demonstrated that the NPC exists as a protein scaffold embedded in the nuclear membrane, exhibiting 8-fold rotational symmetry about a central channel through which cargo can pass between cytoplasm and nucleus (Hinshaw et al., 1992). Proteomic studies that account for the NPC's 8-fold rotational symmetry and 2-fold symmetry across the nuclear envelope demonstrate the entire NPC weighs ~44 MDa in *S. cerevisiae* and ~66 MDa in rat (Alber et al., 2007b; Cronshaw et al., 2002; Rout et al., 2000). Still the NPC is composed of only about 30 proteins called nucleoporins or Nups, which assemble into biochemically distinct subcomplexes. These subcomplexes exist within the NPC in accordance to the symmetry described above and represent stable building blocks from which the structural scaffold of the NPC is built. Each subcomplex is composed of between 2 and 10 Nups, and these subcomplexes are known to also exist in mitotic extracts following breakdown of the NPC in higher eukaryotes (Matsuoka et al., 1999). Due to the high degree of internal symmetry, multiple copies of each Nup are present throughout an intact NPC. The general architecture of the NPC is conserved between yeast and vertebrates and improved methods in cryo-electron tomography, field-emission in-lens scanning electron microscopy (FEISEM), 4pi confocal microscopy, and cell fixation techniques, serve as the best method for such sweeping observations (Beck et al., 2007; Lim et al., 2008a; Maco et al., 2006; Stoffler et al., 2003). Based on structural data and increasingly powerful methods of structural modeling, it is clear that only a small number of structural domains make up the NPC: α-helical stacked domains, β propellers, and coiled-coil domains (Berke et al., 2004; Cronshaw et al., 2002; Devos et al., 2004; Weirich et al., 2004). The fourth distinct domain, found to compose at least one-third of total NPC mass, is the phenylalanine-glycine-repeat domain. This flexible,
unstructured Phe-Gly rich domain harbors repeats of FG dipeptides separated by polar spacers of varying lengths. It is believed that highly disordered and hydrophilic FG-repeat regions represent the primary interaction site between nuclear transport receptors (NTRs) and the NPC, thereby playing a direct role in the transport of import- and export-complexes through the NPC (Ben-Efraim and Gerace, 2001). Only a small subset of Nups, including (in S. cerevisiae) Pom152, Pom34, and Ndc1 have transmembrane domains and are the Nups responsible for anchoring the NPC scaffold into the nuclear envelope (Hetzer and Wente, 2009).

Although the NPC is built around a rigid scaffold responsible for structural integrity, the NPC remains a dynamic macromolecular machine. Evidence is emerging to suggest that the cytoplasmic ring, the nuclear basket, and the luminal spoke ring of NPC undergo significant conformational rearrangements to adjust for translocation events (Beck et al., 2007). Aside from itself being a dynamic and flexible machine, some nucleoporins are dynamic and can shuttle on and off the NPC (Rabut et al., 2004). While some nucleoporin subcomplexes represent rigid scaffolds from which the NPC is framed, including the structurally well-characterized Y-complex (Brohawn et al., 2009; Kampmann and Blobel, 2009; Lutzmann et al., 2002), other nucleoporin subcomplexes remain something of an enigma, with less structural and biochemical data available to specifically define their functional role in the NPC.

As with many Nups, the Nup62 complex was originally identified in extracts of fractionated rat liver nuclei as a biochemically stable subcomplex composed primarily of three proteins (Davis and Blobel, 1986; Rout and Blobel, 1993; Snow et al., 1987). In mammals the homologous proteins are: Nup62, Nup54, and Nup58 respectively. In several instances this manuscript will compare the two trimeric complexes and we will invoke the yeast nomenclature for generalizations made between the two complexes. The size and sequence of Nsp1, Nup57, and Nup49, differ among different eukaryotes, however the overall structural characteristics remain concordant. Nsp1, Nup57, and Nup49 are symmetrically localized within the NPC as
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

judged by immunogold electron microscopy (Rout et al., 2000). Each member of the Nsp1 complex is an essential component for both function and structure of the NPC, as determined with deletions and point-mutants (Grandi et al., 1995; Hurt, 1988; Wente et al., 1992; (Strawn et al., 2004). In addition, *in vitro* experiments such as immunodepletion of the Nsp1 complex during reconstitution of the NPC with rat nuclei and *Xenopus* egg extracts, disrupts transport of NLS tagged cargo and even the formation of NPCs as revealed by scanning EM (Finlay et al., 1991; Mutvei et al., 1992). In addition to the three major components of the Nsp1 complex a fourth protein, Nic96 (yeast) / Nup93 (metazoa), is present as a partially stable component of the Nsp1 complex (Grandi et al., 1993; Grandi et al., 1995; Guan et al., 1995). It is believed that Nic96 serves as a bridging molecule, linking together the heterotrimeric Nsp1 complex with the heteropentameric Nic96 subcomplex.

As the first crystal structures of karyopherins and their complexes with: Ran, NLS-/NES-cargos, and the FG-repeats of Nups, propelled our knowledge of the mobile phase of nucleocytoplasmic transport, the answers to most questions regarding assembly, structure, and function of the NPC will ultimately depend on architectural determination of Nups and NPC subcomplexes at atomic resolution. As a cornerstone of NPC architecture and function, the trimeric Nsp1 complex is a primary target for structural studies. The Nsp1 complex occupies the central channel of the NPC and plays a key regulatory role in the transport of cargo across the nuclear membrane (Fahrenkrog et al., 1998). Domain architecture of each component from the Nsp1 complex maintains a conserved arrangement of secondary structure consisting of a coiled-coil domain, approximately 200 residues in length, flanked by fiber-like extensions of unstructured amino acid sequence containing Phe-Gly-rich (FG) domains (Hu et al., 1996; Schwarz-Herion et al., 2007). FG-repeat domains are found in a total of ~11 different Nups and remarkably account for almost 13% of NPC's total mass to fill the central channel of the NPC (Brohawn et al., 2009). It is well established that FG-repeats are the primary interaction sites
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

between the NPC and nuclear transport receptors (NTRs) that shuttle cargo during nucleocytoplasmic transport (Denning et al., 2003; Isgro and Schulten, 2007a, b; Macara, 2001; Peters, 2005). However it is disputed how FG-repeats interact to form a transport barrier that prohibits molecules greater then ~ 40 kDa from passing through the NPC by diffusion, unless bound to a karyopherin (KAP) (Frey and Gorlich, 2007; Lim et al., 2007; Peters, 2009; Rout et al., 2003). The coiled-coil domains from each member of the Nsp1 complex interact to form a trimeric assembly that lines the inner-channel of the NPC, while the FG-repeat regions are largely unstructured and as such are nonessential for formation of the Nsp1 complex (Bailer et al., 2001; Finlay and Forbes, 1990; Hu et al., 1996; Strawn et al., 2004).

Nic96 is one of the most abundant nucleoporins inside the eukaryotic cell, contributing to nearly 10% of the total NPC mass with an estimated 32-48 copies per intact NPC (Cronshaw et al., 2002; Rout et al., 2000). The Nic96 subcomplex is a key structural module of the NPC’s core scaffold and bridges together the Nsp1 complex with the heteroheptameric Y-complex. Although an essential component of the NPC scaffold, little biochemical data has been gathered to detail the molecular workings of the Nic96 subcomplex. Like members of the Nsp1 complex Nic96 is essential for both NPC assembly and function, as shown in a variety of eukaryotes (Allende et al., 1996; Galy et al., 2003; Grandi et al., 1993; Osmani et al., 2006). Unlike members of the Nsp1 complex Nic96 is not an FG-repeat Nup, but does have two structural domains common to other Nups, a 150-residue long coiled-coil domain at the N-terminus followed by a large α-helical domain roughly ~600 residues in length (PDB codes 2QX5, 2RFO) (Jeudy and Schwart, 2007; Schrader et al., 2008b). The helical domain of Nic96 contains 30 helices arranged in a J-like fashion with the N-terminus starting in the middle of the elongated helical stack. The C-alpha backbone of Nic96 traverses up one side of the molecule, makes a U-turn, and then continues past the N-terminus to the other end of the elongated molecule. Three other α-helical scaffold nucleoporins (Nup84, Nup85, and Nup145C) have since been structurally characterized
and shown to adopt the same unique fold, pointing to a common ancestor (Brohawn et al., 2008). These four scaffold nucleoporins share a common α-helical domain that to date, outside of the NPC, has only been identified in Sec31, an essential component of the COPII vesicle coat. The similarity between these five proteins was not expected, as sequence conservation is so low that no structural specific relationship was previously inferred (Alber et al., 2007a; Hsia et al., 2007). This unique structural domain has been termed, the Ancestral Coatomer Element (ACE1), provides structural evidence that the nuclear pore complex and vesicle coats, notably COPII, share a common origin (Devos et al., 2004).

The N-terminal coiled-coil domain of Nic96 mediates a direct interaction between Nic96 and the Nsp1 trimeric complex (Grandi et al., 1995). In addition to the strong interaction with Nsp1, Nic96 has been reported to bind with other members of the Nic96 complex, including: Nup53, Nup188, and Nup192. These interactions however appear to be significantly weaker then the interaction with Nsp1 as shown in multiple pull-down experiments (Alber et al., 2007a; Kosova et al., 1999; Nehrbass et al., 1996; Zabel et al., 1996). In addition to using the C-terminal domain for localization to the NPC it appears that the N-terminal coiled-coil domain of Nic96 is equally important for assembling the nuclear pore complex (Schrader et al., 2008b).

To understand the interaction between members of the Nsp1 trimeric complex and Nic96, we have reassembled the tetrameric complex made between the coil-coiled core of the Nsp1 trimeric complex and full-length Nic96. Using liquid chromatography, and biophysical characterization we have dissected the interactions responsible for assembly of this large coiled-coil complex. In addition to analysis of the yNsp1 complex from yeast, we have also reassembled the homologous complex from Rat, the rNup62 complex. Based on this analysis and previous data we can now see that not only is the N-terminal coiled-coil domain of Nic96 essential for interaction with the Nsp1 triple complex, but also is essential for incorporation of Nic96 into the nuclear pore complex.
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

Results

Reassembly of the trimeric yNsp1 / rNup62 subcomplex

A biophysical and crystallographic study of the yNsp1 complex requires not only the reassembly of the triple complex in vitro, but milligram amounts of recombinant protein for crystallization trials and biophysical characterization experiments. This is difficult because, although the coiled-coil domains of yNsp1, yNup57, and yNup49 can individually be expressed as N-terminal His-tag fusion proteins in *E. coli*, the proteins are largely insoluble. Under denaturing conditions, the proteins can be resolubilized with urea, purified via Ni-affinity chromatography, and mixed in stoichiometric amounts to reassemble the triple-complex. However, binding by the coiled-coil domains is not entirely specific, resulting in assemblies other than the desired triple-complex being formed. For instance, a hetero-dimeric complex between yNup57 and yNsp1 is a standard impurity in this assembly procedure. In this study the core of the yNsp1 complex has been reassembled in vitro using only the coiled-coil domains, expressed from a modified polycistronic vector. Polycistronic vectors have become a critical tool for expression studies of large complexes and are especially useful when components are insoluble if expressed alone (Selleck and Tan, 2008; Tan, 2001). For our experiments a bicistronic pET-Duet vector was modified to contain a third ribosome binding for coexpression of all three components with a 6xHis-tag fused to the N-terminus of yNup49 in the first cassette of the vector, followed by yNup57 in the second cassette, and yNsp1 in the final and third cassette (Fig. 3.1). Mapping the interaction between members of the yNsp1 complex was done in a systematic fashion. By removing one predicted helix at a time with N- and C-terminal deletion constructs of yNup49, and yNup57, yNsp1, we could identify a recombinantly expressed trimeric complex with optimal solubility (Fig. 3.1 and Fig 3.3). Placing the His-tag on yNup49 is done to purify away the hetero-dimeric yNup57•yNsp1 complex. Mutants lacking one or more helices demonstrate a range in stability and solubility, however one exemplary construct that includes
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

yNup49360-466, yNup57354-534, and yNsp1579-824 (Fig. 3.2) yields about 20 mg of trimeric complex purified from 4 L of liquid E. coli culture, sufficient quantities for crystallographic and biophysical experiments.

