The architecture of the nuclear pore complex

Thesis by

Daniel Hanyang Lin

In Partial Fulfillment of the Requirements for
the degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2017
(Defended June 2nd, 2017)

ACKNOWLEDGEMENTS

I am indebted to many, many people for supporting me, pushing me, and teaching me over the years. A brief chronological acknowledgement of their contributions:

I thank my parents first and foremost. My father inspired a love of science and logic – from semi-forced geometry proofs at a very young age to programming a visual basic game to practice my times tables – and did everything in his power to provide for our family. My father was the first PhD in his family but made the hard choice of making sure his children could live comfortably rather than chase an uncertain future in science. My mother always demanded the best out of her two sons, made sure we were healthy and well, tolerated us as we grew up, and always made sure we could focus on succeeding.

In addition, I'm grateful to my brother for always answering questions about how about computers work, for sharing almost every interest and hobby, and for brotherly competition and friendship.

My uncle Cai Dongqing welcomed me into his home in Guangzhou for a summer and his lab where I did my first experiments, for which I will always be grateful. In addition, Li Zhen, the graduate student who hosted me, went to great lengths to take care of me while I was there.

I am extremely grateful to Professor Niraj Tolia for taking me into his lab at the Washington University School of Medicine for 3 years, for giving me my own project, and introducing me to x-ray crystallography, as well as for an undergraduate research experience that is more

remarkable upon every reexamination. Aditya Radhakrishnan, who was working in the Tolia lab at the time I was looking for a lab to join, inadvertently set me on this path.

I thank Professor André Hoelz, who captured my imagination in a short presentation at recruitment, for tolerating my growing pains in the lab, for pushing me to have the highest standards, for providing unprecedented opportunities, for recruiting and assembling such an incredible team of people, and for finding ways to support me at every opportunity.

The past and current members of the Hoelz lab are an incredible group of people, with whom I have spent most of life these past five years: Tobias Stuwe, who was a mentor and amazing scientific colleague, but more importantly a compassionate friend and role model; Chris Bley, Ferdinand Huber, and Leslie Collins, who joined the lab in rapid succession shortly after I did, I thank for friendship, endless advice and support, and general mischief; Emily Rundlet, George Mobbs, and Thibaud Perriches, with whom I spent a very long year hammering out experiments by day and unwinding at night; Stefan Petrovic, who could always be counted on for fascinating late night discussions and his sense of humor; Karsten Thierbach, for always had something wise to share; Ana Correia, simply the greatest collaborator, who goes beyond the call of duty to help everyone both in and out of the lab; Sarah Cai, who was an unbelievable undergraduate researcher and possessed both inspiring fearlessness and dedication in tackling problems.

I also must thank the members of my committee, Professor Bil Clemons, Professor Shu-ou Shan, and Professor Doug Rees, for years of patient guidance and thoughtful advice. In addition, I'm grateful for all the help and support from the rest of the Caltech community, including the members of the Shan, Clemons, Rees, and Bjorkman labs.

Lastly, I will always be grateful to my close friends during these years, Ben Gross, Amanda Mock, Mikhail Hanewich-Hollatz, Lincoln Collins, Jeff Bosco, Luke Urban, Jacques Edeline, and Norah Kissell for a lifetime's worth of fun and memories, for exposing me to new perspectives, and inspiring me to be a better person. I am especially thankful that I met Mikhail and Amanda, because they have taught me so much about being a decent individual – mostly by example.

ABSTRACT

Nucleocytoplasmic transport, the regulated trafficking of macromolecules in and out of the nucleus, occurs primarily through nuclear pore complexes (NPCs). NPCs are massive macromolecular machines embedded in the nuclear envelope which generate ~40 nanometer transport channels to facilitate transport. Because of its size and complexity (~1000 subunits, ~120 MDa), the structure of the NPC has remained poorly understood. This thesis presents a bottom-up approach to understanding the structure and function of the NPC through reconstitution of the proteins and structural and biochemical studies. The first three chapters present work towards determining the composite structure of the symmetric core of the NPC. X-ray crystal structures are described for many of the components of the symmetric core. This includes a heterohexameric coat nucleoporin complex containing Nup120, Nup85, Nup145C, Sec13, Seh1, and Nup84, revealing how these proteins assemble into one of the main subcomplexes in the NPC. Reconstitution of the symmetric core components and analysis of the protein-protein interaction between the components provides a detailed biochemical map for the protein interaction network in the NPC. X-ray crystal structures of overlapping fragments facilitate the generation of accurate atomic model for full-length proteins. An iterative, sequential docking approach is developed to dock these models into a cryoelectron tomographic reconstruction of the human NPC, yielding a composite model for the structure of the symmetric core of the NPC. In the next two chapters, this analysis is extended to the cytoplasmic-specific decorations of the NPC. The structure of the C-terminal domain of Nup358 is reported and its catalytic activity is described. Lastly, reconstitution of human DDX19 activation by the NPC reveals mechanistic insight into how the NPC directly regulates the last step of mRNA export.

PUBLISHED CONTENT AND CONTRIBUTIONS

Tobias Stuwe[#], **Daniel H. Lin**[#], Leslie N. Collins, Ed Hurt, André Hoelz (2014). Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins, ***Proc. Natl. Acad. Sci. U.S.A.***, 111(7):2530-2535.

DOI: 10.1073/pnas.1311081111

Contributions: cloning, protein expression and purification, protein crystallization, structure determination, protein-protein interaction experiments, structure analysis

Tobias Stuwe[#], Ana R. Correia[#], **Daniel H. Lin**, Marcin Paduch, Vincent T. Lu, Anthony A. Kossiakoff, André Hoelz (2015). Architecture of the nuclear pore complex coat, ***Science***, 347(6226):1148-1152.

DOI: 10.1126/science.aaa4136

Contributions: structure determination and validation, structure analysis, computational docking into EM reconstruction

Daniel H. Lin[#], Tobias Stuwe[#], Sandra Schilbach, Emily J. Rundlet, Thibaud Perriches, George Mobbs, Yanbin Fan, Karsten Thierbach, Ferdinand M. Huber, Leslie N. Collins, Andrew M. Davenport, Young E. Jeon, André Hoelz (2016). Architecture of the symmetric core of the nuclear pore, ***Science***, 352(6283):aaf1015

DOI: 10.1126/science.aaf1015

Contributions: cloning, protein expression and purification, protein crystallization, structure determination, structure analysis, protein-protein interaction experiments, computational docking into EM reconstruction

Daniel H. Lin, Stephan Zimmermann, Tobias Stuwe, Evelyn Stuwe, André Hoelz (2013). Structural and functional analysis of the C-terminal domain of Nup358/RanBP2, ***J. Mol. Biol.***, 425(8):1318-1329.

DOI: 10.1016/j.jmb.2013.01.021

Contributions: cloning, protein expression and purification, structure determination, structure analysis, protein-protein interaction experiments, enzymatic activity experiments

[#]denotes co-first authors

TABLE OF CONTENTS

INTRODUCTION

Eukaryotic cells segregate their genome from the rest of the cell by enclosing it in a double lipid bilayer – the nucleus. In doing so, cells protect genetic information from potential damage. One of the most fundamental consequences of this is the spatial separation of the transcription and post-transcriptional mRNA processing in the nucleus from the ribosomal translational machinery in the cytoplasm. In addition, cells gain powerful mechanisms to regulate gene expression by controlling which factors are allowed into the nucleus or actively removed from it. However, both benefits require a mechanism to traffic macromolecules in or out of the nucleus with high fidelity: mRNAs cannot be exported out of the nucleus before they are fully mature; similarly, only the proteins the cell needs in the nucleus should be able to access it. This problem is magnified by the scale of the macromolecules that require passage, which range from small folded proteins to entire mRNAs coated with dozens or more proteins.

The nuclear pore complex, also known as the NPC, serves as the gateway to the nucleus (Hoelz et al., 2011). The NPC facilitates traffic in and out of the nucleus by generating pores in the nuclear envelope large enough to conduct entire mRNAs or ribosomal particles (Pante and Kann, 2002). Unlike most other membrane transporters, the NPC does not open and close to regulate nucleocytoplasmic transport. To regulate transport, the NPC generates a passive diffusion barrier that prevents most macromolecules from freely diffusing through the transport channel. To cross this barrier, macromolecules require the assistance of dedicated transport factors.

Most proteins and some small RNAs are transported by a family of proteins called karyopherins (reviewed in Cook et al., 2007)) In canonical transport, karyopherins recognize potential cargoes through short sequences. Karyopherins are able to cross the diffusion

barrier and when bound to a cargo they shuttle the cargo across with them. The directionality of transport is ensured by the small GTPase protein called Ran, which adopts different conformations in the GTP and GDP bound state. Because the GTPase activating protein for Ran, RanGAP, is localized to the cytoplasmic side of the NPC and the guanine exchange factor for Ran, RCC1, is localized exclusively to the nucleus, Ran exists in the GDP-bound state in the cytoplasm and the GTP-bound state in the nucleus. RanGTP disassembles import complexes when they reach the nucleus, causing the release of import cargoes into the nucleus. In contrast, RanGTP is required for the formation of export complexes, but once the karyopherin and cargo reach the cytoplasmic side, RanGAP activates hydrolysis of GTP to GDP, causing the release of the export cargo into the cytoplasm. However, compared to the extensive structural characterization of karyopherin-mediated transport, much less is known about the architecture of the NPCs through which cargoes must pass.

NPCs were first observed by electron microscopy in the 1950s (Callan and Tomlin, 1950; Watson, 1959). Most of our understanding of the gross features of the NPCs have come from continued improvement and application of electron microscopy on a variety of model systems. To a rough approximation, NPCs are toroidal in shape, and in humans their dimensions are approximately 700 Å in height, 1200 Å in diameter, with a 400 Å transport channel in the center (Bui et al., 2013) (Figure 1). Each NPC generates a pore by stabilizing the fusion of the inner and outer nuclear membranes. Thus NPCs are massive macromolecular machines directly embedded in the nuclear envelope.

Each NPC contains approximately 1000 subunits, totaling to an estimated molecular weight of 120 MDa. However, there are only 34 unique proteins that make up the NPC, depending on the species (Figure 2). To assemble such a large complex, the NPC utilizes extensive symmetry, so that there are multiple copies of all the unique components in each NPC. From early electron microscopy studies, it was apparent that the NPC possesses

eightfold rotational symmetry around the central transport channel (Gall, 1967). The core components of the NPC are additionally symmetric across the plane of the nuclear envelope, meaning that they are present on both the nuclear and cytoplasmic sides. This core is referred to as the symmetric core (Figure 1). There are also cytoplasmic- and nuclear-specific decorations that break this two-fold symmetry but still possess eightfold symmetry, which are called the asymmetric nucleoporins – specifically the cytoplasmic filaments and nuclear basket.