The purification procedure described above has the inherent flaw of purifying soluble, but unbound yNup49, in addition to the entire Nsp1 complex. Because the yNsp1 complex binds to other components of the NPC, and is composed entirely of coiled-coils, there may exist regions on the outer surface conducive to auxiliary protein-protein interactions. Left unoccupied, these typically hydrophobic patches can bind additional non-stoichiometric amounts of yNup49, resulting in the heterogeneous purification of the yNsp1 trimeric complex and yNup49. To identify hydrophobic patches that mediate coiled-coil contacts, the hydrophobicity of helices was analyzed by hydrophobic cluster analysis (HCA) using the program Drawhca and the mammalian rNup62 complex (Callebaut et al., 1997). After identifying hydrophobic patches, specific amino acids from each helical region were mutated to charged residues using site-directed mutagenesis. With this method we were able to significantly increase solubility of the rNup62 complex with mutations in the first and third helix of rNup62 (Fig. 3.2). Our HCA based mutagenesis experiments also uncovered four critical hydrophobic patches, necessary for stable complex formation. Two mutations of rNup58 completely abolish purification of trimeric complex. Mutations to the fifth helix of rNup54 abolish interaction between rNup54 and rNup62, as rNup54 and rNup62 no longer copurify in the presence of excess rNup58. Further mapping experiments described below have confirmed that the C-terminal helix of rNup54/yNup57 is responsible for binding directly with rNup62/yNsp1.

Not surprisingly, the trimeric complex from R. norvegicus (rNup62 complex) behaves quite differently from the trimeric complex reassembled using S. cerevisiae genes (the yNsp1 complex). The yNsp1 complex is more soluble then the rNup62 complex and does not require
solubility-based point mutations. For purification of soluble protein, using the full coiled-coil
domain from yNspl and yNup57 is tolerable, but where the rNup62 complex was optimized with
a longer version of rNup58 (Fig. 3.1), the yNspl complex will not tolerate using a yNup49
construct that includes all four helices of the coiled-coil domain. As the solubility of yNup49
drops so does the solubility of the trimeric complex. In addition to solubility differences, the
yNspl trimeric complex has two auto-proteolytic sites in yNup57. These sites are not present in
rNup54 (Fig. 3.1). The two proteolytic sites have been mapped to a loop between the predicted
fourth and fifth helix of the coiled-coil domain. Between amino acids 480 and 498 is an
unstructured loop, non-conserved among yNup57 orthologs. With N-terminal sequencing and
mass spectroscopy the primary proteolytic site of yNup57 was mapped to residue 482. A
proteolytic site no doubt leads to a mixed population of yNup57, with a low percentage of full-
length protein intact after several days at room temperature (Fig. 3.3). A second construct was
modified to remove the loop entirely, as well as remove C-terminal helix 5. In this process we
reconfirmed our HCA analysis on rNup54 and found out that yNup57-helix-5 is critical for
binding directly to yNspl. If removed, yNspl dissociates from the triple complex during gel
filtration and ion exchange purifications, leaving a core heterodimeric complex between yNup49
and yNup57. This allowed us to map the helical interactions between coiled-coil domains of the
Nsp1 trimeric complex (Fig. 3.3).

**Mapping interactions within the Nsp1 complex**

Based on a series of helical truncations at the N- and C-terminus of yNup57 we believe
the heterodimeric interaction between yNup49 and yNup57 constitutes a core scaffold of the
trimeric yNspl complex. The presence of consecutive heptad repeats in a peptide sequence,
denoted \[abcdefg\]_n, serves as a powerful tool for the prediction of coiled-coil domains that can
be evaluated with several different prediction programs (Lupas et al., 1991; McDonnell et al.,
Although the identification of coiled-coil domains can be fairly simple, designating interacting helices within large coiled-coil assemblies is not trivial. Predicting the structure of coiled-coil assemblies is made difficult due to the two orientations that each helix can make (coiled-coils can be parallel or anti-parallel), the assembly into coiled-coils composed of two, three, four or five helices, and the potential axial shifts in heptad interactions. There are however studies demonstrating that it is possible to discriminate between two-stranded versus three-stranded assemblies, providing insight into assembly order and plausible binding pairs (Apgar et al., 2008; Wolf et al., 1997; Woolfson and Alber, 1995). Using hydrophobic cluster analysis, coupled with the mutation of residues from 11 hydrophobic clusters to charged, discrete point mutations can have a drastic effect on solubility as well as assembly. As demonstrated with mutations 5 and 6 to rNup58 in one-step affinity tag purifications, disruption of either rNup58 helix results in complete disassembly of the trimeric complex (Fig. 3.2). Mutations 1, 8, and 11 are similar to wild-type and produce similar levels of soluble trimeric complex (Fig. 3.2). We rationalize these hydrophobic patches likely have little effect on trimeric complex assembly. Mutations 2, 7, 9, and 10, appear to increase solubility of the trimeric complex, while maintaining stable complex assembly. As a result of this analysis, mutations 7 and 9 were specifically used to increase solubility of the rNup62 trimeric complex for all further biophysical experiments described herein. Interestingly, constructs containing mutation 3 or 4 in the fifth helix of Nup54 have the most aberrant trimeric assembly. Only in these constructs is the solubility of Nup58 much higher than wild-type, however there is far less rNup54 and rNup62 present after the one-step Ni-affinity purification. Because all proteins are stably expressed, the purification of excess rNup58 likely indicates that these helices play a critical role in the full trimeric assembly. This data is in agreement with additional mapping studies done using rNup54. As described above, the scissile site between helices four and five of rNup54 led us to truncation at the C-terminus of rNup54. Upon removal of the fifth and most C-terminal helix of
yNup57 at residue 512, purification of two distinct species is evident during gel filtration experiments (Fig. 3.3). It is clear that after removing the fifth helix, yNsp1 dissociates from the trimeric complex. This results in the purification of two distinct species, full trimeric assembly and a smaller heterodimeric "core" assembly between yNup49 and yNup57. Further truncations at the C-terminus of yNup57 to residue 474, in addition to removal of the first two helices at the N-terminus, results in a stable dimeric "core" interaction between yNup57 and yNup49. This "core" interaction represents the smallest interaction possible between members of the Nsp1 trimeric complex, as further truncations disrupt dimer assembly. This core heterodimeric assembly between yNup49 and yNup57 has been used in further biophysical experiments, alongside the larger trimeric, and tetrameric assembly with full-length Nic96.

**Biophysical characterization of the Nsp1 complex**

Biophysical characterization of the purified Nsp1/Nup62 trimeric complex and smaller components is critical for understanding the structure and biochemical nature of such a large coiled-coil protein assembly. This analysis gives us information regarding size, stoichiometry, and homogeneity of all purified assemblies. Size exclusion chromatography (SEC) provides a rough approximation of size and stoichiometry for macromolecular protein complexes. Protein from the yNsp1/rNup62 complex elute as a single stable peak at a volume roughly equivalent to twice the calculated mass, as compared with globular protein standards. However, the shape of macromolecules can dramatically affect the elution volumes obtained with SEC. In the purification of rod shaped assemblies it is difficult to distinguish size differences with SEC because the stokes radius can remain constant, even upon addition of more components. Based on SEC alone, it is unclear if the yNsp1/rNup62 complex exists as a heterotrimer or a heterohexamer. In the purification of a tetrameric complex including Nic96 it is possible that the extended Nic96 molecule with the N-terminal coiled-coil domain exhibits the same stokes radius
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

as the yNsp1•yNic96 tetrameric complex. Because of these limitations a more appropriate and advanced method for measuring mass, stoichiometry, and shape is necessary.

Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) allows for the measurement of molecular weight of a soluble protein independent of shape. Two types of experiments can be performed: sedimentation velocity and sedimentation equilibrium. Data collected using sedimentation velocity is influenced by both size and shape of the protein being studied. However, AUC data obtained using sedimentation equilibrium is unaffected by shape of the molecule and is sensitive only to the mass of the protein in native buffer conditions. Used in combination, these methods can provide precise measurements of molecular mass under native conditions, and even provide information regarding stoichiometry of the macromolecule based on fluctuations of thermodynamic parameters for self-associating systems (Hansen et al., 1994; Lebowitz et al., 2002).

The only known partial structure of any yNsp1/rNup62 complex member is an anti-parallel hairpin, homo-dimeric assembly of rNup58_{327-411} (PDB code 2OSZ) (Melcák et al., 2007). The crystal structure rNup58_{327-411} forms a dimer of dimers, with the core interaction being a tight dimer between two α-helices connected by a short loop (Melcák et al., 2007). The crystal structure suggests that the core homodimer of rNup58 is a stable component of the entire rNup62 complex. Numerous van der Waals interactions exist between the two helices with a total buried surface area of 1,262 Å². With a contact area of this size and complementarity it seems likely that the rNup58 dimer is representative of the rNup58 state found within the full rNup62 trimeric assembly. To test this observation and determine the stoichiometry of components within the yNsp1/rNup62 trimeric assembly, several variants of the assembly were tested using an analytical ultracentrifuge. With sedimentation velocity experiments we determined that the rNup62/yNsp1 complex exists in a 1:1:1 ratio between
components in both rat and yeast versions with only one molecule per assembly (Table 3.1). In addition to the trimeric assembly, dimeric complexes between rNup54•rNup62 and yNup49•yNup57 were tested and shown to exist in a 1:1 ratio with only one molecule each per assembly. However, in agreement with the crystal structure, a preparation of the rNup58 orthologue yNup49 was shown to self-associate into a higher order species in sedimentation velocity experiments. This higher-order species appears to be a homotetramer. Under physiological conditions in the cell, our results suggest that the yNsp1/rNup62 complex exists as a stable assembly with only one copy of each component, with 1:1:1 stoichiometry. Self-association of yNup49/rNup58, likely a non-physiological interaction, is only observed when the two other binding partners are limiting and can self-associate to form a unique higher order assembly in absence of the trimeric complex partners.

**Assembly of a tetrameric assembly between the Nsp1 complex and Nic96**

Previous studies have demonstrated that the Nsp1 complex interacts directly with the Nic96 complex via the coiled-coil domain of Nsp1 and the N-terminal coiled-coil domain of Nic96 (Grandi et al., 1993; Grandi et al., 1995; Guan et al., 1995). Although more recent work demonstrates this interaction with careful pulldown experiments (Alber et al., 2007a), complete reassembly *in vitro* has not previously been described. The N-terminal domain of Nic96 is a predicted coiled-coil domain that spans approximately 160 residues prior to a 30 residue long loop, followed by the ACE1 domain, spanning residue 200 to the C-terminus. Interaction between Nic96 and Nsp1 does not depend on the C-terminal ACE1 domain (Grandi et al., 1995). The N-terminal coiled-coil domain is insoluble unless expressed with the highly charged and soluble C-terminal fragment of yNic96. For this reason, reconstitution of a tetrameric yNsp1•yNic96 complex was only achieved in the presence of full-length yNic92 with the yNsp1 trimeric complex described above, yNup49_380-466, yNup57_354-534, and yNsp1_579-824. For reassembly
of this large complex a two-vector system was used, with the Nsp1 complex expressed from a pET-Duet-1 (ampicillin resistance) based tricistronic vector and full-length yNic96 expressed from pET-28a (kanamycin resistance). Both yNic96 and yNsp1 are fused to a 6x His tag at the N-terminus. Following copurification with Ni-NTA resin the tetrameric complex is purified away from smaller assemblies and other contaminants by ion-exchange chromatography, followed by size exclusion chromatography (SEC). With SEC, the yNsp1/rNup62 trimeric complex elutes at a volume equivalent to twice the calculated mass and at roughly the same position as the tetrameric yNsp1•yNic96 complex (Fig. 3.5). From our analytical ultracentrifugation data we calculate a high frictional ratio (f/f₀) above 1.5, highly indicative of a rod-like conformation for each protein assembly (Table 3.1). yNup49 purified alone is the only sample to have a f/f₀ below 1.5, which suggests that the protein adopts shorter, more spherical assemblies in the absence of yNup57 and yNsp1. The yNsp1/rNup62 trimeric complex exists as a hetero-trimer in a 1:1:1 ratio, with one copy of each component per assembly, and an early elution volume with SEC is due to the large stokes radius of the elongated coiled-coil assembly. Surprisingly, binding extended rod-shaped Nic96 (Jeudy and Schwartz, 2007; Schrader et al., 2008b), to the Nsp1 trimeric assembly does not affect the elution volume obtained via SEC (Fig. 3.5), indicating that the two elongated structures may collapse into a more compact structure or maintain an equivalent stokes radius.

Discussion

Reconstitution of a tetrameric protein assembly between the Nsp1 trimeric complex and FL-Nic96 represents the first reported assembly between proteins of the Nsp1 and the Nic96 nuclear pore subcomplexes. The stable trimeric yNsp1/rNup62 complex has been reconstituted from S. cerevisiae and R. norvegicus respectively and exhibits the shape of a long rod, as definitively shown by analytical ultracentrifugation and size exclusion chromatography. Data from analytical ultracentrifugation also demonstrates that the trimeric assembly exhibits a stable
1:1:1 stoichiometry between the three coiled-coil domains of \( y\text{Nspl}, y\text{Nup57}, \) and \( y\text{Nup49} \). Mapping experiments reveal a defined interaction between \( y\text{Nup57} \) and \( y\text{Nspl} \) mediated by helices 1, 2, and 5 of \( y\text{Nup57} \) and the coiled-coil domain of \( y\text{Nspl} \). In addition to the trimeric assembly between \( y\text{Nup49}, y\text{Nup57}, \) and \( y\text{Nspl} \), a stable dimeric "core" has been isolated between \( y\text{Nup49} \) and \( y\text{Nup57} \). The \( y\text{Nup49} + y\text{Nup57} \) core complex also exists in a stable stoichiometric 1:1 ratio, as seen with SEC and analytical ultracentrifugation. Although this study provides considerable insight into the assembly and interaction between the \( y\text{Nspl} \) and \( y\text{Nic96} \) subcomplexes, further experiments will probe the molecular details of such a large structural assembly, namely by crystallographic analysis.

In this study we have focused on the coiled-coil domains that mediate contact between \( y\text{Nic96} \) and the \( y\text{Nspl} \) trimeric complex, as \( y\text{Nic96} \) tethers the dynamic \( y\text{Nspl} \) complex to the structural scaffold of the NPC. The structural scaffold is made up of the Y-complex, composed of seven proteins that serve as a scaffold lattice for the more dynamic Nups that reside on either face of the NPC and that line the inner channel. The \( y\text{Nic96} \) complex probably contains at least five stable components and most likely has a similarly important architectural role as the Y-complex. The ACE1 domain of \( y\text{Nic96} \) is a stable rod-like domain, and other proteins from the \( y\text{Nic96} \) complex, such as \( y\text{Nup188} \) and \( y\text{Nup192} \), are also large helical units. \( y\text{Nic96} \) is related to three of the five helical proteins that build the Y-complex. \( y\text{Nup157}/y\text{Nup170} \) of the \( y\text{Nic96} \) complex is structurally related to \( y\text{Nup133} \) of the Y-complex (Whittle and Schwartz, 2009). The next stage of NPC structural studies will likely focus on the \( y\text{Nic96} \) complex, such that the lattice scaffold can be completely understood. The anchor points for the \( y\text{Nspl} \) complex will then become more apparent and will likely bring us closer to a comprehensive functional characterization.

**Methods**
**Protein purification**

A bacterial expression construct containing rNup58<sub>328-415</sub>, rNup54<sub>348-495</sub>, and rNup62<sub>322-523</sub> from *Rattus norvegicus* was constructed for reassembly of the rNup62 subcomplex. The genes were cloned in a pET28a vector (EMD Biosciences), modified to be a tricistronic vector containing three independent ribosome binding sites for each ORF and expressed as a 6xHis-tag fusion protein engineered to contain a protease 3C site after the N-terminal affinity tag. The rNup62 subcomplex construct containing an extended rNup58<sub>239-415</sub> was similarly cloned using the original rNup62 subcomplex vector as a template. The yNsp1 subcomplex, yNup49<sub>360-466</sub>, yNup57<sub>345-534</sub>, and yNsp1<sub>579-824</sub>, was cloned into pET-Duet 1, modified to include a protease 3C site after the N-terminal 6xHis-tag fused to yNup49 and to contain a third RBS site for expression of a third ORF. Full-length Nic96 was cloned as a 6xHis-tag fusion protein in pET28a, engineered to include a 3C site after the N-terminal affinity tag. All proteins were expressed in *Escherichia coli* strain BL21(DE3) RIL (Stratagene). Cells were grown at 30°C in Luria-Bertani broth supplemented with 0.4% glucose to OD<sub>600</sub> = 0.8 and induced with 0.2 mM IPTG at 18°C for 18 hours. Bacterial cell pellets harboring His-tagged proteins were suspended in 50 mM potassium phosphate pH 8.0, 500 mM NaCl, 20 mM imidazole, 3 mM β-mercaptoethanol (β-ME), and lysed using a french press. Crude lysate was supplemented with 200 μM phenylmethanesulfonyl fluoride and clarified by centrifugation at 9,000 g for 20 minutes. The soluble fraction was then incubated with 1 ml Ni-NTA per 1000 ODs for 1 hour at 277 K and loaded onto a disposable column (Pierce). The column was washed with 4 bed volumes of 50 mM potassium phosphate pH 8.0, 500 mM NaCl, 20 mM imidazole, 3 mM β-ME, and eluted with 6 bed volumes of 10 mM Tris/HCl pH 8.0, 50 mM NaCl, 150 mM imidazole, 3 mM β-ME. Eluted protein was dialyzed against 10 mM Tris/HCl pH 8.0, 250 mM NaCl, 0.1 mM EDTA, 1 mM DTT, for 1 hour before the 6xHis-tag was cleaved with 3C protease. The rNup62 trimeric assembly was purified using anion exchange chromatography on a HiTrapQ column (GE Healthcare) via a
linear NaCl gradient, followed by purification with size exclusion chromatography using a Superdex S200 26/60 column (GE Healthcare) in 10 mM Tris/HCl pH 8.0, 200 mM NaCl 1.0 mM DTT, 0.1 mM EDTA. The Nsp1 trimeric complex alone does not bind tightly to a Q column and as such was not purified with a Q column. The tetrameric assembly between full-length Nic96 and the Nsp1 complex was assembled by coexpression on the pET28a and pET-Duet-1 vectors respectively in Escherichia coli strain BL21(DE3) RIL (Stratagene), and purified with Ni-NTA resin as described above. The tetrameric complex between FL-yNic96 and the yNspl trimeric complex does bind to a Q column and elutes at 200 mM NaCl. The tetrameric assembly was purified on a Q column prior to further purification using a Superdex S200 26/60 column (GE Healthcare) equilibrated in 10mM Tris/HCl pH 8.0, 250mM NaCl, 0.1 mM EDTA, 1 mM DTT.

**Hydrophobic cluster analysis**

Hydrophobic cluster analysis (HCA) was conducted using Mobyle@RPBS (Callebaut et al., 1997; Gaboriaud et al., 1987). Hydrophobic patches were identified with HCA and two residues within each patch were mutated using an updated QuickChange Site-Directed Mutagenesis System (Stratagene) protocol (Zheng et al., 2004). Hydrophobic residues were mutated to large hydrophilic residues and the resultant construct was tested for trimeric complex assembly and solubility with the one-step Ni-NTA purification protocol described above.

**Mapping helical interactions**

Using N-terminal sequencing via Edman degradation and mass spectroscopy we were able to identify the two degradation products described in Fig. 3.1 as yNup57, truncated in a loop just prior and within helix 5. In an attempt to trim yNup57 past the proteolytic site we identified several key interacting domains between yNup57 and yNspl. By utilizing PCR techniques to truncate helices at the N- and C-terminus of yNup57 we were able to map
interactions among members of the yNsp1 complex. To construct a yNup49-yNup57 dimeric complex, helices 1, 2, and 5, were removed from yNup57, using the yNsp1 construct described above as template. Although removal of these helices does abolish interaction between yNup57 and yNsp1, yNsp1 was removed from the tricistronic vector to increase overall purity in purification of a dimeric yNup49-yNup57 complex.

**Analytical ultracentrifugation**

Several purified rNup62 and yNsp1 complex assemblies were gel-filtered in 10 mM Tris/HCl pH 8.0, 200 mM NaCl, 0.5 mM TCEP, and 0.1 mM EDTA immediately prior to the velocity sedimentation experiments. Analytical ultracentrifugation experiments were carried out with an Optima XL-A centrifuge using An50Ti (6 hole, equilibrium runs) or An60Ti (4 hole, velocity runs) rotors. Samples for sedimentation velocity (440 µL sample or 450 µL buffer) were loaded into Epon-charcoal filled 2 channel centerpieces, fit with sapphire windows, and spun at 42,000 rpm. Sedimentation velocity data was analyzed globally to generate a c(s) distribution using the continuous distribution model in SEDFIT (Schuck, 2000). The data was fit from 0.5 to 10 s-13 with an estimated vbar = 0.7300 cm3/g, η = 1.0182 cP, and ρ = 1.002 g/cm3.
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

Figures – Chapter 3

**Nsp1 Trimeric Complex**

- **Nsp1**
  - Domain architecture in cartoon form with largely unstructured FG-repeat domain in grey and the predicted coiled-coil domain in blue.
  - Orange boxes above each coiled-coil domain represent the minimal sequence of each component necessary to assemble the trimeric complex shown to right.
- **Nup57**
- **Nup49**

**Nup62 Trimeric Complex**

- **Nup62**
  - Domain architecture in cartoon form with largely unstructured FG-repeat domain in grey and the predicted coiled-coil domain in blue.
  - Orange boxes above each coiled-coil domain represent the minimal sequence of each component necessary to assemble the trimeric complex shown to right.
- **Nup54**
- **Nup58**

**Figure 3.1 - Domain architecture of both the Nsp1 and Nup62 trimeric complex**

The domain architecture of the Nup62 trimeric complex from rat and the homologous Nsp1 trimeric complex from yeast are represented in cartoon form with the largely unstructured FG-repeat domain in grey and the predicted coiled-coil domain of each protein in blue. The mammalian trimeric Nup62 complex is composed of Nup62, Nup54, and Nup58, the yeast orthologues are Nsp1, Nup57, and Nup49 respectively. Orange boxes above each coiled-coil domain represent the minimal sequence of each component necessary to assemble the trimeric complex shown to right. The coomassie-stained-SDS polyacrylamide gels shown to the right demonstrate the results of electrophoresis of each purified trimeric assembly. Each component is labeled, and the two stars below Nup57 in the Nsp1 trimeric assembly represent two degradation products that result from proteolytic cleavage at the N-terminus of helix 5 of Nup57. Two versions of the Nup62 assembly have been used in this study, with a short and long form of Nup58. This difference is noted with an orange and yellow box above the Nup58 coiled-coil domain, and is shown with the SDS gels at the right.
Figure 3.2 - Hydrophobic cluster analysis of the Nup62 trimeric assembly
As typically seen with coiled-coil assemblies, hydrophobic interactions form the key interactions for the
large coiled-coil Nup62 trimeric assembly. Hydrophobic cluster analysis (HCA) was used to identify hydrophobic patches that serve as potential interaction sites between helices of Nup62 trimeric assembly. In the diagrams shown above hydrophobic patches are outlined and hydrophobic residues are shown in green. In addition to abolishing interactions among helices, HCA was used to identify hydrophobic patches that could potentially impair solubility of the trimeric assembly. In both instances two hydrophobic residues, typically isoleucine/leucine and denoted by an arrow and number inside of the identified hydrophobic patches, were mutated to aspartic acid residues except for mutation 5 which was mutated to arginine. (b) The wild-type trimeric complex, as well as all mutations listed above, are visualized with this coomassie-stained-SDS polyacrylamide gel following one-step affinity tag purifications. The mutations are listed by number and have drastically different effects on the overall solubility and stability of the trimeric complex. Notably, mutations 3, 4, and 10 appear to disrupt the trimeric assembly, as evident in the increase in Nup58 purification but less than stoichiometric amounts of Nup62, and Nup54. Mutations 2, 7, and 9 appear to increase the overall solubility of the trimeric complex while maintaining stability among all members. Both mutations to Nup58, mutations 5 and 6, completely disrupt interaction within the known hairpin structure of Nup58 and leads to insoluble Nup58 and complete lack of a Nup62 trimeric assembly.
Figure 3.3 - Systematic deletion of helices from Nup57 map an interaction between Nup57 and Nsp1
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

The helical domains of Nup57 were systematically removed at both the N- and C-terminus to identify a minimal binding domain between Nup49 and Nup57, as well as identify the helical interactions between Nup57 and Nsp1. (a) A trace of the purification peak of the Nsp1 trimeric complex demonstrates that Nsp1, Nup57, and Nup49 are purified in stoichiometric amounts. (b) By removing the 5th helix of Nup57 a critical interaction between Nup57 and Nsp1 is lost. This is made evident in the gel filtration trace on the left where two obvious species elute at different volumes. To the right is the coomassie-stained-SDS polyacrylamide gel showing that the two species separated are the Nsp1 trimeric assembly, as well as a small assembly between Nup49 and Nup57. (c) In addition to the 5th helix, the 1st and 2nd helices also play a critical role in binding Nup57 with Nsp1. Upon removal of these three helices, interaction between Nup57 and Nsp1 is abolished leaving the core dimeric interaction between Nup49 and Nup57.