While the gross architectural features of the NPC can be deciphered from electron microscopy, a mechanistic understanding how the NPC accomplishes its many functions while gatekeeping the nucleus requires atomic-resolution insight. This has proved challenging despite even the latest technological advances. Because it is a membrane embedded protein complex, it is difficult (if not impossible) to isolate an NPC without perturbing its structure. Because of its tremendous size and flexibility, the NPC has also presented challenges for structural analysis *in situ*. Structure determination methods such as electron microscopy require averaging and the conformational heterogeneity of individual NPCs limits the resolution of the resulting structures. Despite these challenges, tremendous progress has been made in electron tomographic reconstructions of the NPC, with the best available reconstructions approaching 20 Å in resolution (Eibauer et al., 2015; Kosinski et al., 2016). However, these reconstructions do not permit *de novo* modeling of the structure of the NPC. Instead they provide an envelope into which high-resolution models of the individual proteins can be assembled.

The work presented in this thesis utilizes a complementary, bottom-up approach to determining the structure of the NPC. High-resolution biochemical and structural analyses performed with recombinant proteins provide not only the building blocks to assemble the structure of the NPC, but also reveal the rules and driving forces that hold the NPC together.

This analysis builds on a decade of structural work done by André Hoelz and others, which itself was built on decades of genetic and biochemical work to identify the proteins that make up the NPC, which are called nucleoporins (reviewed in Hoelz et al., 2011).

The nucleoporins themselves present many challenges that must be overcome. A significant challenge for structure determination techniques is that large parts of the nucleoporins are not structured. This feature is in fact critical for NPC function, because the diffusion barrier itself is generated by intrinsically disordered sequences (often stretches of hundreds of amino acids) called FG repeats. FG repeats are enriched in motifs containing phenylalanine and glycine residues, usually embedded in polar sequences depleted of charge. FG repeats have a propensity for self-assembly and have been observed to form a separate phase *in vitro* (Frey et al., 2006). Many nucleoporins possess a stretch of FG repeats and the assembly of these sequences in the central transport channel creates the passive diffusion barrier.

There are also nucleoporins which possess large stretches of unstructured sequences that are not FG repeats. These sequences also present a challenge to structural determination, but the work presented here as well as studies by other groups has revealed that these sequences are critical for holding the NPC together (Fischer et al., 2015; Lin et al., 2016; Stuwe et al., 2015a). Besides these unstructured domains, the remaining mostly nucleoporins contain large, structured domains. Often these domains are composed of structural repeats: α-helices assembling into α-helical solenoids or β-strands that form β-propellers. These nucleoporins are the structural building blocks of the NPC, but a primary challenge has been purifying the recombinant proteins for high-resolution structure determination. Since most human nucleoporins have not been successfully reconstituted, fungal nucleoporins have been the predominant source of structural information.

The first three chapters of this thesis present work towards understanding the structure of the symmetric core of the nuclear pore complex using fungal proteins. The symmetric core forms three concentric rings: an inner ring that surrounds the central transport channel and cytoplasmic and nuclear outer rings that sit above the nuclear envelope on either side of the NPC. The nucleoporins that make up these rings in the symmetric core are divided into three major complexes: (1) the coat nucleoporin complex, (2) the adaptor nucleoporin complex, and (3) the channel nucleoporin trimer. The coat nucleoporin complex (CNC) is the primary component of both the cytoplasmic and nuclear outer rings. Early structural insights into NPC architecture came primarily from x-ray crystal structures of components of the CNC, either individually or in complexes with each other (Debler et al., 2008; Hsia et al., 2007; Nagy et al., 2009; Seo et al., 2009). In contrast, very little was known about the atomic structure of components of either the adaptor nucleoporin complex or channel nucleoporin trimer.

The first chapter describes the 7.4 Å-resolution x-ray crystal structure of a ~400 kDa hexameric CNC subcomplex from *S. cerevisiae* and exemplifies the bottom-up approach (Stuwe et al., 2015b). The crystal structure both revealed the relative arrangement of the subunits and provided the first structural information on the central interaction hub that holds the complex together. While the crystal structure was relatively low resolution, there were high-resolution structures most of the individual components. In determining the structure, the high-resolution structures were essentially docked into a low-resolution envelope. As a result, high-resolution information such as the position of specific residues and the associated known biochemistry is contextualized in the overall structure of the entire complex. These insights were confirmed by a 4.1 Å crystal structure solved by Thomas Schwartz's group using a smaller fragment from a different fungus, *Myceliophthora thermophile* (Kelley et al., 2015).

The second chapter describes the x-ray crystal structure of a ~110 kDa fragment of a nucleoporin in the adaptor nucleoporin complex, called Nup192 from the thermophilic fungus *C. thermophilum* (Stuwe et al., 2014). Most structural work on the CNC was performed using *S. cerevisiae* nucleoporins, primarily because the proteins could be expressed and purified recombinantly. A similar approach has not been technically feasible for *S. cerevisiae* nucleoporins in the adaptor nucleoporin complex. However, tremendous progress has been made since Ed Hurt's group demonstrated that nucleoporins from *C. thermophilum* are biochemically more stable, finally making adaptor nucleoporin complex components accessible for structural and biochemical studies (Amlacher et al., 2011). The crystal structure of the *C. thermophilum* Nup192^{NTD} is the first crystal structure of a *C. thermophilum* nucleoporin. Using the structure, we designed a large panel of surface mutants and successfully mapped the binding site of one of its interaction partners, Nup53. This approach served as a prototype for our extensive analysis of the entire interaction network of the adaptor nucleoporin complex.

In the third chapter, the reconstitution of every component of the symmetric core from *C. thermophilum* is described (Lin et al., 2016). We reconstituted complexes containing nearly every symmetric core component and systematically identified the interaction partners for each nucleoporin. Further dissection of the minimal interaction regions provided a high-resolution biochemical interaction map, which revealed that unstructured sequences were the primary driver of adaptor nucleoporin complex assembly. Using *C. thermophilum* proteins, we captured three of these interactions in high-resolution crystal structures (Nup170^{NTD}•Nup53, Nup170^{CTD}•Nup145N, Nic96•Nup53). We also solved crystal structures of large fragments of Nup170 and Nup192, which we used as templates to assemble structures of the full-length proteins. At this point, we had acquired accurate models

for virtually every protein in the symmetric core built from high-resolution x-ray crystal structures.

At the same time, higher resolution cryoelectron tomographic reconstructions of the human NPC were being reported by Martin Beck's group (von Appen et al., 2015). With a nearly complete inventory of crystal structures, it was therefore possible to build up the structure of the entire symmetric core. Because of the limited resolutions of the maps (at best 23 Å), placement of even very large crystal structures remains far from unambiguous, especially because many nucleoporins share common features and shapes. However, by iteratively placing the largest and most distinctive structures first and removing the corresponding density from the subsequent search, we could assign the position and orientation of nearly the entire symmetric core. The resulting composite structure defines the stoichiometry of most nucleoporins, reveals interaction surfaces that are critical for NPC assembly, identifies the primary sites of contact with the nuclear envelope, and provides additional insight into the molecular organization of the diffusion barrier. Importantly, the composite structure was confirmed by a structure determined independently by Martin Beck's group (Kosinski et al., 2016).

The next two chapters focus on understanding the function of nucleoporins in the cytoplasmic-specific asymmetric nucleoporins, primarily using human proteins. While the symmetric core is the scaffold that holds the NPC together and generates the diffusion barrier, the asymmetric nucleoporins interact extensively with the nucleocytoplasmic transport machinery. For example, RanGAP, the GTPase activating protein for Ran, is localized to the NPC by the vertebrate-specific nucleoporin Nup358. The fourth chapter presents a structure-function analysis of the very C-terminal domain of human Nup358, which displays a high sequence similarity to the peptidyl-prolyl-isomerase cyclophilin A (Lin et al., 2013). We determined the x-ray crystal structure of the $Nup358^{CTD}$ and found that this domain also

possess peptidyl-prolyl-isomerase activity despite an altered active site. We showed that the active site also mediates binding to the HIV-1 CA protein, which was later confirmed by a crystal structure of the complex reported by Leo James' group (Bichel et al., 2013).

Finally, the fifth chapter presents a structure-function analysis of cytoplasmic-specific nucleoporins involved in the terminal step of mRNA export. mRNAs are exported from the nucleus through a pathway distinct from karyopherin-mediated transport. Instead they are recognized by a heterodimeric transport factor called Nxf1•Nxt1, which similar to karyopherins is capable of shuttling bound mRNAs across the diffusion barrier. The directionality of this pathway is ensured by spatial activation of the DEAD-box helicase DDX19 by the cytoplasmic-specific asymmetric nucleoporins Gle1, Nup42, and Nup214, leading to removal of Nxf1•Nxt1 from mRNAs (Stewart, 2010). We discovered that Nup42 binding to Gle1 is a critical determinant of Gle1 stability, allowing us to reconstitute human DDX19 activation. Crystal structures and ATPase assays provided new insights into the mechanism of DDX19 activation by Gle1, revealing significant differences from the fungal system. Lastly, mapping of human disease mutations in Gle1 uncovered a strong link between Gle1 stability and human disease. Thus by building on our understanding of the structure of the symmetric core, we have acquired important insights into how the NPC participates in nucleocytoplasmic transport and its role in human disease.

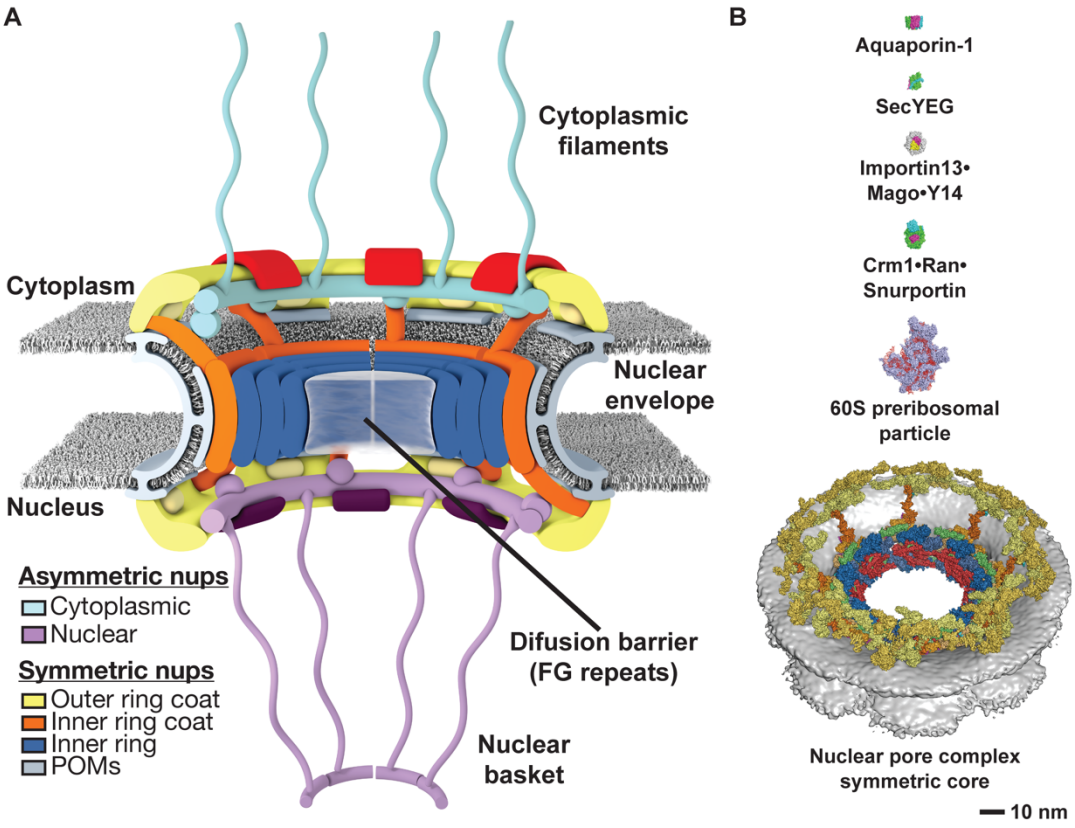


Fig. 1.
Overall architecture of the nuclear pore complex. (A) Cartoon schematic of the NPC. (B) Size comparison of other membrane transport channels and NPC transport cargoes to the structure of the NPC.