Figure 3. 4 - Analytical ultracentrifugation of Nsp1 / Nup62 assemblies
Analytical ultracentrifugation is a valuable tool for assessing the stoichiometry of the yNsp1/rNup62
assemblies presented in this manuscript. It had been previously reported that rNup58 exists as a large homo-tetramer assembly suggesting that assembly of the yNsp1/rNup62 complex would proceed with multiple copies of each component per full assembly (Melcák et al., 2007). This does not seem to be the case as based on our results recorded in Table 3.1. For every assembly only one copy of each molecule is present, except for our study of yNup49 alone. yNup49 is the yeast orthologue of rNup58, and in our study yNup49 exists as a tetramer in solution. Here the data and fits, in addition to residuals in the form of a bitmap are shown for all of our analytical ultracentrifugation experiments. The peak for a calculated sedimentation coefficient value is shown to demonstrate homogeneity of each purified assembly.
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

Figure 3.5 - Purification of a tetrameric assembly between full-length Nic96 and the trimeric Nsp1 complex
The purification of a tetrameric assembly is possible by co-expressing both full-length Nic96 and the Nsp1 trimeric assembly in E. coli. (a) Analytical gel filtration experiments demonstrate that full-length Nic96 and the Nsp1 coiled-coil trimeric assembly elute at roughly the same volume. When co-expressed and analyzed with gel filtration, the elution volume of the tetrameric assembly is not significantly different from the elution volume of full-length Nic96 alone. (b) Although not significantly observed with analytical gel filtration, analytical studies with anion exchange chromatography show a dramatic shift in elution conditions on a Hi-Trap Q column between full-length Nic96, the Nsp1 trimeric complex, and the
Chapter 3: Biophysical Analysis of the Tetrameric Assembly Between Nic96 and the Trimeric Nsp1/Nup57/Nup49 Complex

combined tetrameric assembly. The tetrameric assembly elutes as stable peak at 200 mM NaCl, while the Nsp1 trimeric complex does not bind tightly to a Q column and elutes below 100 mM NaCl. (c) Following purification by anion exchange chromatography, the tetrameric assembly between full-length Nic96 and the Nsp1 trimeric complex elutes as a stable peak as shown with gel filtration and coomassie-stained SDS-polyacrylamide gel analysis.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Expected MW</th>
<th>Calculated MW</th>
<th>RMSD</th>
<th>f/f0</th>
<th>S-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nup62 Triple</td>
<td>61.92</td>
<td>60.9</td>
<td>0.0059</td>
<td>1.68</td>
<td>3.23</td>
</tr>
<tr>
<td>Extended Nup58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nup62 Triple</td>
<td>51.92</td>
<td>58.3</td>
<td>0.0092</td>
<td>1.56</td>
<td>3.45</td>
</tr>
<tr>
<td>Dimeric Nup54 - Nup62</td>
<td>41.17</td>
<td>43.9</td>
<td>0.0045</td>
<td>1.70</td>
<td>2.64</td>
</tr>
<tr>
<td>Nup49</td>
<td>12.47</td>
<td>47</td>
<td>0.0145</td>
<td>1.41</td>
<td>3.33</td>
</tr>
<tr>
<td>Nsp1 Triple</td>
<td>60.63</td>
<td>57.94</td>
<td>0.0051</td>
<td>1.70</td>
<td>3.16</td>
</tr>
<tr>
<td>Dimeric Nup49- Nup57</td>
<td>34.56</td>
<td>34.5</td>
<td>0.0234</td>
<td>1.60</td>
<td>2.38</td>
</tr>
</tbody>
</table>
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

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


*equal contribution

Experimental contributions: James R. Partridge and Renuka Pandya conducted all experiments.
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

Introduction

Targeted modification of proteins with ubiquitin satisfies several essential functions inside the eukaryotic cell. Many processes such as, cell-cycle regulation, receptor trafficking, signal transduction, epigenetic regulation of gene expression, DNA repair, regulated proteolysis, and proteasome targeting are all under partial control by ubiquitin (Ub) labeling (Bernassola et al., 2008; Glickman and Ciechanover, 2002). The ability of ubiquitin to function in such a diverse array of biological functions is explained by the existence of distinct ubiquitin labeling signals recognized by an equally diverse set of ubiquitin-binding domains found in several regulatory proteins (Hicke et al., 2005; Hurley et al., 2006). Ubiquitin is linked to target proteins in three steps via a series of enzymes that act in a defined sequence. The Ub cascade begins with an activating E1 enzyme that requires ATP to generate a high-energy thioester intermediate. An E2 ubiquitin-conjugating enzyme then transfers the activated Ub molecule, via a high-energy thioester intermediate, to a substrate that has previously been bound to a member of the ubiquitin ligase family of enzymes, or E3 enzyme (Christensen et al., 2007; Pickart, 2001). E3 enzymes serve as the primary determinant for substrate recognition and come from one of two families. The RING (really interesting new gene) family of E3 enzymes function primarily as a rigid scaffold to facilitate interaction between the E2-Ub complex and a substrate targeted for ubiquitination (Zheng et al., 2002; Zheng et al., 2000). The second class of E3 enzymes is the HECT (homologous to E6AP C-terminus) domain family, all of which possess a highly conserved ~ 350 residue long domain at the C-terminus, responsible for catalyzing transfer of Ub from E2 to substrate. In contrast to the RING E3s, the HECT domain forms a direct thioester intermediate with Ub prior to active transfer of Ub to substrate in a highly mechanistic fashion. Each HECT domain contains an absolutely conserved cysteine near the C-
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

terminus that is responsible for formation of the Ub-E3 thioester intermediate

(Huibregtse et al., 1995; Scheffner et al., 1995).

**Mechanism of ubiquitination mediated by E3 HECT domains**

Approximately 600 E3s exist in the human proteome and 28 of those belong to the HECT domain family (Li et al., 2008). HECT domains generally share a conserved structural architecture consisting of a stable N-lobe and a mobile C-lobe responsible for transferring charged Ub from E2 to substrate. The need for a mobile and flexible enzyme is made apparent in recent structural and mechanistic studies. Early crystal structures between the E3 HECT domain E6AP and an E2, UbcH7, mapped the distance between the E3 catalytic cysteine, present on the C-lobe, and the E2 binding site, from the N-lobe, at 41 Å (PDB accession code 1D5F and 1C4Z) (Huang et al., 1999). Further crystallographic studies on the E3 HECT domain WWP1 confirmed movement of the C-lobe, and that a flexible hinge between the N- and C- lobes of the HECT domain allows for movement of the C-lobe to transfer Ub from E2 to substrate (PDB accession code 1ND7) (Verdecia et al., 2003). Modeling an E2 enzyme docked at the E2 binding site of the WWP1 structure demonstrates that the gap between the E3 catalytic cysteine and E2 is only 16 Å. Despite this detailed structural knowledge of E2 and E3 HECT domains, little is known regarding the molecular details of Ub exchange between E2-E3-substrate.

In a study of two different E3 HECT domains it was shown that E3 enzymes could adopt distinct methods for ubiquitinating substrate. E6A6 uses a mechanism termed indexation, where a full poly-Ub chain is built on the E6AP catalytic cysteine prior to substrate transfer. Another E3 HECT domain, KIAA10, utilizes a sequential addition method where one Ub molecule at a time is transferred to substrate to form a poly-Ub chain (Wang and Pickart, 2005). These differences highlight the characteristically
modular nature of the E3 HECT domain, an enzyme that requires high-level flexibility while maintaining a level of specificity critical for governing ubiquitination of target substrate.

**Structural characteristics of the E3 HECT domain**

The E3 HECT domain is particularly flexible for mechanistic reasons. It is likely that this flexibility enables the E3-ligase HECT domain to adopt variable mechanisms for poly-Ub labeling of substrate. Based on crystallographic data from previously solved structures we know that the HECT domain contains two major lobes referred to as the N- and C-lobe (Fig. 3.2). The large N-lobe contains the E2 binding region, while the C-lobe contains the catalytic cysteine responsible for binding Ub during transfer of Ub from E2 to target substrate. While the crystallographic structure of the E6AP HECT domain bound to Ub demonstrates a gap of 41 Å between E2 active site and the catalytic cysteine, the WWP1 structure shows that the C-lobe rotates ~100° to close this gap. Further experiments to inhibit rotation at the unstructured hinge between the N- and C-lobes leaves little doubt that a flexible hinge is an functional feature of HECT domains that enables rotation of the C-lobe to facilitate substrate labeling (Verdecia et al., 2003).

Some HECT domains require adaptor proteins to catalyze E3 activity. In a study that describes the structure of another E3 HECT domain Smurf2 (PDB accession code 1ZVD), the authors describe a phenomenon by which catalytic activity of Smurf2 is greatly enhanced by the addition of Smad7, a protein that belongs to the TGFβ family of proteins involved in multiple cell signaling pathways as a regulatory transcription factor (Varelas et al., 2008). Addition of the N-terminal domain (NTD) domain of Smad7 is shown to link the E2 (Ubch7) and Smurf2, as Smad7 can bind individually to either component alone. Additionally, Smad7 appears to have a regulatory role. E2 enzymes
such as UbcH6 and Ubc10 show decreased activity in the presence of Smad7. Although not entirely necessary for the entire family of E3 HECT domains, it appears that some E3s require a secondary regulatory protein to facilitate cooperation between E2 and E3 to increase specificity of substrate labeling. A recent paper from Brenda Schulman's group describes a trimeric crystallographic structure between E2, E3, and ubiquitin (Kamadurai et al., 2009). In this structure (PDB codes 3JVZ and 3JWO), an E3 HECT domain NEDD4L and E2 enzyme, UbcH5B, are docked at the HECT domain E2 binding site, with a single ubiquitin molecule covalently bound to UbcH5B. To suppress conjugation between E3 and ubiquitin, the catalytic cysteine of NEDD4L was modified to either alanine or serine. In this stunning crystallographic complex, the ubiquitin molecule is positioned directly between E2 and E3 enzymes, primed for transfer to E3.

Conserved hydrophobic patches on the E3 C-lobe appear to stabilize a non-covalent interaction with ubiquitin, effectively positioning ubiquitin against E3 so that the thioester transfer reaction can take place between catalytic cysteines of E2 and E3 enzymes. Overall the structure of NEDD4L in the NEDD4L-UbcH5B-ubiquitin complex is similar to the HUWE1 HECT domain with fluctuations to the N-lobe at the E2 binding site, significant for E2 specificity. The authors further elaborate that modification to substrate specificity can likely be made based on changes to the HECT domain C-lobe, as was seen by swapping the entire C-lobe for another from a second HECT domain (Kim and Huibregtse, 2009). This also stems from the observation that modifying hydrophobic residues on the C-lobe, distant from the both the catalytic cysteine and non-covalent ubiquitin-interaction site modulate substrate-labeling activity (Salvat et al., 2004). These observations make it clear that the E3 HECT domain has developed a variety of mechanistic methods to regulate E2 and substrate specificity for accurate and processive regulation of the ubiquitin pathway.
An important question regarding HECT domain function is the control of ligase activity and specificity. Here we present a functional analysis of the HECT domain of the E3 ligase HUWE1, based on crystal structures, and show that a single, N-terminal helix significantly stabilizes the HECT domain. We observe that this element modulates HECT domain activity, as measured by self-ubiquitination induced in the absence of this helix, as distinct from its effects on Ub conjugation of substrate Mcl-1. Such subtle changes to the protein may be at the heart of the vast spectrum of substrate specificities displayed by HECT domain E3 ligases.

HUWE1 (also called ARF-BP1, Mule, Lasu1, Ureb1, E3 histone, and HectH9) is a large 482-kDa HECT domain E3 Ub ligase implicated in the regulation of cell proliferation, apoptosis, and DNA damage response (Adhikary et al., 2005; Chen et al., 2005; Hall et al., 2007; Herold et al., 2008; Zhao et al., 2008; Zhong et al., 2005). We recovered this enzyme with immunoprecipitation using Ub C-terminal electrophilic probes (Love et al., 2009). Following an initial biochemical characterization (Love et al., 2009), we completed a structural and biophysical analysis of the HECT domain to understand modulation of its robust in vitro activity. Here we present crystal structures of the HUWE1 HECT domain and characterize a structural element that both stabilizes this domain and modulates its activity. This structural element, the α1 helix, is an important component of the HECT domain that largely restricts its autoubiquitination activity, while only nominally affecting Mcl-1 ubiquitination activity.

Results

Structure of the HUWE1 HECT domain

We first attempted to crystallize the HUWE1 HECT domain using a fragment defined by the founding member of the HECT domain E3 ligase family, E6AP (PDB
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

codes 1C4Z and ID5F) (Huang et al., 1999). The crystals diffracted to only 3.5 Å (data not shown), with fairly high temperature factors indicating vibrational disorder within the protein crystals. By sequence comparison, we noted the significance of a conserved N-terminal helix that seals the hydrophobic core of the N-lobe in the structures of WWP1, Smurf2, and Nedd4-like (residues 546-560 in WWP1 and 371-387 in Smurf2) (Ogunjimi et al., 2005; Verdecia et al., 2003). The presence of this helix is conserved in over 13 HECT domain E3 ligases based on sequence comparison (Fig. 4.1a) and (Verdecia et al., 2003), highlighting its structural importance. The α1 helix has been previously described as a critical element for structural stability, yet an element dispensable for HECT domain function (Huang et al., 1999; Verdecia et al., 2003). Initial model-building into the 3.5 Å electron density showed a noticeable hydrophobic groove on the surface of helices 5, 11, 12, and 13, possibly indicating an additional helix being bound here. As sequence conservation drops off N-terminal of this α1 helix (Fig. 4.1b), we hypothesized that this element is an important part of the HECT domain. We note that expression of the HECT domain of Nedd4 yielded soluble folded product when the homologous helical segment was included in the expression construct, but was not successful in its absence (E. Maspero and S. Polo, personal communication). We therefore asked whether addition of helix α1 would not only assist in our crystallographic efforts but also affect the catalytic activity of the HECT domain.

Addition of helix α1 greatly stabilizes the HECT domain, as made evident in thermal denaturation experiments (Fig. 4.1c), shifting the transition midpoint by 16°C from 44 °C to 60°C. Although thermal stability differs between the two versions of the HECT domain, the level of secondary structure remains the same (Fig. 4.1d), indicating that the absence of helix α1 does not lead to unfolding, but to less rigidity of the domain. We solved the structure of the helix-extended HECT domain by molecular replacement...
using the E3 ligase WWP1 (Verdecia et al., 2003) as a search model. The final model (R/R\textsubscript{free} 19.3%/24.5%) was built and refined to 1.9 Å resolution (Table 1). The structure of HUWE1 HECT domain closely resembles that of WWP1, with which it shares 41.3% sequence identity (Fig. 4.2a). The HUWE1 HECT domain contains two distinct lobes similar to previously determined HECT domain structures (E6AP, Smurf2, WWP1). The larger N lobe (residues 3993-4252) contains the E2 binding region, and the smaller C lobe (residues 4259-4374) contains the conserved catalytic cysteine (C4341). The N lobe is composed of 13 α-helices and 7 β-strands, and the C lobe is composed of 4 α-helices and 4 β-strands. Residues 4253-4258 form the hinge that connects the two lobes. A rotary movement about this linker likely repositions the N and C lobes to bring the catalytic cysteine of the cognate Ub-loaded E2 in proximity to its E3 counterpart (Verdecia et al., 2003). Like WWP1, HUWE1 is oriented in an inverted T shape (⊥), in which the C lobe is positioned over the middle of the N lobe, with approximately 800 Å\textsuperscript{2} of contact surface area (Fig. 4.2b). Hydrogen bonds between Glu4248 (N lobe) and Ser4304 (C lobe), as well as Gln4245 (N lobe) and Gln4298 (C lobe), and a salt bridge between Glu4246 and Lys4295, stabilize the ⊥ conformation. The ⊥ conformation is further stabilized by water-mediated hydrogen bonds between the two lobes, involving residues Arg4130, Glu4147, Ser4148, and Glu4244 from the N lobe and Gln4298, Thr4340, Gly4302, and Lys4295 in the C lobe. The orientation of the N and C lobes of the HUWE1 HECT domain differs from the more open conformation observed in the crystal structures of E6AP and Smurf2 (Fig. 4.2c) (Huang et al., 1999; Ogunjimi et al., 2005), although we cannot exclude the possibility that crystal contacts influence the observed orientation of the C lobe. The stabilizing nature of helix α1 is apparent from the extended structure, as it closes the hydrophobic core of the N-lobe (Fig. 4.2a, d).
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

The most notable difference between HUWE1 HECT domain and previously solved structures concerns the E2 binding region (residues 4150-4200). Most of the hydrophobic residues in WWP1 that mediate contact with the E2 are similar to those in HUWE1, obvious from the alignment between HECT E3 ligases (Fig. 4.2e). The HUWE1 E2 binding region differs from that of WWP1, in that it contains additional structured elements – mainly ordered β-strands not previously identified. The well-ordered β-strands in the E2 binding region of the HUWE1 HECT domain extend farther from the helical core of the protein than seen in the structure of WWP1, and the loop is folded back on itself to complete the β3 strand and form the α8 helix (Fig. 4.2a). It is possible that HUWE1 uses its unique E2 binding region to interact with a specific set of E2 enzymes in vivo that differ from WWP1.

A comparison of our two HUWE1 HECT domain structures shows that they are nearly identical with respect to the positioning of the N and C lobes (Fig. 4.7a, b). For the structure lacking the α1 helix, the additional β-strands and α-helix seen in the E2 binding region remain unresolved, likely due to the low resolution data and high temperature factors.

Catalytic activity of the HECT domain

As addition of the α1 helix to the HECT domain stabilized the protein, we asked whether the presence of this structural element affected HECT domain catalytic activity. We hypothesized its addition might confer altered catalytic properties to the HECT domain compared to its helix-lacking counterpart. We therefore examined the ability of the HECT domain to catalyze self-ubiquitination in the presence of E1 and E2 (UBE2L3) enzymes, an ATP regenerating system, and [32P]-Ub (Fig. 4.3). The use of [32P]-Ub allowed us to quantify the amount of Ub adducts formed and calculate initial rates of
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

product formation in HECT-domain limiting conditions. In this assay, the HECT domain catalyzes the formation of a complex mixture of self-ubiquitinated species (Fig. 4.3a, b) that are not observed in absence of the HECT domain (lane marked “N”). Immunoblotting using an anti-His antibody confirmed that these species are ubiquitinated E3 enzyme, as it is the only species in this reaction that contains a polyhistidine tag (data not shown). The pattern of autoubiquitination observed is similar regardless of the presence of the α1 helix – both versions form multi- and polyubiquitinated species (Fig. 4.3a, b). Although the pattern of product formation is similar, the presence of the α1 helix suppresses the autoubiquitination activity of the HECT domain by more than 25-fold (Fig. 4.3c). As autoubiquitination is observed for many Ub ligases and is often used as a criterion of E3 Ub ligase activity, we sought to further characterize the reasons for its modulation.

The autoubiquitination reaction described above produces a complex mixture of products. We examined HECT domain activity in a single-turnover reaction to monitor the first round of Ub addition to the HECT domain. This assay encompasses two steps. In the first step, E2~Ub thioester is generated by incubating E1, E2, an ATP regenerating system, and Ub. After the E2~Ub thioester has formed, this reaction is quenched by the addition of EDTA to prevent further E1-catalyzed activation of Ub. In the second step, the HUWE1 HECT domain is added, and Ub is chased from the E2~thioester onto the HECT domain (Eletr et al., 2005). The use of a mutant version of Ub, in which all lysines are mutated to arginine (K0 Ub), prevents polyubiquitin chain formation on the HECT domain. Ub-conjugated HECT domain is visualized using anti-Ub immunoblot (Fig. 4.3d). We find that the HUWE1 Δα1 HECT domain shows increased activity under single turnover conditions compared to the HECT domain.
containing helix α1 (Fig. 4.3d), confirming the rate differences observed in the autoubiquitination assay.

**Thioester formation in the HECT domain**

Catalysis by HECT domain E3 enzymes is a multi-step process. The E3 enzyme binds Ub-loaded E2 and substrate, followed by Ub transfer between the E2 and E3 catalytic cysteines. The E3 then catalyzes isopeptide bond formation between Ub and a lysine residue on the substrate, which may be the E3 itself, Ub, or another protein. We next determined whether the presence of the α1 helix affects this upstream step, in which the catalytic cysteine of the E3 enzyme forms a thioester bond with ubiquitin.

We first attempted to assay thioester formation using the wild-type HECT domain, but the enzyme efficiently catalyzes formation of the isopeptide bond on a time scale too fast to measure (Singer et al., 2008; Wang and Pickart, 2005). Instead, we analyzed thioester formation using a four-amino acid, C-terminal truncation of the HECT domain (Salvat et al., 2004). This truncation removes a crucial determinant for isopeptide bond formation, a conserved phenylalanine located four amino acids from the C terminus of most HECT E3s (Salvat et al., 2004). HUWE1 Δ4 displays a diminished rate of isopeptide bond formation, allowing us to monitor formation of thioester-linked Ub to the enzyme. After incubation of the HECT domain with E1, the E2 UBE2L3, Ub, and an ATP regenerating system, the reaction is quenched with SDS-PAGE loading buffer with or without β-mercaptoethanol, and following electrophoretic resolution, is analyzed by anti-Ub immunoblot. The presence of the α1 helix greatly reduces the rate of thioester formation (Fig. 4.4), proportional to its suppression of autoubiquitination activity. The α1 helix, located on the back surface of the N lobe, is clearly not sufficiently close to interact with the E2 binding region of the HECT domain (Fig. 4.2a), suggesting
that the vibrational disorder in this protein contributes to the HECT domain-E2 interaction.