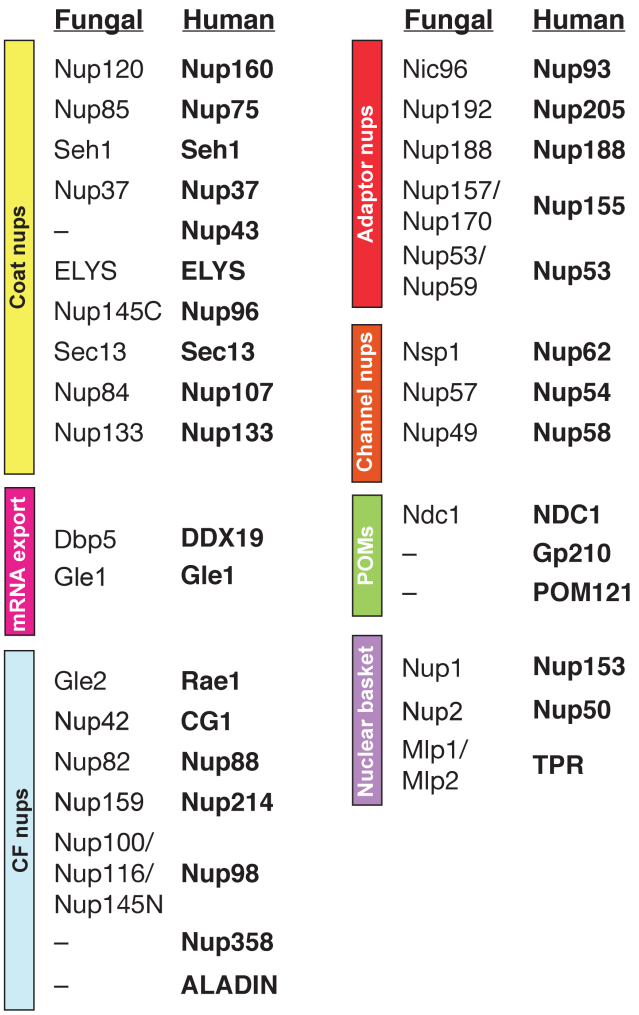


Fig. 2.
Nucleoporins are highly conserved. Nucleoporins are organized by subcomplexes. The fungal homologues are shown on the left and the human homologues are shown on the right.

REFERENCES

Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D.P., Arumugam, M., Bork, P., and Hurt, E. (2011). Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. Cell *146*, 277-289.

Bichel, K., Price, A.J., Schaller, T., Towers, G.J., Freund, S.M., and James, L.C. (2013). HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. Retrovirology *10*, 81.

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

Callan, H.G., and Tomlin, S.G. (1950). Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. Proc R Soc Lond B Biol Sci *137*, 367-378.

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

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

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

Fischer, J., Teimer, R., Amlacher, S., Kunze, R., and Hurt, E. (2015). Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. Nat Struct Mol Biol *22*, 774-781.

Frey, S., Richter, R.P., and Gorlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science *314*, 815-817.

Gall, J.G. (1967). Octagonal nuclear pores. J Cell Biol *32*, 391-399.

Hoelz, A., Debler, E.W., and Blobel, G. (2011). The structure of the nuclear pore complex. Annu Rev Biochem *80*, 613-643.

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

Kelley, K., Knockenhauer, K.E., Kabachinski, G., and Schwartz, T.U. (2015). Atomic structure of the Y complex of the nuclear pore. Nat Struct Mol Biol *22*, 425-431.

Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A.L., Wan, W., Bui, K.H., Hagen, W.J., Briggs, J.A., Glavy, J.S.*, et al.* (2016). Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science *352*, 363-365.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N.*, et al.* (2016). Architecture of the symmetric core of the nuclear pore. Science *352*, aaf1015.

Lin, D.H., Zimmermann, S., Stuwe, T., Stuwe, E., and Hoelz, A. (2013). Structural and functional analysis of the C-terminal domain of Nup358/RanBP2. J Mol Biol *425*, 1318-1329.

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

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

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

Stewart, M. (2010). Nuclear export of mRNA. Trends Biochem Sci *35*, 609-617.

Stuwe, T., Bley, C.J., Thierbach, K., Petrovic, S., Schilbach, S., Mayo, D.J., Perriches, T., Rundlet, E.J., Jeon, Y.E., Collins, L.N.*, et al.* (2015a). Architecture of the fungal nuclear pore inner ring complex. Science *350*, 56-64.

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

Stuwe, T., Lin, D.H., Collins, L.N., Hurt, E., and Hoelz, A. (2014). Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. Proc Natl Acad Sci U S A *111*, 2530-2535.

von Appen, A., Kosinski, J., Sparks, L., Ori, A., DiGuilio, A.L., Vollmer, B., Mackmull, M.T., Banterle, N., Parca, L., Kastritis, P.*, et al.* (2015). In situ structural analysis of the human nuclear pore complex. Nature *526*, 140-143.

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

CHAPTER 1

ARCHITECTURE OF THE NUCLEAR PORE COMPLEX COAT

This chapter was adapted from:

Tobias Stuwe[#], Ana R. Correia[#], **Daniel H. Lin**, Marcin Paduch, Vincent T. Lu, Anthony A. Kossiakoff, André Hoelz (2015). Architecture of the nuclear pore complex coat, ***Science***, 347(6226):1148-1152.

ABSTRACT

The nuclear pore complex (NPC) constitutes the sole gateway for bidirectional nucleocytoplasmic transport. Despite half a century of structural characterization, the architecture of the NPC remains unknown. Here, we present the crystal structure of a reconstituted ~400 kDa coat nucleoporin complex (CNC) from *S. cerevisiae* at a 7.4-Å resolution. The crystal structure revealed a curved Y- shaped architecture and the molecular details of the coat nucleoporin interactions forming the central "triskelion" of the Y. A structural comparison of the yeast CNC with an electron microscopy reconstruction of its human counterpart suggested the evolutionary conservation of the elucidated architecture. Moreover, thirty-two copies of the CNC crystal structure docked readily into a cryoelectron tomographic reconstruction of the fully-assembled human NPC, thereby accounting for ~16 MDa of its mass.

RESULTS

The nuclear pore complex (NPC) is composed of ~34 different proteins, termed nucleoporins (nups), that assemble in numerous copies to yield a ~120 MDa transport channel embedded in the nuclear envelope (NE) (1). To facilitate the extensive membrane curvature generated in each NE pore, NPCs require a membrane-bending coat. The NPC coat is believed to be formed by an evolutionarily conserved coat nup complex (CNC), the Nup107/160 complex in humans and the Nup84 complex in *S. cerevisiae*, the latter of which is composed of Nup120, Sec13, Nup145C, Seh1, Nup85, Nup84, and Nup133 (1, 2).

We reconstituted a hetero-hexameric CNC containing the yeast nups Nup120, Sec13, Nup145C, Seh1, Nup85, and the Nup84 N-terminal domain (NTD) (Fig. 1, A and B). Our reconstituted CNC did not include Nup133 because this nup is conformationally flexible and loosely associated (2–4). Because the initial crystals of this reconstituted CNC diffracted poorly, we generated a series of conformation-specific, high-affinity synthetic antibodies(sABs) and tested them as crystallization chaperones (5). This approach yielded crystals of the CNC in complex with sAB-57, which allowed us to solve the structure to 7.4 Å by molecular replacement using high-resolution crystal structures of CNC components and the sAB scaffold (fig. S1 and S2) (6–10). The inclusion of a second sAB (sAB-87) produced another crystal form, for which we collected anomalous X-ray diffraction data of Seleno-L-methionine and heavy metal-labeled crystals to confirm the placement of the CNC components (fig. S1, S2 and S3). Because the coat nups in both CNC•sAB complexes adopted the same arrangement, we focused our analysis on the better ordered CNC•sAB-57 structure (fig. S4, S5, and S6).

The CNC adopted a curved Y-shaped structure spanning ~250 Å in length and width, consistent with previous negative-stain electron microscopy (EM) analyses (Fig. 1C) (2–4,

11). The Seh1•Nup85 pair and Nup120 constituted the upper arms of the Y, which were connected to the rest of the CNC through a central triskelion. Sec13•Nup145C•Nup84^{NTD} formed the stalk at the bottom of the triskelion and would attach the tail formed by Nup84^{CTD} and Nup133, which were absent in the structure. Both arms curved out such that the Nup120 β-propeller domain was perpendicular to the plane of the Y. Nup145C organized the CNC through four distinct interaction surfaces contacting nearly every member of the complex. sAB-57 bound at the Nup145C-Nup85 interface and formed crystal packing contacts (Fig. 2 and fig. S4).

The C-terminal domains (CTDs) of Nup145C (residues 553–712), Nup85 (residues 545– 744), and Nup120 (residues 729–1037) converged to form the CNC triskelion. While we observed clear electron density that revealed the connectivity of the three CTDs and their interactions (Fig. 2 and fig. S2), the sequence register in the triskelion was only approximate due to the absence of side chain density. Nup120^{CTD} was sandwiched between Nup85^{CTD} and Nup145C^{CTD} and no direct contacts were observed between Nup85^{CTD} and Nup145C^{CTD} (Fig. 2, A and B). The interactions between Nup85^{CTD}, Nup145C^{CTD}, and Nup120^{CTD} were mediated predominantly by their most C-terminal helices. An additional interaction was made by an N-terminal Nup145C helix bound to a groove in the Nup85^{CTD} surface ~60 Å away from the triskelion center, an interaction that was recognized by sAB-57 (Fig. 2C).