**Substrate ubiquitination catalyzed by the HECT domain**

Having seen that removal of helix α1 destabilizes the HECT domain and increases its autoubiquitination activity; we asked whether this effect is also observed during substrate ubiquitination. The anti-apoptotic Bcl-2 family member Mcl-1 is an in vivo target of HUWE1 (Zhong et al., 2005). HUWE1 recruits Mcl-1 via its BH3-domain, while the HECT domain presented here catalyzes Mcl-1 ubiquitination. Although Mcl-1 is a substrate of the full-length HUWE1, we use this assay with the isolated HECT domain here as a measure of non-self ubiquitination activity with an in vivo verified substrate of HUWE1. We examined initial rates of product formation, under HECT-domain limiting conditions, to determine whether the intrinsic activity of the HECT domain toward substrate is altered by the destabilizing effect of removing the α1 helix (Fig. 4.5a, b). We find that the HECT domain lacking α1 helix is ~5-fold more active in catalyzing Mcl-1 ubiquitination than the more stable HECT domain containing helix α1 (Fig. 4.5c). These results also suggest that autoubiquitination of the HECT domain does not impair catalytic activity toward substrate. The two versions of the HUWE1 HECT domain, which differ 25-fold in autoubiquitination rates, show only a 5-fold difference in their Mcl-1 ubiquitination rates. A similar observation was made for the heterodimeric complex of the minimal catalytic domains of Ring1a/Bmi1, in which autoubiquitination of the Ring1b protein did not affect E3 ligase activity toward its substrate, histone H2A, in an in vitro reconstituted system (Buchwald et al., 2006). We also observe similar differences in autoubiquitination and Mcl-1 ubiquitination activity between the two versions of the HECT domain at 37°C (Fig. 4.6).
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

Catalytic activity of the C4341A mutants

Mutation of the conserved catalytic cysteine to alanine (C4341A) abolishes activity of the HECT domain (Fig. 4.8). In the case of the helix-lacking HECT domain, we consistently observed that the C4341A mutant is capable of transferring a single ubiquitin to self (Fig. 4.8) or Mcl-1 (Fig. 4.8). These species are not generated when the HECT domain is omitted from the reaction (lane marked “N”). Although the failure of mutation of the catalytic cysteine to abolish activity has been previously observed (Adhikary et al., 2005), quantification of the monoubiquitinated species shows that this activity represents at best a minor fraction of wild-type activity (Fig. 4.8).

Discussion

We present here crystal structures of the HUWE1 HECT domain and identify a conserved structural element, helix α1, which stabilizes the HECT domain and tightly modulates its activity. Helix α1 is present in the structure of the WWP1 HECT domain (referred to in the WWP1 structure as H1’) (Verdecia et al., 2003), where the authors note that it plays an obvious role in contributing to HECT domain stability. As the H1’ helix is oriented between the C lobe and domains N-terminal to the HECT domain that presumably mediate protein-protein interactions, the authors suggested that H1’ helix contributes to target protein specificity in reactions catalyzed by the HECT domain. We confirm that the α1 helix is indeed crucial for stability and identify a role for this structural element in modulating HECT domain activity, as judged by autoubiquitination and Mcl-1 ubiquitination assays. Further experiments will determine whether this conserved helix modulates activity of other members of the HECT domain family.

In the absence of the N-terminal α1 helix, the HUWE1 HECT domain gains activity relative to its helix-extended counterpart. What could be the reason for this
unexpected behavior? Deletion of helix α1 might expose hydrophobic residues that trigger assembly of HUWE1 HECT domains into oligomers. Such behavior has been observed in the crystals of E6AP (PDB codes 1C4Z and 1D5F) (Huang et al., 1999). However, HUWE1 HECT Δ α1 behaves as a monomer and is properly folded in solution, as judged by several criteria (Fig. 4.1c and data not shown).

Our structural data on HUWE1 shows that the HECT domain adopts the same conformation regardless of the presence of helix α1, and both variants contain the same helical content (Fig. 4.1c). HUWE1 Δ α1, however, is far more active in catalyzing self-ubiquitination and in single-turnover assays. It also accepts Ub from the E2 UBE2L3 more readily than the helix-extended HECT domain. Furthermore, we observed elevated temperature factors, indicative of conformational flexibility, in the crystal structure of HUWE1 HECT Δ α1; a similar observation was made in the crystals of the E3 Ub ligase IpαH from the bacterial pathogen Shigella flexneri (Singer et al., 2008). We favor the interpretation that removal of helix α1 destabilizes the HECT domain to produce a more relaxed version of the enzyme that exhibits greater intra-domain flexibility. This increased flexibility allows the enzyme to sample more conformational states, thereby increasing its level of activity. Some of these conformational states may resemble the extended HECT domain structures observed in the crystal structures of Smurf2 and E6AP, in which the C-lobe has rotated about the flexible linker that connects the two subdomains of the HECT domain. In this scenario, removal of the α1 helix is analogous to the linker-extension mutations made in WWP1 (Verdecia et al., 2003). The removal of helix α1 may also shift the conformational equilibrium of the HECT domain into an orientation that facilitates the E2-HECT interaction or product release. This possibility is supported by evidence that enzymes exist in a dynamic range of
conformations, and the equilibrium between these different conformers can be shifted by mutation (Eisenmesser et al., 2005).

We did not anticipate that destabilization of the HECT domain would increase enzymatic activity. The Ub transfer reaction involves defined regions including the ordered β-strands that describe the E2 binding region and the catalytic site surrounding residue C4341. However, other steps, such as product release, may contribute to catalytic rate and may be influenced by increased conformational flexibility (Tokuriki and Tawfik, 2009). A correlation between conformational flexibility and promiscuous activity has been observed for several other proteins (Tokuriki and Tawfik, 2009). An example of a flexible enzyme is cytochrome P450, which can adopt a range of different conformations that allow it to act upon a variety of substrates. Among the P450 family of enzymes, the rigid CYP2A6 enzyme exhibits limited substrate specificity, whereas the highly flexible CYP3A4 is far more promiscuous (Tokuriki and Tawfik, 2009). In the case of HUWE1 HECT domain, the α1 helix may serve to impose a constraint on the inherent flexibility of the catalytic domain, thus fine-tuning enzymatic activity.

Autoubiquitination is often used as a criterion of E3 Ub ligase activity and, for some ligases, has been proposed as a mechanism of self-regulation of stability and downstream signaling functions (Varfolomeev et al., 2007; Wiesner et al., 2007). Our data show that this type of activity can be largely suppressed by minor extensions of what has been considered the core catalytic domain. We note, however, that our study focuses on the HECT domain of a multi-domain protein, and there may exist other structural elements in the 482-kD Huwe1 protein that affect its activity. This has been observed for the E3 Ub ligases Smurf2 (Wiesner et al., 2007), IpaH9.8 (Singer et al., 2008), and SspH2 (Quezada et al., 2009), in which domains N-terminal to the catalytic domain suppress autoubiquitination activity. While autoubiquitination is clear evidence of
catalytic activity of Ub ligases, the functional relevance of this reaction remains to be established for many E3s, including HUWE1. The increase in activity seen in the absence of helix α1 appears to stem from increased conformational flexibility in the enzyme. It is difficult to rationalize this behavior from static crystal structures, yet increased thermal motion observed in the Δhelix α1 structure is at least an indirect indicator. The significance of thermal motion within a protein with respect to reaction parameters is an emerging theme (Lange et al., 2008). HECT domains may have diverged to arrive at their extant spectrum of substrates by modulating flexibility (Tokuriki and Tawfik, 2009), in addition to more tractable changes of surface properties. We conclude that the activity of the HUWE1 HECT domain is tightly modulated, through restriction of conformational space rather than steric considerations, by the presence of a 19-residue helix α1.

Methods

Plasmids

HECT domain constructs of HUWE1 (amino acids 3993-4374 or amino acids 4012-4374) were cloned into a modified pET-28a plasmid (Novagen) containing a human rhinovirus 3C (HRV3C) protease site to generate an N-terminal His$_6$ fusion protein for use in biochemical assays. The catalytic mutants C4341A and C4099A/C4184A/C4367A were generated using site-directed mutagenesis (Stratagene). For biochemical assays with radiolabeled substrate, Flag-Mcl-1 (amino acids 1-327) and Ub were both cloned with an N-terminal protein kinase A (PKA) site for $^{32}$P-labeling into pET-16b and pET-28a with the HRV3C site, respectively, as previously reported (Love et al., 2009). UBE2L3 was cloned into the pET-28a plasmid (Novagen).
**Bacterial protein expression and purification**

All versions of the HUWE1 HECT domain were expressed and purified as previously reported (Love et al., 2009). ['{}^{32}\text{P}\text{-}labeled proteins were purified and labeled as previously described (Love et al., 2009). Native UBE2L3 was expressed in Rosetta (DE3) cells (Novagen). UBE2L3 was precipitated from bacterial lysate by addition of saturated ammonium sulfate to 90%. The precipitated protein was resuspended in 50 mM HEPES pH 7.4, 200 mM NaCl and purified by gel filtration (Superdex 75 PC 3.2/30, GE Healthcare).

**Circular dichroism**

HUWE1 Δ α1 and HUWE1 + α1 HECT domains were dialyzed into 5 mM HEPES pH 7.5, 100 mM NaCl immediately prior to the scanning and melting CD experiments using an AVIV model 202 CD spectrometer. HUWE1 Δ α1 HECT domain at 2.4 μM and HUWE1 + α1 HECT domain at 2.8 μM were used for scanning experiments between 195 and 280 nm at 25°C. CD signal at 222 nm of 4.8 μM HUWE1 Δ α1 and 5.6 μM HUWE1 + α1 was recorded every 2 degrees over a 20-94 °C temperature ramp with 2 minutes of equilibration time at every step.

**Biochemical assays**

Reaction mixtures (10 μl) for HUWE1 autoubiquitination assay contained 100 nM human E1 (Ube1, Boston Biochem), 5.6 μM E2 (UBE2L3), HECT domain, and 60 μM ['{}^{32}\text{P}\text{-}Ub with an ATP regenerating system (50 mM Tris [pH 7.6], 5mM MgCl₂, 5 mM ATP, 10 mM creatine phosphate, 3.5 U/ml creatine kinase). Reactions were incubated at room temperature and aliquots were removed after the indicated amount of time and terminated in reducing SDS-PAGE sample buffer. Samples were boiled 10 min. and
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

separated on 8% Tris-tricine SDS-PAGE. Dried gels were exposed to a phosphor screen. \[^{32}\text{P}\]-Ub bands were visualized by autoradiography and quantification of data was performed using a phosphorimager. Background correction for each sample was performed by subtracting the counts from an equivalent area of the gel from a lane containing all reaction components except E3 enzyme (lane marked "N"). The \[^{32}\text{P}\]-Ub signal from this lane was used to convert the observed sample counts to a concentration value. The concentrations of HUWE1-[\(^{32}\text{P}\)]-Ub were measured and rates of product formation determined by fitting the initial linear data points to a least squares regression line.

**Mcl-1 ubiquitination assay**

Reaction mixtures (10 µl) for the Mcl-1 ubiquitination assay were set up as above except with the addition of 5 µM \[^{32}\text{P}\]-Flag-Mcl-1 and 100 µM Ub (Sigma). Reactions were quenched with reducing sample buffer and separated on 10% SDS-PAGE. Bands from dried gels were analyzed as above.

**Thioester assay**

Reaction mixtures for the thioester assay (10 µl) contained 100 nM human E1 (Ube1, Boston Biochem), 5.6 µM E2 (UBE2L3), 2 µM HUWE1Δ4 HECT domain, and 60 µM Ub (Sigma) with an ATP regenerating system described above. Reactions were incubated at room temperature and aliquots were removed after the indicated amount of time and terminated in 4M urea and incubated 15 min. at 30°C, or terminated in reducing SDS-PAGE sample buffer. Samples were boiled 10 min., separated on 10% Tris-glycine SDS-PAGE, and analyzed by immunoblot using anti-Ub antibody (Sigma).
**Single-turnover assay**

For the single-turnover assay, the E2−Ub thioester was generated in a 20 μl reaction containing 200 nM E1 (Boston Biochem), 8 μM E2, ATP regenerating system described above, 60 μM mutant Ub in which all lysines are mutated to arginine (K0 Ub) (Boston Biochem), and 1 μg/μl BSA incubated 25 min. at room temperature. Formation of the E2−Ub thioester was quenched with 50 mM EDTA on ice for 5 min. The E2−Ub thioester was diluted into chase mixture containing 2 μM HECT domain, 100 mM NaCl, 50 mM EDTA, and 1 μg/μl BSA, or the same reaction components lacking the HECT domain (labeled "N"). Reactions were incubated at room temperature and aliquots were removed after the indicated amount of time and terminated in either 4 M urea and incubated 15 min. at 30°C, or in reducing SDS-PAGE sample buffer. Samples were boiled 10 min., separated on 10% Tris-glycine SDS-PAGE, and analyzed by immunoblot using anti-Ub antibody (Sigma).