Consistent with our structural data, we reconstituted a stoichiometric complex between Nup120 and Nup85^{CTD} as monitored by size-exclusion chromatography interaction experiments (fig. S7A). Furthermore Nup120 failed to interact with Sec13•Nup145C in the absence of Nup145C^{CTD} (fig. S7, B and C). The interaction between Seh1•Nup85 and Sec13•Nup145C depended on the presence of an N-terminal Nup145C fragment (residues

75–125) (fig. S7, D and E). Further mapping identified a region of Nup145C (residues 75–109) that was sufficient for Nup85^{CTD} binding (fig. S7F), confirming that this fragment bridged the two subcomplexes. Nup120^{CTD} and Nup85^{CTD} were essential for the formation of the CNC (fig. S8 and S9). These data were consistent with published CNC cross-linking data of three different species, more so than the models generated by coarse-grained analysis (3, 11, 12). Lastly, to validate their placement in the structures, we confirmed the interactions of sAB-57 and sAB-87 with Seh1•Nup85•Nup145C$^{1-123}$ and Nup120^{NTD}, respectively (fig. S10). Next, we compared the CNC structure to previously determined EM reconstructions of the yeast and human CNCs (3, 4). While the EM reconstruction of the yeast CNC recapitulates its overall shape, significant deviations were apparent (fig. S11). No density is observed in the EM structure for the U-shaped tip of Nup85. The overall shape of the crystal structure was also consistent with the human CNC EM reconstruction, which contains the crystallized evolutionarily conserved core as well as two additional human components, Nup37 and Nup43 (Fig. 3A). Our crystal structure did not account for additional EM density directly adjacent to the Seh1•Nup85 and Nup120 arms, which reportedly accommodate Nup43 and Nup37, respectively (Fig. 3A) (3). In the human CNC, Nup43 appears to bind to the same site as sAB-57 on the Seh1•Nup85 arm of the Y. The major difference between the CNC crystal structure and both EM reconstructions is the curvature of the arms of the Y, and thus the orientation of the Nup120 β-propeller was substantially different in the crystal structure (Fig. 3A). The flatness of both EM reconstructions suggests that these deviations may be a result of EM sample preparation. Despite this, the degree of similarity between the yeast CNC crystal structure and the human CNC EM reconstruction suggested substantial evolutionary conservation of the CNC architecture.

The higher-order arrangement of CNCs in a fully assembled NPC has been debated, with several models proposed based on various structural, biophysical, or computational

approaches (3, 7, 13, 14). Given the evolutionary conservation of the CNC architecture, we tested whether our crystal structure could be docked into the ~32-Å resolution tomographic reconstruction of an intact human NPC (3). Indeed, an unbiased 6-dimensional search combined with a cross-correlation analysis confidently docked 32 copies of the CNC crystal structure in the tomographic reconstruction, yielding a model for the NPC coat (Figs. 3B, 4A and S12). These results agreed with the stoichiometry and approximate localization previously proposed based on crosslinking mass spectrometry and the docking of the human CNC EM reconstruction (3). However, the crystal structure fit the tomographic reconstruction substantially better than the human CNC EM reconstruction (Fig. 3B).

The NPC coat was formed by 32 copies of the CNC arranged in four eight-membered rings (Fig. 4, A and B). The eight CNCs in each ring were oriented horizontally with their long axis positioned parallel to the surface of the NE in a head-to-tail fashion. On each side of the NE a pair of inner and outer CNC rings emerged up to ~210 Å (Fig. 4A). These rings were separated by a ~280 Å gap, yielding a total height of ~700 Å. The diameters of the outer CNC rings were slightly larger than those of the inner CNC rings, spanning ~1,200 Å and ~1,050 Å, respectively (Fig. 4B). While the CNCs in both rings were arranged with the same directionality, each CNC from the outer ring was offset from its mate in the inner ring by ~120 Å in a clockwise direction (Fig. 4B). Moreover, the tandem CNC rings on the nuclear and cytoplasmic side of the NE possessed the same handedness and were related by two-fold rotational symmetry (Fig. 4A and S12D).

The unambiguous placement and orientation of the coat nups and their conserved surfaces allowed for an investigation into the details of their interactions in the assembled NPC coat. Each CNC was situated on top of the NE and was oriented such that the plane of the Y was nearly perpendicular to the membrane, with the Nup120 and Seh1•Nup85 arms

pointed at or away from the membrane, respectively (Fig. 4A). Only two interfaces appeared to be responsible for oligomerization of individual CNCs into the NPC coat. The inner and outer rings only interacted where the top of the triskelion of each inner ring CNC met the bottom of the Nup84-Nup145C interface of each outer ring CNC (Fig. 4C). Nup120 was oriented such that its Nup133 binding site was directly adjacent to the density assigned to the N- terminus of Nup133, consistent with our previous findings that this interaction is responsible for CNC ring formation and critical for NPC assembly (Fig. 4D) (9). This Nup120 orientation also pointed the apex of its β-propeller directly towards the NE, which was the only membrane contact that we observed in our model (Fig. 4D). This region of the Nup120 β-propeller domain also contains a conserved surface patch on its side (9) that may serve as a NE anchor point for the entire NPC coat, either through a direct interaction with the membrane or via a membrane-anchored nup, as previously reported (15).

The NPC coat architecture is dissimilar to other structurally characterized membrane coats. Whereas the latter are generated by homotypic vertex elements (16, 17), the NPC coat is formed by heterotypic interactions of its asymmetric CNC protomers. Given the location of the CNC rings above and below the NE, other nups likely play a role in generating the complex curvature of the NE pores. While the placement of the CNCs in the NPC coat did not directly address the organization of the central transport channel (fig. S13), it accounted for ~16 MDa of the total mass of the NPC, bridged the resolution gap between low- resolution EM analyses and high-resolution crystallographic studies, and suggested the evolutionarily conservation of its architecture.

ACKNOWLEDGMENTS

We thank C.J. Bley, W.M. Clemons, A.M. Davenport, O. Dreesen, A. Patke, D.C. Rees, S.O. Shan, P. Stravropoulos, and K. Thierbach for critical reading of the manuscript, K. Kato and K. Kato for their contributions at the initial stages of this project, L.N. Collins for technical support, D. King for mass spectrometry analysis, E. Hurt and P. Loppnau for material, S. Koide for providing the phage display library, S. Gräslund for the pSFV4 vector, P. Afonine for advice regarding structure refinement in PHENIX, and J. Kaiser and the scientific staff of SSRL Beamline 12-2 and the APS Beamline GM/CA-CAT for their support with X-ray diffraction measurements. We acknowledge the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program for their support of the Molecular Observatory at the California Institute of Technology. The operations at the SSRL and APS are supported by the Department of Energy and the National Institutes of Health. T.S. was supported by a Postdoctoral Fellowship of the Deutsche Forschungsgemeinschaft. D.H.L. was supported by a National Institutes of Health Research Service Award (5 T32 GM07616). A.A.K. was supported by National Institutes of Health Awards (U01 GM094588, U54 GM087519) and the Chicago Biomedical Consortium. A.H. was supported by Caltech startup funds, the Albert Wyrick V Scholar Award of the V Foundation for Cancer Research, the 54th Mallinckrodt Scholar Award of the Edward Mallinckrodt, Jr. Foundation, and a Kimmel Scholar Award of the Sidney Kimmel Foundation for Cancer Research. The coordinates and structure factors have been deposited with the Protein Data Bank with accession codes 4XMM and 4XMN. The authors declare no financial conflicts of interest.

MATERIALS AND METHODS

Protein expression and purification. DNA fragments encoding full-length Sec13, Nup145C (residues 74–712), full-length Seh1, and Nup85 (44-744) were PCR amplified and cloned as pairs into the pET-Duet1 expression vector (Novagen). Nup145C and Nup85 were cloned into the first multiple cloning site using BamHI and NotI restriction sites, whereas Sec13 and Seh1 were cloned into the second multiple cloning site using NdeI and XhoI restriction sites. Nup84 (residues 1-451) was cloned into a modified pET28a vector, which contains an N-terminal hexahistidine tag followed by a PreScission protease cleavage site, using NdeI and NotI restriction sites (*18*). The expression construct for Nup120 was described previously (*2*). The selected synthetic antibody (sAB) fragments of sAB-57 and sAB-87 were cloned into the pSFV4 vector (Peter Loppnau, Structural Genomics Consortium, University of Toronto) using the restriction sites NcoI and SalI, and subsequently digested using SalI and BsaI and religated to obtain the C-terminal hexahistidine tag. The details of the bacterial expression constructs are listed in table S1.

All proteins were expressed in *Escherichia coli* BL21-Codon-Plus (DE3)-RIL cells (Stratagene) in Luria-Bertani media. Seleno-L-methionine-labeled (SeMet) Sec13•Nup145C, Seh1•Nup85, and Nup84^{NTD} were produced in a synthetic medium that suppresses methionine biosynthesis, following standard protocols. For all nucleoporins, expression was induced at an OD600 of 0.8 with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG), followed by growth at 18 °C for 18 hours. Cells were harvested by centrifugation and resuspended in a buffer containing 20 mM TRIS (pH 8.0), 500 mM NaCl, 20 mM imidazole, 4 mM β-mercaptoethanol (β-ME), and complete EDTA-free protease inhibitor mixture (Roche).

For purification, cells were lysed with a cell disruptor (Avestin) and DNase I (Roche) was added to the lysate before centrifugation at 30,000 × g for 1 hour. The supernatant was filtered through a 0.45-μm filter (Millipore) and loaded onto a nickel-nitrilotriacetic acid (Ni-NTA) column (Qiagen) equilibrated in 20 mM TRIS (pH 8.0), 500 mM NaCl, 20 mM imidazole, and 4 mM β-ME. Protein was eluted with a linear gradient of 20 mM TRIS (pH 8.0), 500 mM NaCl, 500 mM imidazole, and 4 mM β-ME. Protein-containing fractions were pooled, incubated with either PreScission (GE Healthcare) or ULP1 protease, and dialyzed overnight at 4 °C against a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT. Next the protein was loaded onto a Mono Q 10/100 GL ion-exchange column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT and eluted using a NaCl gradient. Protein-containing fractions were concentrated in a centrifugal filter (Millipore) and loaded onto a HiLoad Superdex 200 16/60 gel filtration column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT. Protein-containing fractions were pooled and concentrated to 20 mg/mL for biochemical interaction experiments and CNC reconstitution.

sABs were expressed, harvested, and lysed in a similar fashion, but the cells were induced at an OD600 of 0.9 with 0.25 mM IPTG and grown at 25 °C for 18 hours. After lysis, the lysate was incubated at 65 °C for 30 minutes and then cooled on ice for 15 minutes before centrifugation. Protein-containing fractions from the Ni-NTA affinity purification were pooled and loaded on to a 5 mL HiTrap MabSelect SuRe column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 500 mM NaCl, 20 mM imidazole, and 4 mM β-ME. The protein was eluted using a linear gradient of 0-100 % elution buffer, containing 0.1 M sodium citrate (pH 3.2). To rapidly increase the pH after

elution, the fractions were collected into tubes containing 200 µl of 1 M TRIS (pH 9.0). The eluted fractions were dialysed against a buffer containing 20 mM TRIS (pH 8.0) and 100 mM NaCl and concentrated to 10 mg/ml for biochemical interaction experiments and crystallization.