**Crystallization of HUWE1 HECT domain**

Crystallization experiments with purified HUWE1 HECT domain including the N-terminal His6 tag, HRV3C protease site and with C4099A, C4184A, and C4341A mutations, were set up in 96-well sitting drop trays using commercially available sparse-matrix screens (Hampton Research, Qiagen). The initial crystals were improved in hanging-drop vapor diffusion setups. The HUWE1 + α1 HECT domain crystallized by mixing 1 μl of protein sample concentrated to 17 mg/ml with a 1 μl solution containing 0.1 M citric acid (pH 5.2) and 1.8 M (NH₄)₂SO₄. Birefringent crystals in the shape of thick rods with dimensions of approximately 80 x 40 x 40 μm grew within two days of incubation at 18 °C. The HUWE1 Δ α1 HECT domain crystallized by mixing a 1 μl
solution containing (Na/K)$_2$PO$_4$ (pH 6.5) and 1.4 M (Na/K)$_2$PO$_4$. Thin rod-shaped crystals grew within 10 days at 23°C.

Data collection and processing

For native X-ray diffraction studies, crystals were cryoprotected by soaking in 0.1 M citric acid (pH 5.2), 1.8 M (NH$_4$)$_2$SO$_4$, 12% glycerol for 30 sec prior to vitrifying in liquid nitrogen. X-ray diffraction data were collected on a single cryogenized crystal at beamline 24ID-E, Advanced Photon Source (Argonne, IL, USA), summarized in Table 4.1. Data were processed using DENZO and SCALEPACK (Otwinowski, 1993). The crystals belong to the monoclinic space group C2 and diffracted to 1.9 Å. Initial phases were obtained by molecular replacement using PHASER from the CCP4 crystallographic program suite (Collaborative Computational Project, 1994; McCoy, 2007), with the coordinates of the E3 ligase WWP1 (PDB accession code 1ND7) as search model. The final model was refined at resolution of 1.9 Å using PHENIX (Verdecia et al., 2003). Details of refinement are given in Table 4.1. Figures were made using PyMol (DeLano, 2002).

Acknowledgements

We thank Christian Schlieker for helpful discussions and critical reading of the manuscript, and Fenghe Du and Xiaodong Wang at UT Southwestern Medical Center for providing cDNA plasmids for HUWE1 (Mule) and Mcl-1. RKP is supported by a US Department of Defense Breast Cancer Research Program award (W81XWH-06-1-0789). KRL was supported by an NIH postdoctoral fellowship (F32 AI63854).
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

The abbreviations used are: HECT, homologous to E6AP C-terminus; Ub, ubiquitin; HUWE1, HECT, UBA and WWE domain containing 1. Protein coordinates have been deposited in the RCSB Protein Data Bank (PDB code 3H1D).
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

Figures – Chapter 4

Figure 4.1 - The α1 helix stabilizes the HUWE1 HECT domain
(a) Multiple sequence alignment of helix α1 with a diverse set of human HECT E3 ligases. Residue conservation is indicated by degree of shading ranging from orange (most conserved) to light yellow (least conserved). Secondary structure is illustrated with α-helices as cylinders and β-sheets as arrows. (b) Multiple sequence alignment with a diverse set of human HECT E3 ligases indicating that sequence conservation drops off N-terminal to the α1 helix. The N-
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

terminus of the HECT domain + al helix is indicated. (e) Thermostability of the HUWE1 HECT domain was measured in a CD melting experiment. HUWE1 HECT domain, +/D al helix, wild-type or with cysteine mutations, was heated in a circular dichroism cuvette and unfolding measured at 222 nm as a loss of helical content. Deletion of helix al results in reduction of thermostability. (d) A CD scan demonstrates structural similarities of the +/D al helix versions of the HECT domain.

Figure 4. 2 - Structure of the HUWE1 HECT domain
(a) Stereo view of HUWE1 HECT domain (residues 3993-4374) showing the N- and C-lobes connected by the hinge loop. Helix al is colored green. The N-lobe contains the E2 binding region and the C-lobe contains the catalytic cysteine (C4341). (b) Overlay of HUWE1 (blue) and WWP1 (orange; PDB 1ND7) crystal structures. (c) Overlay of HUWE1 (blue) and Smurf2 (orange; PDB 1ZVD) crystal structures. (d) Helix al plays a significant role in mediating
hydrophobic contacts that maintain the core of the HUWE1 HECT domain. Hydrophobic residues in the α helix, Phe3994 and Phe4001, pack into hydrophobic pockets in the N lobe. Arg3998 and Asp4009 form hydrogen bonds stabilizing the N-lobe. Lys4014 and Tyr4119, C-terminal to the α helix, orient the α helix to further stabilize the N-lobe. (e) Multiple sequence alignment of E2 binding region with a diverse set of human HECT E3 ligases. Residues important for E2 binding are indicated with blue circles.
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

Figure 4.3 - E3 Ubiquitin ligase activity of HUEW1 HECT domain

The autoubiquitination activity of HUEW1 HECT domain was tested using 60 µM [32P]-Ub as substrate and (a) 2 µM WT D α1 or (b) 2.9 µM WT + α1 HECT domains incubated with UBE1, UBE2L3, and an ATP regenerating system. (Note the different time scale for the two variants of the HECT domain). HECT domain is omitted in the lane marked “N”. Asterisk denotes a likely ubiquitin polymer; double asterisk denotes likely mono-ubiquitinated UBE2L3. Concentrations of HECT domain were chosen to obtain initial rate conditions. (c) Ligation activity of the WT D α1 and WT + α1 HECT domains in the autoubiquitination assay. Activity is given as the ratio between initial velocity (pmol total [32P]-Ub product/min) and total enzyme concentration E (pmol). Errors are the standard deviations calculated from three independent experiments, shown in parenthesis. (d) Single turnover assay monitoring transfer of Ub from the UBE2L3-Ub thioester to a lysine in the WT D α1 and WT + α1 HECT domains. The UBE2L3-Ub thioester is generated in a pulse reaction containing E1, UBE2L3, ATP regenerating system, and a Ub mutant in which all lysines are mutated to arginine (K0 Ub). Ub is chased from the E2 enzyme to HECT domain added to the reaction. Ub-conjugated HECT domain is visualized by anti-Ub.
immunoblot. Samples were terminated in reducing or non-reducing sample buffer as indicated. Panel marked “N” is a chase reaction performed in absence of HECT domain and terminated in non-reducing sample buffer.

Figure 4.4 - Detection of Ub-thioesters in the HUWE1 HECT domains
Ub-thioester assay with the indicated HUWE1 HECT domain proteins. Purified HECT domains containing a C-terminal truncation of the last four amino acids were incubated with E1, UBE2L3, an ATP regenerating system, and Ub for the indicated amount of time. Reactions were stopped with 4M urea and non-reducing sample buffer (upper panel) or reducing sample buffer (lower panel), separated by SDS-PAGE, and analyzed by immunoblot with anti-Ub antibody.
Figure 4.5 - Substrate ubiquitination activity of the HUWE1 HECT domains

(a,b) The Mcl-1 ubiquitination activity of HUWE1 HECT domain was tested using 5 μM [32P]-Mcl-1 as substrate and (a) 100 nM WT Δ α1 or (b) 300 nM WT + α1 HECT domains incubated with UBE1, UBE2L3, Ub, and an ATP regenerating system. HECT domain was omitted from the lane marked "N". Concentrations of HECT domain were chosen to obtain initial rate conditions. (c) Ligation activity of the HECT domains in the Mcl-1 ubiquitination assay. Activity is given as the ratio between initial velocity (pmol total [32P]-Ub product/min) and total enzyme concentration E (pmol). Errors are the standard deviations calculated from three independent experiments, shown in parenthesis. (d, e) Graph showing percent ubiquitinated Mcl-1 as a function of time in the reactions shown in panels (a) and (b) catalyzed by HUWE1 Δ α1 (d) or HUWE1 + α1 (e) HECT domains.
Figure 4.6 - Ubiquitination activity of the HUWE1 HECT domains at 37°C
(a) The autoubiquitination activity of HUWE1 HECT domain at 37°C was tested using 60 μM Ub as substrate and 2 μM WT HECT domains incubated with UBE1, UBE2L3, and an ATP regenerating system. Reaction mixtures were separated on SDS-PAGE and analyzed by anti-Ub immunoblot. (b) The Mcl-I ubiquitination activity of HUWE1 HECT domain at 37°C was tested using 5 μM Flag Mcl-I as substrate and 100 nM WT HECT domains incubated with UBE1, UBE2L3, Ub, and an ATP regenerating system. Reaction mixtures were separated on SDS-PAGE and analyzed by anti-Flag immunoblot.
Figure 4.7 - Both HUWE1 Δa1 and +a1 HECT domains maintain the \( \perp \) conformation
(a) Overlay of the \( \text{C}_\alpha \) backbone of the HUWE1 + a1 HECT domain (orange) and HUWE1 D a1 HECT domain (blue). (b) The \( \text{C}_\alpha \) backbone of the HUWE1 + a1 HECT domain fit into electron density of the HUWE1 D a1 structure demonstrates the similarity of these structures.
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

Figure 4. 8 - Mutation of the catalytic cysteine (C4341) to alanine abolishes activity of the HECT domain

(a, b) The autoubiquitination activity of C4341A HECT domain was tested using 60 μM [32P]-Ub as substrate and (a) 11.6 μM C4341A Δ α1 or (b) 12.1 μM C4341A + α1 HECT domains incubated with UBE1, UBE2L3, and an ATP regenerating system. Aliquots of the reaction mix were removed after the indicated amount of time, quenched in reducing SDS sample buffer, and separated on SDS-PAGE. HECT domain is omitted in the lane marked “N”. Asterisk denotes a likely ubiquitin polymer. (c, d) Substrate ubiquitination activity was measured using 5 μM [32P]-Mcl-1 as substrate and (a) 11.6 μM C4341A Δ α1 or (b) 12.1 μM C4341A + α1 HECT domains incubated with UBE1, UBE2L3, Ub, and an ATP regenerating system. HECT domain is omitted in the lane marked “N”. (e, f) Ligation activity of the HECT domains in the (e) autoubiquitination or (f) Mcl-1 ubiquitination assay. Activity is given as the ratio between initial velocity (pmol total [32P]-Ub product/min) and total enzyme concentration E (pmol). Errors are the standard deviations calculated from three independent experiments, shown in parenthesis.
Table 4.1 – Crystallographic data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>HUWE1 HECT + α1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
</tr>
<tr>
<td>Wavelength</td>
<td>0.9793</td>
</tr>
<tr>
<td>Space group</td>
<td>C2</td>
</tr>
<tr>
<td>Cell dimensions:</td>
<td></td>
</tr>
<tr>
<td>a,b,c (Å)</td>
<td>119.6, 56.6, 69.6</td>
</tr>
<tr>
<td>β (°)</td>
<td>122.5</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>30,847</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>30-1.9 (1.93-1.9)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>30-1.9</td>
</tr>
<tr>
<td>Nonhydrogen Atoms</td>
<td>3190</td>
</tr>
<tr>
<td>Water Molecules</td>
<td>357</td>
</tr>
<tr>
<td>R_work b</td>
<td>16.6</td>
</tr>
<tr>
<td>R_free c</td>
<td>22.9</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.01</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.20</td>
</tr>
<tr>
<td>B factors (Å²)</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>30.0</td>
</tr>
<tr>
<td>Water</td>
<td>39.4</td>
</tr>
<tr>
<td>Coordinate error (Å)</td>
<td>0.68</td>
</tr>
<tr>
<td>Ramachandran plot d</td>
<td></td>
</tr>
<tr>
<td>Most favored (%)</td>
<td>93.1</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>6.9</td>
</tr>
<tr>
<td>Generously allowed (%)</td>
<td>0</td>
</tr>
<tr>
<td>Disallowed (%)</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ R_{\text{sym}} = \frac{\sum |I_i - \langle I \rangle|}{\sum I_i} \text{ where } I_i \text{ is the intensity of the } i\text{th observation and} \langle I \rangle \text{ is the mean intensity of the reflection.} \]
Chapter 4: A Structural Element Within The HUWE1 HECT Domain Modulates Self- and Substrate Ubiquitination Activities

\[ R_{\text{f.i.m.}} = \frac{\sum_{hkl} |l(hkl)| - \langle l(hkl) \rangle}{\sum_{hkl} l(hkl)} \]

where \( l(hkl) \) is the observed intensity and \( \langle l(hkl) \rangle \) is the average intensity of multiple observations of symmetry-related reflections.

\[ R_{\text{p.i.m.}} = \frac{\sum_{hkl} |l(hkl)| - \langle l(hkl) \rangle}{\sum_{hkl} l(hkl)} \]

where \( l(hkl) \) is the observed intensity and \( \langle l(hkl) \rangle \) is the average intensity of multiple observations of symmetry-related reflections.

\[ R_{\text{work}} = \frac{\sum_{hkl} |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum_{hkl} |F_{\text{obs}}|} \]

\[ R_{\text{free}} = \frac{\sum_{hkl} |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum_{hkl} |F_{\text{obs}}|} \]

Highest resolution shell is shown in parenthesis.