Reconstitution of CNC complexes. Seh1•Nup85•Nup120 (Trimer1) was purified by co-lysis of cells expressing Seh1•Nup85 or Nup120, following the protocol described above. For the reconstitution of Sec13•Nup145C•Nup84^{NTD} (Trimer2), purified Sec13•Nup145C was mixed with a 1.2 fold molar excess of Nup84^{NTD}, incubated on ice for 30 minutes, and loaded onto a HiLoad Superdex 200 16/60 gel filtration column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT. Trimer1 and Trimer2 containing fractions were pooled, concentrated, mixed with a 1.2 molar excess of Trimer2, incubated on ice for 1 hour, and injected onto a HiLoad Superdex 200 16/60 gel filtration column equilibrated in the same buffer. Fractions containing the reconstituted CNC were pooled and concentrated to 10 mg/ml for sAB interaction experiments. For SeMet labeled CNC, SeMet-labeled Seh1•Nup85, Sec13•Nup145C, and Nup84^{NTD} were purified and used instead of the native proteins. SeMet labeling of Nup120 rendered the protein insoluble in our bacterial expression system and thus native Nup120 was used for the reconstitution of the SeMet-labeled CNC. For the generation of CNC•sAB complexes, native or SeMet-labeled CNC were mixed with 1.5 fold molar excess of sAB-57 or a 1:1 mixture of sAB-57 and sAB-87 and loaded onto a HiLoad Superdex 200 16/60 gel filtration column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT. DTT was included in the buffer as it was necessary for CNC stability and had

no effect on the integrity of the sABs. Fractions containing the various CNC complexes were pooled and concentrated to 10 mg/mL for crystallization.

sAB selection and characterization. The generation and screening of conformation-specific sABs has been described previously (*5*). Briefly, a modified yeast CNC was reconstituted with a $Nup84^{NTD}$ variant that harbored an N-terminal avi-tag. The complex was biotinylated in a 2 mL reaction by incubating 40 μM protein with a buffer containing 50 mM BICINE (pH 8.3), 100 M biotin, 10 mM ATP, 10 mM magnesium acetate, and 30 μg biotin ligase (BirA) at 30 °C for 2 hours. After labeling, protein was buffer exchanged using a 5 mL HiTrap Desalting column (GE Healthcare) equilibrated with a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT and purified again using a HiLoad Superdex 200 16/60 gel filtration column equilibrated in the same buffer. The extent of $Nup84^{NTD}$ biotinylation and efficiency of capture were tested by incubating 25 μg of protein with 50 μL of Streptavidin MagneSphere particles (Thermo Scientific), washing once with 50 μL of a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT, and resolving the bound proteins on a SDS-PAGE gel. Four rounds of competitive selection were performed using 100 nM (round 1), 50 nM (round 2), 10 nM (round 3), and 10 nM (round 4) biotinylated protein target and a phage display library according to previously published protocols (*5*). In case of sAB-57 biotinylated yeast CNC was used and to eliminate sABs that recognized unassociated CNC components, 1 μM of non-biotinylated CNC subunits (Seh1•Nup85, $Nup120^{NTD}$, Sec13•Nup145C, and $Nup84^{NTD}$) were used as competitors in all solutions during the last three rounds of selection. Phages were preincubated with competitors for 1 hour at room temperature. sAB-87 was obtained in a selection where biotinylated $Nup120^{NTD}$ was used and no competition

was performed. After successful selection, the specificity of candidate sABs was tested against the assembled biotinylated yeast CNC, as well as individual biotinylated subunits using a single point competitive ELISA assay (*5*). Only sequence-unique sABs with the desired binding properties were nominated for further biochemical characterization. To evaluate the binding affinity and specificity of the selected sABs, 1.5-fold molar excess of sAB was incubated with the reconstituted CNC or individual CNC components and loaded onto a MonoQ 5/50 GL ion-exchange column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT and eluted using a NaCl gradient. Interacting sABs eluted with the CNC components, whereas non-interacting sABs eluted prior to the gradient step. Initially, only sABs that specifically interacted with the fully assembled CNC were systematically tested in crystallization trials. To improve the diffraction properties of the CNC•sAB-57 crystals, additional sABs with the ability to bind individual CNC components were systematically screened for crystal formation. The addition of sAB-87 to the CNC•sAB-57 complex yielded a new crystal form with distinct packing and different space group.

Protein crystallization, heavy metal derivatization and data collection. Protein crystallization was carried out at 21 °C in hanging drops consisting of 1.0 μL protein solution (2 mg/ml) and 1.0 μL reservoir solution. Crystals appeared in the monoclinic space group C2 with one copy of the CNC•sAB-57 complex in the asymmetric unit. The crystals were improved by microseeding, which resulted in crystals that grew as thin plates with maximum dimensions of ~30×300×300 μm^3 within 1 week. Crystals used for diffraction experiments were grown in 0.1 M MES, pH 6.7, 5 % (w/v) PEG 20000, and 3% (v/v) ethanol. Crystals were cryoprotected by gradually supplementing the drop with 36 % (v/v)

ethylene glycol (in 1 % steps, every 5 minutes) and flash frozen in liquid nitrogen. Crystals of the CNC changed morphology after the inclusion of a second sAB (sAB-87) and appeared in the orthorhombic space group $P2_12_12_1$ with one copy of the CNC•sAB-57•sAB-87 complex in the asymmetric unit. Crystals grew to maximum dimensions of ~50×100×150 μm^3 within 1 week. Crystals used for diffraction experiments were grown in 0.1 M MES, pH 6.5, 5 % (w/v) PEG 20000, and 20 mM SrCl. Crystal were cryoprotected by serial transfers into solutions containing 5 %, 10 %, 15 %, 20 % and 25 % (v/v) ethylene glycol supplemented reservoir solution. This crystal form typically diffracted to a resolution limit of ~9 Å. During a systematic heavy metal derivative screen individual SeMet labeled crystals were identified that yielded X-ray diffraction data to a resolution limit of 7.6 Å after soaking with 1 mM potassium hexachloroosmate (K_2OsCl_6). Native crystals were derivatized by adding 0.2 μL of a saturated tantalum bromide cluster (Ta_6Br_{14}) solution to the crystallization drop and incubated for 1 week. X-ray diffraction data were collected at 100 K at beamline BL12-2 at the Stanford Synchrotron Radiation Source (SSRL) and beamline GM/CA-CAT 23ID-D at the Advanced Photon Source (APS) on a Pilatus3 detector. Thousands of CNC crystals were screened to yield the reported X-ray diffraction datasets.

Structure determination and model building. X-ray diffraction data was processed with XDS (*19*). The structures for both crystal forms were solved by iterative cycles of molecular replacement (MR) using Phaser (*20*).

For structure determination of the first crystal form containing the CNC•sAB-57 complex in the space group C2, Phaser was run with the assumption that the asymmetric unit (ASU) harbored one CNC•sAB-57 complex (~450 kDa), corresponding to a solvent

content of ~83 %. The crystal structures of the *S. cerevisiae* CNC components Sec13•Nup145C (PDB ID 3IKO) Seh1•Nup85 (PDB ID 3F3F), Nup120^{NTD} (PDB ID 3F7F), Nup84^{NTD} (PDB ID 3IKO), and a structure of the sAB scaffold (PDB ID 3PGF) were used sequentially as search models with a model variance of 100 % sequence identity (*6-10*). MR was performed in the above search order and the top solutions were taken from each MR search to look for the next molecule. During each MR round, Phaser robustly obtained solutions with clear separation from other solutions after the packing test with Log Likelihood Gain (LLG) values and refined translation function Z-scores (TFZ) of: (1) LLG=43, TFZ=9.0 (Sec13•Nup145C), (2) LLG=94, TFZ=11.6 (Seh1•Nup85), (3) LLG=190, TFZ=6.9 (Nup120^{NTD}), (4) LLG=202, TFZ=8.1 (Nup84^{NTD}), and (5) LLG=310, TFZ=8.5 (sAB) (fig. S1A, B).

The correctness of the final solution output from Phaser was assessed on the following criteria: (1) clear separation of the best scoring solutions from the remaining solutions at every step, (2) very high TFZ scores after each step, as TFZ scores above 8 usually indicate a definite solution, (3) increasing LLG scores at each step indicating that each additional molecule was improving the solution, (4) an internal test that the Nup84^{NTD} was placed in the same orientation as previously determined in the Sec13•Nup145C•Nup84^{NTD} crystal structure (*8*), despite no *a priori* information restricting it to that location, (5) the overall shape of the solution was consistent with low resolution EM reconstructions, and most importantly (6) the appearance of strong additional features in the calculated electron density maps of the final solution. Most strikingly, strong positive difference density for the helices of the triskelion were clearly visible in the |Fo|-|Fc| map output from PHENIX (fig. S2A). Density modification of the MR solution using RESOLVE (*21*) yielded an improved electron density map with additional density for loops connecting the new helices despite no additional model building (fig. S2A). Furthermore,

no additional density was observed in the solvent channels (fig. S4A). Model building was performed with COOT (*22*). The α-helical C-terminal domains of Nup145C, Nup85, and Nup120 formed distinctive arrays of tubular electron density at 7.4 Å, into which we were able to place idealized α-helices. As the C-terminal domains of Nup145C, Nup85, and Nup120 are connected to their respective N-terminal domains by short loops, a preliminary model for the connectivity and directionality of the helices was traced starting from the C-terminus of each previously determined structure. This preliminary model was validated by comparison with the helical arrangement in the *S. pombe* homolog of Nup120, which could be structurally aligned with the helices assigned to the C-terminal domain of Nup120. Once all of the helices were successfully assigned to each protein, the connectivity of all three proteins could be assigned with the aid of the helix and loop lengths from a secondary structure prediction (fig. S3D). As the electron density does not possess features to assign the sequence register, the numbering in the structure is approximate and only reflects the order and directionality of each helix and thus we modeled the triskelion with the sidechains truncated at the Cβ position.

For structure determination of the second crystal form, containing the CNC•sAB-57•sAB-87 complex, which grew in the space group $P2_12_12_1$, sequential Phaser searches for the Sec13•Nup145C•Nup84^{NTD} heterotrimer (PDB ID 3IKO), Nup120^{NTD} (PDB ID 3F7F), and the sAB scaffold structure (PDB ID 3PGF) produced clearly separated solutions with the following scores: (1) LLG=239, TFZ=13.3 (Sec13•Nup145C•Nup84^{NTD}), (2) LLG=653, TFZ=12.2 (Nup120^{NTD}), and (3) LLG=887, TFZ=11.2 (sAB-87) (fig. S1C, D). Despite exhaustive attempts with both the normal Phaser pipeline and brute-force translation and rotation searches, no MR solutions were identified for Seh1•Nup85 (PDB ID 3F3F) and sAB-57 (PDB ID 3PGF) and these molecules are likely disordered in the

crystal. The resulting maps were comparable in quality to those of the C2 crystal form (fig. S2B). The arrangement of Sec13, Nup145C, Nup84^{NTD}, and the Nup120^{NTD} in the final MR solution is the same as that in the structure of the CNC•sAB-57 complex (fig. S6). The correctness of the solution was confirmed by the calculation of an anomalous difference Fourier map using the phases from the MR solution, which revealed peaks for 20 selenium sites and 1 Os site (fig. S3A-C). Improved phases were obtained with MR-SAD in Phaser using phases from the MR solution and the 21 anomalous scatterers. Subsequent density modification revealed clear tubular density for the triskelion helices, including density for the Nup85^{CTD}, which could be readily docked in the same conformations as observed in the CNC•sAB-57 structure. Additional confirmation of the correctness of the solution was obtained by calculating an anomalous difference Fourier map using anomalous X-ray diffraction data obtained from a CNC•sAB-57•sAB-87 complex crystal derivatized with Ta_6Br_{14}, which revealed 8 tantalum bromide cluster sites (fig. S3A-C).

Of the 20 selenium peaks observed, 11 aligned with the expected selenium sites in the previously determined structures of Nup84^{NTD}, Nup145C, and Sec13. An additional 8 selenium peaks were present in the newly built helices of Nup85^{CTD} and Nup145C^{CTD}, which were used to confirm the directionality and approximate sequence assignment of the helices. The final selenium site aligned with the last methionine present in Nup85^{NTD}, but no additional sites were observed for the remainder of the Seh1•Nup85^{NTD} heterodimer (fig. S3A). No additional electron density was visible for Seh1•Nup85^{NTD} in density modified maps either, despite room being available in the lattice to accommodate the molecules (fig. S4B). Thus, this part of the structure is presumed to be disordered in this crystal form.

Structure refinement. Refinement of both structures was performed with heavy restraints using PHENIX, with 1 group B-factor per residue with similarity restraints and positional refinement with secondary structure restraints and reference model restraints for the portions of the structure for which there were high-resolution structures (*23*). We elected to use models re-refined by the PDB_REDO server, as they had superior geometrical parameters to the previously deposited structures (*24*). The best strategy for B-factor refinement was determined by comparing the results of test refinements using the following strategies: 1 B-factor per residue with similarity restraints, 2 B-factors per residue with similarity restraints, 1 B-factor per group, and 1 B-factor per group with TLS parameters (fig. S5A, B). We additionally tested the output of a refinement strategy of 1 B-factor per residue without similarity restraints to ensure that B-factors were meaningfully restrained. Refinement with 1 B-factor per residue with similarity restraints yielded the lowest R-factors and realistic B-factors that were smoothly distributed across the model (fig. S5B). Therefore, we elected to use that strategy with no TLS parameters for the final refinement. The final models of the C2 and $P2_12_12_1$ crystal forms yielded average B factors for the overall model of 716.5 $Å^2$ and 536.5 $Å^2$, respectively, with comparable B factors for all protein chains (fig. S5D, E). These B-factors include the overall B of the crystal, as is the standard method of reporting B-factors in PHENIX. The resolution limits for both data sets were determined by using the paired refinement technique described by Karplus and Diederichs (*25*). Paired refinements were performed in 0.2 Å steps from 8.0 Å to 7.0 Å and the resolution limits were selected conservatively before resolution steps that did not improve the model (fig. S5F, G). The final models of the C2 and $P2_12_12_1$ crystal forms, refined to a 7.4-Å and a 7.6-Å resolution, yielded R_{free} and R_{work} values of 35.3 %, 33.0 %, and 34.7 %, 31.8 %, respectively. The stereochemical properties of the two structures were determined by MolProbity (*26*). The CNC complex structures reported here have similar

Ramachandran statistics as the search models used for Sec13•Nup145C•Nup84^{NTD}, Seh1•Nup85, and Nup120^{NTD}. The newly built triskelion has perfect stereochemical parameters with no residues in the disallowed region of the Ramachandran plot. For details of the data collection and refinement statistics see table S2.

Analytical size-exclusion chromatography. Protein-protein interaction experiments were carried out on a Superdex 200 10/300 GL gel filtration column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM NaCl, and 5 mM DTT. The various combinations of the yeast CNC components were mixed and incubated for 30 minutes on ice using a 1.2 molar excess of the smaller proteins. Complex formation was monitored by injection of the pre-incubated proteins or the individual components onto the gel filtration column. All proteins were analyzed under identical buffer conditions and complex formation was confirmed by SDS-PAGE of the protein-containing fractions, followed by Coomassie brilliant blue staining.

EM docking. The crystal structure of yeast CNC was docked into the negative stain EM reconstructions of the yeast complex (EMDB-5151) and the human complex (EMDB-2443) using the Fit in Map function in the UCSF Chimera software package (*27*).

The crystal structure of the yeast CNC was docked into the cryoelectron EM tomographic reconstruction of the human NPC (EMDB-2444) using an exhaustive, unbiased six-dimensional search using a Cα trace of the CNC structure with the program ESSENS from the Uppsala Software Factory package RAVE (*28*). The rotations were sampled in 10° steps across α, β, and γ for a total of 26,011 rotations, which were each tested at all of the 366,980 grid points which had a map value greater than 1.5. Each

combination of rotation and grid point was scored by the K-minimum sum function over the lowest scoring 60 % of the atoms against the average of the 8 nearest grid points, as implemented in ESSENS (*28*). This exhaustive scoring method produced a clear separation of 65 top scoring placements from the remaining orientations (fig. S12A). The positioning of the top 65 placements in the EM reconstruction was further refined and rescored using an orthogonal scoring method with the Fit in Map tool of UCSF Chimera (*27*), which we used to calculate the cross-correlation of the EM map with a simulated map calculated at 34 Å for each docked model (fig. S12B).

Analysis of these solutions and their placements revealed a clear separation between the top 32 solutions and the remainder of solutions. Because of the eight-fold rotational symmetry in the map, unique solutions are each composed of 8 solutions related by rotational symmetry. As a result, the top 32 solutions form four unique rings, all of which are compatible when simultaneously placed into the NPC. The remaining solutions could be classified as one of the following: (1) solutions that refined into one of the above orientations upon refinement, (2) solutions with moderate scores lower than the top scoring 32 orientations and could be discarded due to clashes (fig. S12C), or (3) low scoring solutions that yielded much worse fits than the top 32 solutions upon refinement. Despite the presence of additional features in the map for cytoplasmic filament nucleoporins and associated mRNA export factors, the two CNC rings on the cytoplasmic face and the two CNC rings on the nucleoplasmic face are identical (fig. S12D). This additional unbiased test was taken as final confirmation that this stoichiometry and orientation of CNCs reflects their organization in the NPC.

Illustration and figures. Structural figures were generated using PyMOL (www.pymol.org).

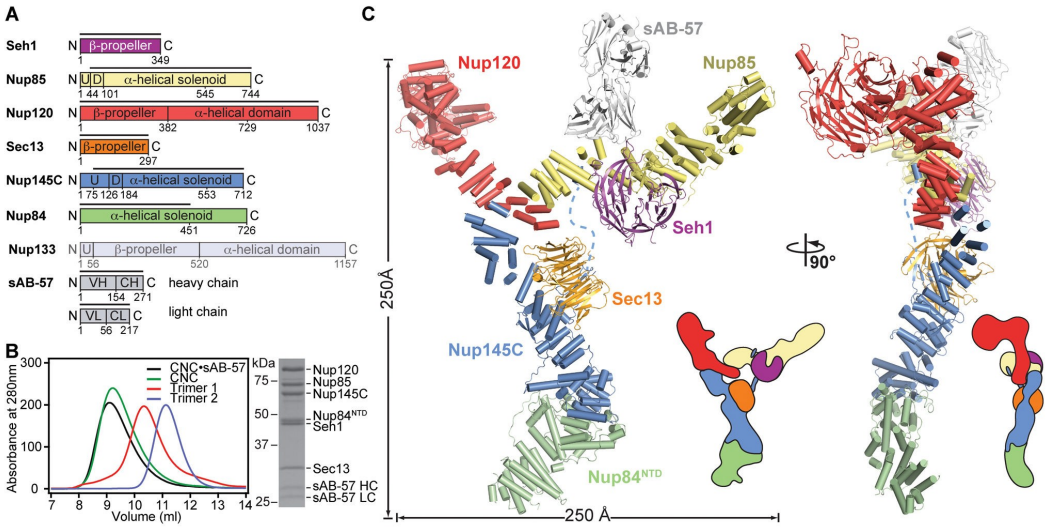


Fig. 1.

Overall architecture of the CNC. (**A**) Domain structures of the yeast coat nups and sAB-57. Black lines indicate the crystallized fragments. U: unstructured, D: domain invasion motif, VH: heavy chain variable region, CH: heavy chain constant region, VL: light chain variable region, CL: light chain constant region. (**B**) Reconstitution of the yeast CNC•sAB-57, lacking Nup133. Elution profiles from a Superdex 200 10/300 column are shown for Nup120•Seh1•Nup85 (Trimer 1), Sec13•Nup145C•Nup84^{NTD} (Trimer 2), CNC, and CNC•sAB-57 (left). SDS-PAGE gel of the reconstituted CNC•sAB-57 used for crystallization (right). (**C**) Cartoon and schematic representations of the yeast CNC•sAB-57 crystal structure viewed from two sides.

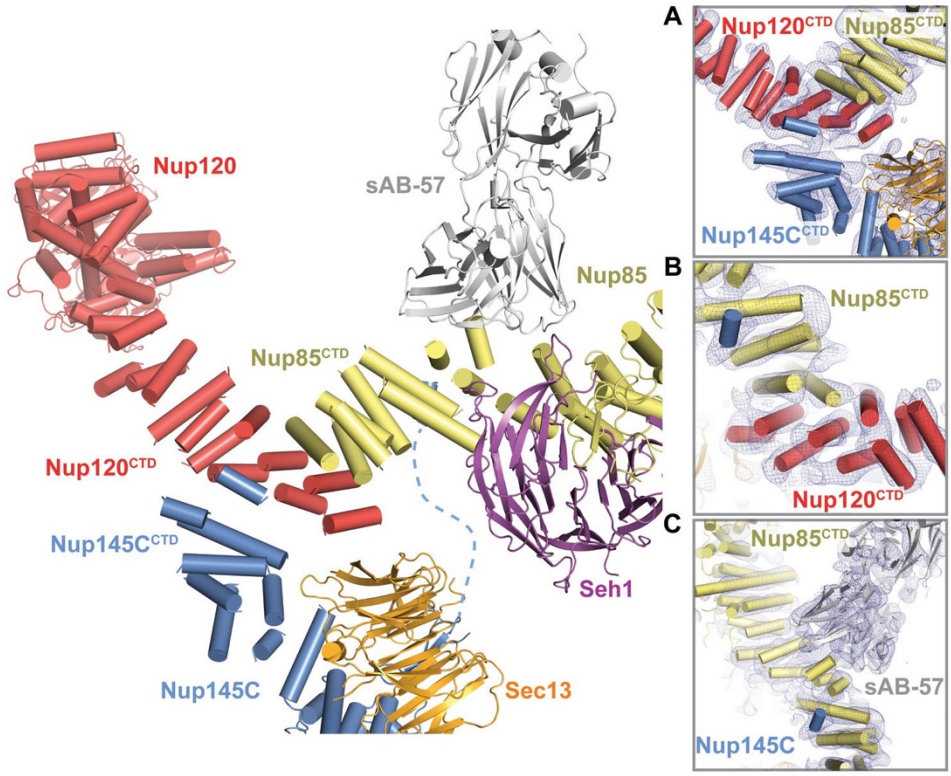


Fig. 2.

Architecture of the CNC triskelion. Cartoon representation of the triskelion formed by Nup120, Nup85 and Nup145C. Insets (**A–C**) depict magnified views for the interactions between (**A**) Nup120^{CTD}, Nup85^{CTD}, and Nup145C^{CTD} (**B**) Nup120^{CTD}, Nup85^{CTD}, and N-terminal Nup145C helix; and (**C**) Nup145C, Nup85^{CTD}, and sAB-57. The density modified electron density map is contoured at 1.0 σ.