Calculated with the program MolProbity (Davis et al., 2007).
Chapter 5: Conclusion
Summary

The body of work presented in this dissertation comes from a collection of three distinct projects that utilize a crystallographic and biophysical approach towards understanding protein-protein interactions. Two studies have focused on characterization of reassembled nucleoporin subcomplexes and the third project concerns the structure and function of an E3 HECT domain ubiquitin ligase. Details and results for each project have been summarized below, with some thoughts on what future studies might entail.

Ran-binding zinc fingers from Nup153 bind weakly at the switch I region of RanGDP. The strongest binding constant is measured at \(-5\) pM for two out of four Nup153 zinc fingers and \(\sim 50\) pM for the other two. RanGTP also binds to Nup153 zinc fingers albeit with lower affinity. Considering the weak binding constant and symmetric localization across NPC, the Ran-binding zinc finger domain helps to facilitate proper regulation of nucleocytoplasmic transport by sequestering RanGTP and RanGDP in close proximity to the NPC. These interactions are likely to be fast on/off interactions that play a subsidiary role in the Ran regulated transport through the NPC.

Reconstitution of a tetrameric complex between the Nsp1 trimeric subcomplex (Nsp1, Nup57, Nup49) and full-length Nic96 marks a significant step in elucidating structural properties that govern assembly of the NPC. The coiled-coil domains of the Nsp1 complex are sufficient to assemble a trimeric complex that interacts with full-length Nic96. Helical deletions at the N- and C-terminus of Nup57 confirm that Nup57 interacts directly with Nsp1 and that this interaction is necessary for copurification of Nsp1 with Nup49 and Nup57. We have also demonstrated that a core interaction domain exists between Nup57 and Nup49. Biophysical analysis with ultracentrifugation demonstrates that the Nsp1 trimeric complex contains one molecule of Nsp1, Nup57, and Nup49 per
assembly. Purification of Nup49 alone forms a tetrameric assembly in the absence of Nup57 and Nsp1, in agreement with previous structural data for the Nup49 homologue Nup58. Assembly of a tetrameric species between the Nsp1 trimeric complex and Nic96 has been characterized with gel filtration and further experiments are necessary to fully characterize this large coiled-coil assembly.

Structural characterization of the HECT domain from the E3 ubiquitin ligase HUWE1 is a significant deviation from our studies on the NPC. This difference is only topical however, as much of the same methods and scientific principles that were used throughout our studies of the NPC have been used to study HUWE1. During crystallization experiments with HUWE1 we found that a small conserved helical element (the α1-helix) at the N-terminus strongly influences substrate- and self-ubiquitination activity. Although the α1-helix strongly enhances thermal stability of HUWE1, the absence of the α1-helix does not disrupt the secondary structure of HUWE1. Structurally, the HUWE1 HECT domain is similar to previously solved HECT domains, exhibiting a conformation of the N- and C-lobes most similar to WWP1. Upon analysis of our structure and other similar HECT domains, it is surprising that a small helix would have such a drastic effect on thermal stability and ubiquitination activity. The α1-helix most likely contributes to stabilizing the N-lobe hydrophobic core and removal is enough to destabilize regulation of self- and substrate-labeling mediated by rotation of the C-lobe about the N-lobe.

**Future Directions**

*A role for Nup153 zinc fingers in nucleocytoplasmic transport*

Several research groups have now demonstrated that both Nup153 and Nup358 bind to RanGDP with low micro-molar affinity via the RanBP2-type zinc finger domain. Crystallographic studies have provided detailed information regarding the interaction
between RanBP2-type zinc fingers and RanGDP that exists in a 1:1 stoichiometric ratio with one molecule of each component per complex (Partridge and Schwartz, 2009; Schrader et al., 2008a). Although we understand the structural details of how RanBP2-type zinc fingers bind preferentially to RanGDP over RanGTP, we do not yet fully comprehend why Nup153 and Nup358 bind RanGDP at either face of the NPC. Several live imaging studies done with microinjected wild-type Ran, labeled with dye, have described accumulation of Ran at either side of the nuclear envelope, and with fluorescence imaging provide evidence that Ran exists at a surprisingly high concentration around the nuclear rim (Hinkle et al., 2002; Paradise et al., 2007; Smith et al., 2002). More recent analysis using EYFP-tagged wild-type Ran in HELa cells describe Ran accumulating at the nuclear envelope with an estimated 800 molecules of Ran/μm² nuclear envelope (Abu-Arish et al., 2009). The authors further explain that although Ran accumulates at the nuclear envelope, Ran is not evenly dispersed across the envelope but rather focused as discrete puncta around nuclear pores. With this in mind the authors calculate that nearly 200 wild-type Ran molecules surround one NPC at any given time (Abu-Arish et al., 2009). Our research supports the notion that a population of Ran binds to the NPC via low μM interactions with RanBP2-like zinc fingers of Nup153 and Nup358. It remains unclear if these interactions are intermediary steps of nucleocytoplasmic transport within the RanGDP-RanGTP cycle, or perhaps these interactions are simply responsible for localizing Ran near the NPC for interactions with transport factors. It is conceivable that Nup153 at the nuclear basket is a designated site for nucleotide exchange mediated by RCC1. RanGTP does bind to RanBP2-type zinc fingers with 50 μM affinity and following nucleotide exchange, RanGTP might remain at the nuclear basket and bind tightly with importin-β-cargo complexes to terminate import and release cargo into the nucleoplasm. These details
Chapter 5: Conclusion

remain unknown and future transport experiments will decipher the molecular basis for
direct NPC interaction with RanGDP mediated by RanBP2-type zinc fingers.

From a structural perspective the interaction between these two proteins is clear. A protein engineering approach allowed us to achieve very high-resolution data for each zinc finger in question, including tandem zinc finger pairs. What remains to be understood is the functional significance behind much of our binding data. We have shown with both isothermal calorimetry and crystallography that individual Nup153 zinc fingers bind favorably with RanGDP versus RanGTP. Nup153 and Nup358 are respectively localized to the nucleoplasmic and cytoplasmic face of the NPC. They each contain a varying number of Ran binding zinc fingers and based on our data Nup358 binds RanGDP with a slightly weaker affinity then Nup153. It is possible that both Nup153 and Nup358 play intermediary roles in the Ran cycle, such as recycling RanGDP back into the nucleus via NTF2 and the binding of RanGDP by RCC1 to mediate nucleotide exchange. In vitro experiments to monitor GTPase activity, localization, and nucleotide exchange could address some of these lingering questions. It is possible that the zinc fingers from Nup153 and Nup358 play a role in nucleotide exchange and GTPase activity, respectively. These questions could be addressed using radiolabeled nucleotides in the presence of components from the Ran pathway, including RCC1 and RanGAP1.

The tetrameric assembly between Nic96 and the Nsp1 subcomplex

The tetrameric assembly between the Nsp1 trimeric subcomplex and Nic96 has been successfully purified and crystallographic studies are ongoing. All Nsp1 complex assemblies have been screened for crystals in 96-well format, however no crystals of an assembled complex have been identified. We crystallized Nup49 in setups with the
trimeric complex, indicating that our purification scheme results in excess Nup49, or Nup49 falls off the trimeric complex. We are currently in the process of designing a series of constructs that utilize a dual-tag purification scheme to eliminate excess Nup49. In addition to removal of Nup49, we have designed a series of dual-tag constructs to increase purity and favor crystallization of the tetrameric coiled-coil assembly. We will also attempt to crystallize some of the Nsp1 assemblies as fusion proteins, fused to smaller proteins known to readily crystallize. We hope that this method will coax the coiled-coil domains into forming a stronger crystallographic lattice and crystals suitable for data collection. In addition to crystallographic studies the tetrameric complex with full-length Nic96 still requires biophysical analysis with analytical ultracentrifugation.

The HUWE1 HECT domain

A balance between conformational flexibility and enzyme promiscuity

A flexible hinge between the N- and C- lobes of the E3 HECT domain regulates conformational flexibility to facilitate rotation of the C-lobe for Ub transfer from E2 to substrate. The C-lobe pivots on this unstructured hinge and rotates over 100° to transfer Ub from docked E2 over 41 Å to substrate. It is an amazing structural accomplishment for the E3 HECT domain to develop such an active mechanism and adopt a range of conformations to regulate ubiquitination. Any modifications towards increased promiscuity by the E3 HECT domain could no doubt prove costly to the cell, as the Ub pathway regulates a range of vital cellular functions. We have focused our studies on a highly conserved helix at the N-terminus of the HUWE1 HECT domain, shown to regulate enzymatic activity (Pandya et al., 2010). Removal of the α1 helix did not detrimentally affect the overall fold of the domain, as measured with circular-dichroism
(CD) scans and shown with crystallographic data, however stability of the enzyme was substantially decreased when measured with CD melting curves. There is a disparate increase in activity between auto-ubiquitination and substrate ubiquitination, potentially demonstrating that the two mechanisms are differentially regulated by the α1 helix.

A case can be made for the importance of enzymes having some potential for conformational variability, given multiple selective pressures from the cellular environment. A recent review describes the significance of conformational variability in several enzyme families that can adopt multiple structural conformations under selective pressure to facilitate a host of cellular functions (Tokuriki and Tawfik, 2009). In the case of the E3 HECT domain, the N- and C-lobes have evolved to employ a mechanism of incredible flexibility that allows the C-lobe to sample a diverse conformational landscape for controlled transfer of ubiquitination between multiple E2 enzymes and an even greater number of substrates. Removal of the N-terminal helix does not necessarily result in a conspicuous change of function. Rather, the N-terminal helix serves as means to restrict conformational space sampled by the C-lobe and potentially directs activity towards specific E2 enzymes and substrates. Where we have used a sledgehammer a surgeon could use a fine scalpel. The helix in question is roughly twenty amino acids long, and within those twenty residues, there are three obvious hydrophobic residues that appear to be critical for maintaining stability of hydrophobic core of the N-lobe and two residues that make hydrogen bonds with the N-lobe. An interesting question is to examine mutations of these residues and observe potential effects on the Ub pathway, or what phenotypes would result in organisms harboring these mutations. It is entirely possible that only a single point mutation can push the balance between regulated activity and stability of this enzyme towards over activity and unregulated promiscuity.
Our structural studies on the HECT domain from HUWE1 unexpectedly uncovered a key element responsible for regulating activity of both self- and substrate-labeling with ubiquitin. From our structural data it is clear which residues from the α1-helix mediate key interactions with the N-lobe of the HECT domain. These residues could be mutated *in vivo* to monitor ubiquitination of model substrates, such as Mcl-1, or self-ubiquitination by HUWE1. To further assess contribution of the α1-helix *in vitro*, stability and activity could be monitored upon removal or mutation of specific residues from the α1-helix with labeling assays described in this manuscript. From a structural and functional perspective, there is still minimal molecular information regarding how the HECT domain facilitates transfer of ubiquitin from E2 to substrate. A recent crystal structure from the Schulman lab describes a trimeric complex between, E2, E3 HECT domain, and Ub, representing the most detailed "image" of the ubiquitin cascade to date. However we still know very little about how: 1) substrate is bound to the HECT domain and, 2) how ubiquitin is transferred between the HECT domain C-lobe and substrate. These questions will likely soon be answered with both structural and biochemical data, greatly advancing our knowledge of the ubiquitin pathway.


References


References


References


References


References


References


References


144


References


References


References


References


References


References