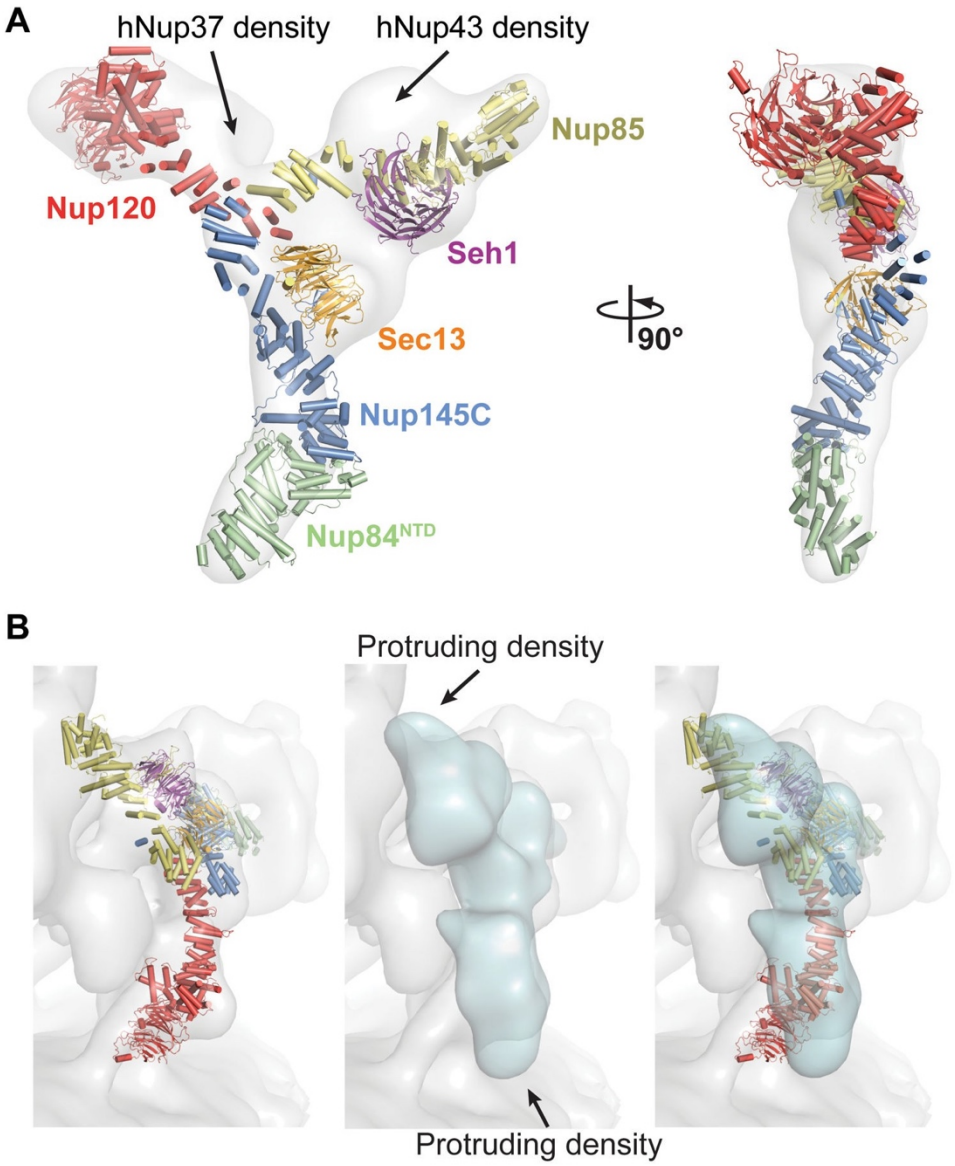


Fig. 3.

Comparison of yeast and human CNCs. (**A**) Fit of the yeast CNC crystal structure into the human CNC negative-stain EM reconstruction (gray) (3). Arrows indicate density accounted for by the additional human coat nups Nup37 or Nup43. (**B**) Comparison of the quality of fit for the yeast CNC crystal structure and human CNC EM reconstruction (cyan) into the intact human NPC cryoelectron tomographic reconstruction (gray) (3). Arrows indicate regions where the human CNC EM reconstruction protrudes from the cryoelectron tomographic reconstruction.

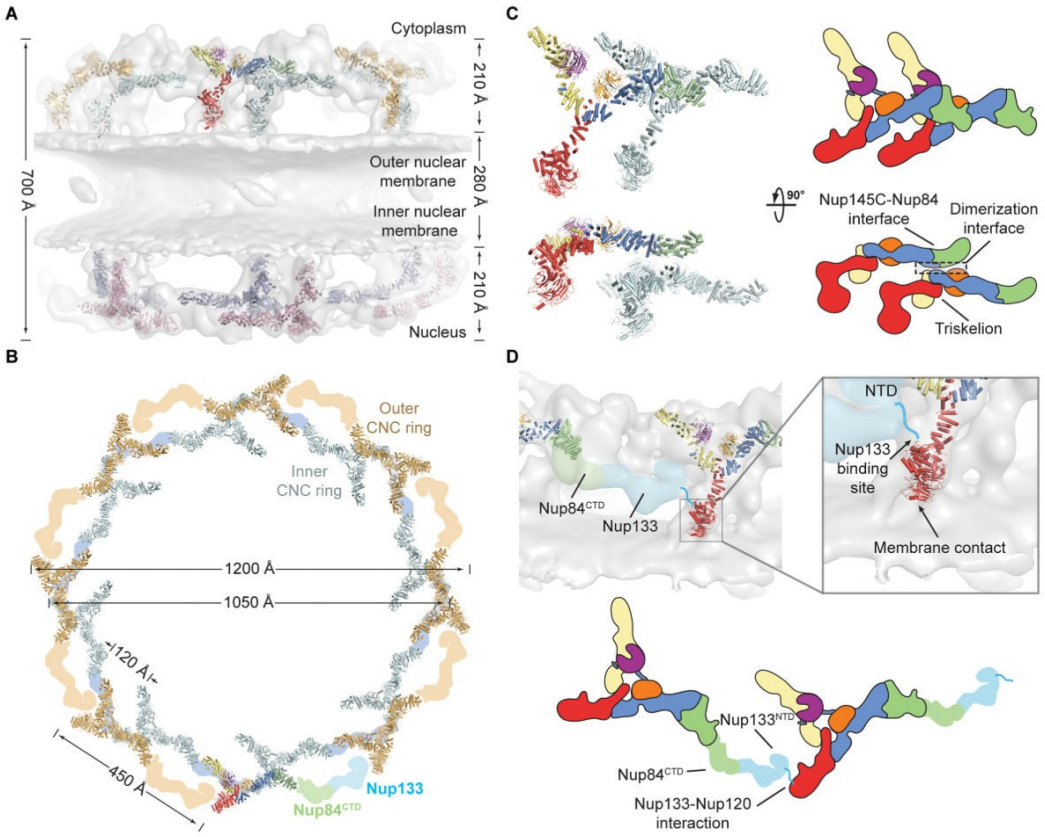


Fig. 4.

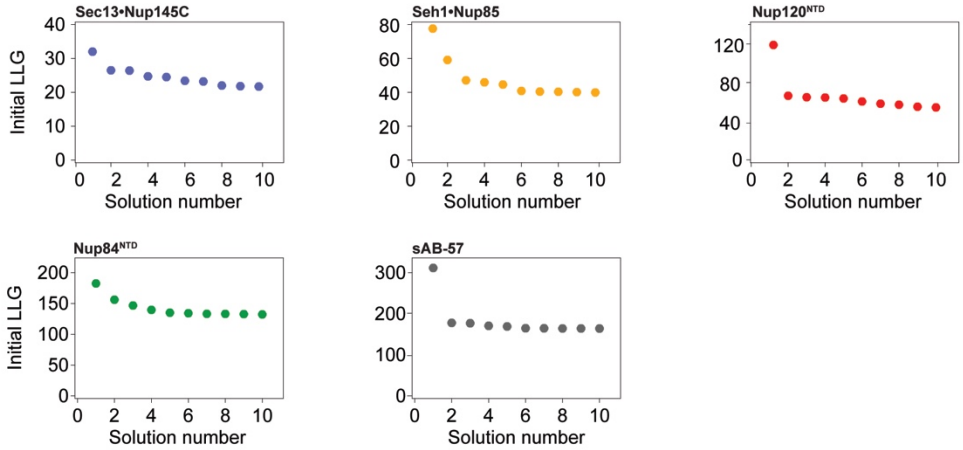
Architecture of the NPC coat. (**A**) 32 copies of the yeast CNC, shown in cartoon representation with a representative subunit colored as in Fig. 1, docked into the cryoelectron tomographic reconstruction of the intact human NPC (3), shown as a gray surface. The outer and inner cytoplasmic and nuclear CNC rings are highlighted in orange, cyan, pink, and blue, respectively. (**B**) Cartoon representations of 16 yeast CNC copies from the cytoplasmic side of the NPC coat. Schematics indicating the positions assigned to Nup84^{CTD} and Nup133, which were not crystallized, are shown. (**C**) Interface between the inner and outer CNC rings. Two views of the yeast CNC and its mate from the inner ring are shown. (**D**) Orientation of the Nup120 β- propeller relative to neighboring coat nups and the membrane. Portions of two CNCs from the cytoplasmic outer ring are shown in cartoon representation. Green and cyan shading indicate the positioning of Nup84^{CTD} and Nup133, respectively. The cyan line represents the N-terminal unstructured segment of Nup133 that binds to Nup120 (9). A schematic representation of the ring-forming Nup120-Nup133 interaction is shown below.

A

CNC•sAB-57
Space group C2

	Rotation function			Translation function					Refinement of solutions			
	solutions	RFZ Top solution	RFZ 2nd solution	solutions	solutions after 75% cutoff	TFZ top	TFZ 2nd	solutions after packing	LLG solution1	LLG solution2	Refined TFZ 1st	Refined TFZ 2nd
Sec13•Nup145C	11	4.3	3.9	219	6	4.9	3.4	2	43	33	9	x
Seh1•Nup85	118	4.35	4.1	37159	2	6.7	5.6	1	94	x	11.6	x
Nup120^{NTD}	27	4.3	4.2	7882	1	8.2	x	1	190	x	6.9	x
Nup84^{NTD}	79	4.4	4.1	18461	3	7.8	6.7	2	220	202	4.7	8.1
sAB-57	83	4.3	4	14652	1	13.3	x	1	310	x	8.5	x

B



C

CNC•sAB-57•sAB-87
Space group $P2_12_12_1$

	Rotation function			Translation function					Refinement of solutions			
	solutions	RFZ Top solution	RFZ 2nd solution	solutions	solutions after 75% cutoff	TFZ top	TFZ 2nd	solutions after packing	LLG solution1	LLG solution2	Refined TFZ 1st	Refined TFZ 2nd
Sec13•Nup145C• Nup84^{NTD}	1	6.9	x	2	2	15	8.3	2	239	102	13.3	x
Nup120^{NTD}	2	4.9	3.9	72	1	27.7	x	1	653	x	12.2	x
sAB-87	3	4.7	4	342	1	20.3	x	1	887	x	11.2	x
Seh1•Nup85	107	3.5	3.45	23320	461	7.3	7	0	x	x	x	x
sAB-57	17	4.5	4	2033	1	18.9	x	0	x	x	x	x

D

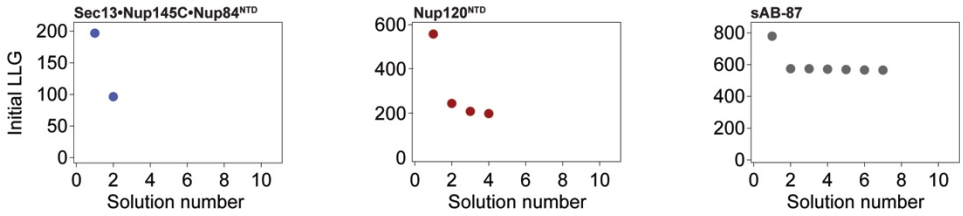


Fig. S1.

Structure determination statistics of the yeast CNC. (**A**) Table of statistics for each step of molecular replacement performed using Phaser for structure determination of the CNC•sAB-57 complex. Each sequential step was performed with the top solutions from the previous step, and a clear separation of the top solutions was apparent with each step. (**B**) Plots of the initial log-likelihood gain (LLG) scores for the top 10 peaks from the translation function step of Phaser for structure determination of the CNC•sAB-57 complex. (**C**) Table of statistics for each step of molecular replacement performed using Phaser for structure determination of the CNC•sAB-57•sAB-87 complex. Each sequential

step was performed with the top solutions from the previous step, and a clear separation of the top solutions was apparent with each step. (**D**) Plots of the initial log-likelihood gain (LLG) scores for the top peaks from the translation function step of Phaser for structure determination of the CNC•sAB-57•sAB-87 complex. For each step, there were only a handful of peaks selected from the rotation function, resulting in fewer peaks in the translation function.

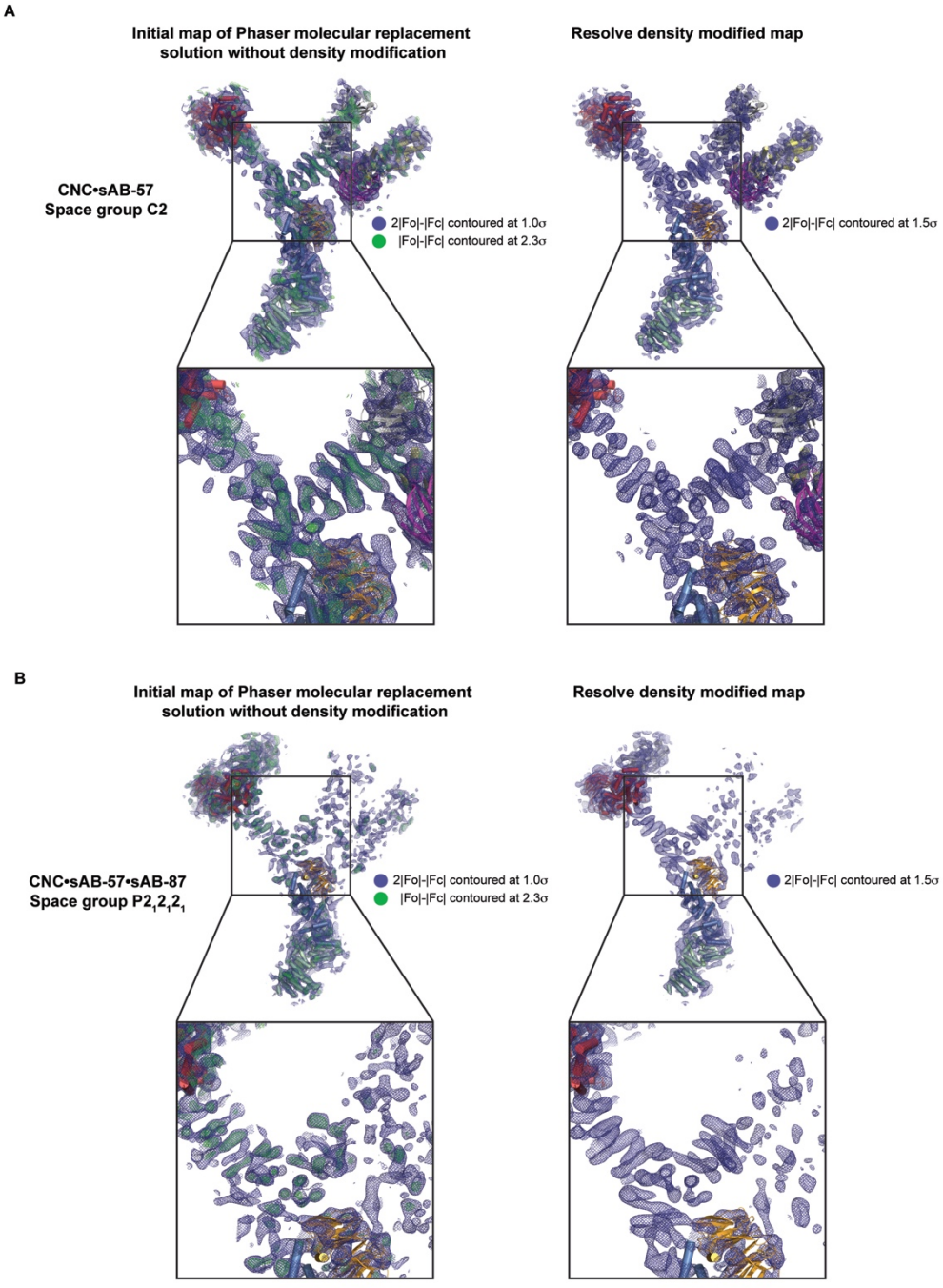


Fig. S2.

Electron density during structure determination of the yeast CNC. The electron density for the initial molecular replacement and after density modification for the crystal structures of the (**A**) CNC•sAB-57 and (**B**) CNC•sAB-57•sAB-87 complexes are shown. Clear density was visible for the triskelion helices after successful placement of previously solved crystal structures (left) and remained after density modification (right), which was performed prior to model building. The models visualized are the direct output from molecular replacement prior to any interpretation.

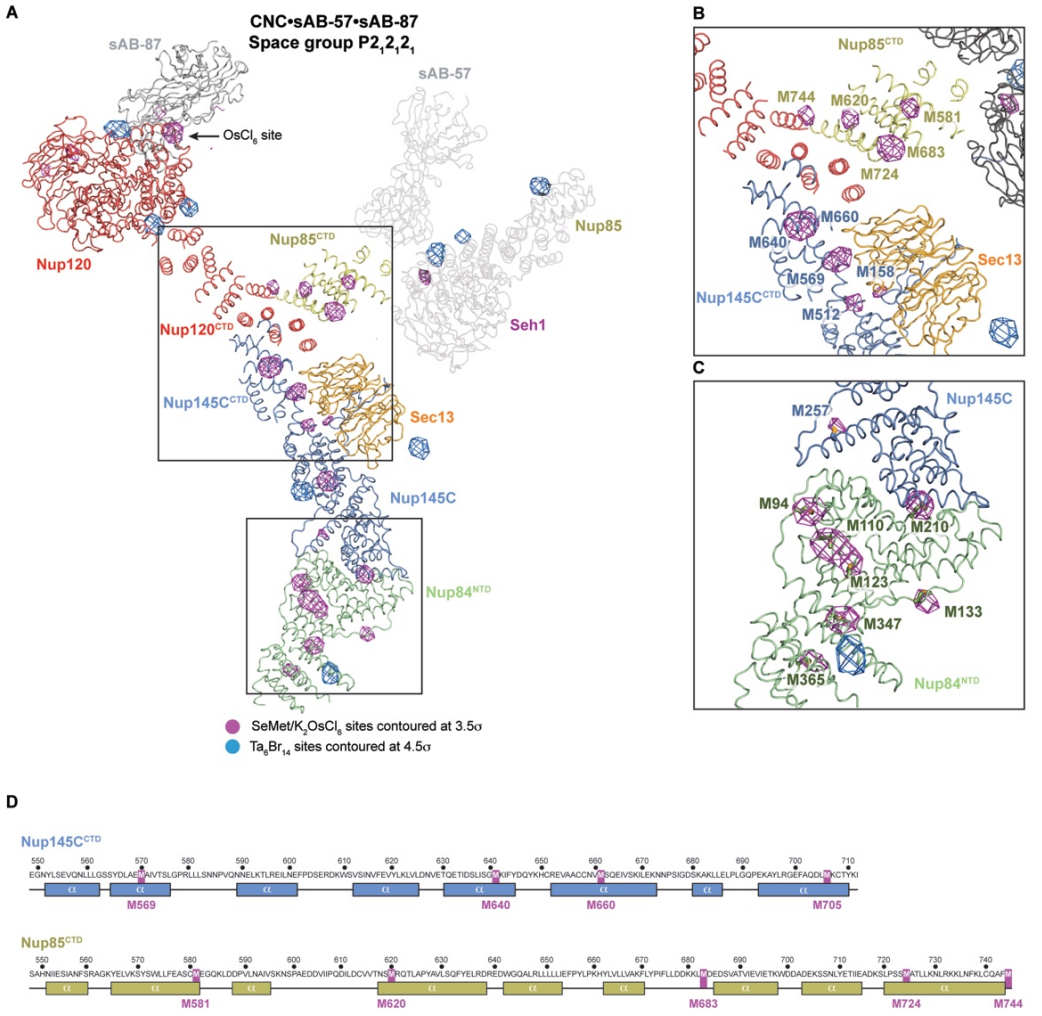


Fig. S3.

Location of anomalous scatterers in crystals of the CNC•sAB-57•sAB-87 complex. Anomalous difference Fourier maps were calculated for X-ray diffraction data collected at the selenium and tantalum peak wavelengths for CNC•sAB-57•sAB-87 crystals grown with SeMet labeled protein and soaked with K_2OsCl_6 or native CNC•sAB-57•sAB-87 crystals soaked with tantalum bromide clusters (Ta_6Br_{14}). (**A**) A ribbon representation of the structure of the CNC•sAB-57•sAB-87 complex, with the anomalous difference Fourier maps of X-ray diffraction data collected at the selenium (purple) or tantalum (blue) peak wavelengths contoured at 3.5 σ and 4.5 σ, respectively. Tantalum peaks adjacent to Nup85^{NTD} and a peak corresponding to the last selenium site in Nup85^{NTD} are visible despite the molecule being disordered in the crystal. (**B**) Close-up view of selenium sites present for the triskelion helices for which there are no high-resolution structures. 8 peaks are visible and confirm the positioning and orientation and approximate sequence assignment in the structure. (**C**) Close-up view of the selenium peaks in Nup84^{NTD} and Nup145C, with stick representations of the SeMet residues highlighting the expected sites. (**D**) Sequence and secondary structure prediction of Nup145C^{CTD} and Nup85^{CTD} with methionine residues highlighted.

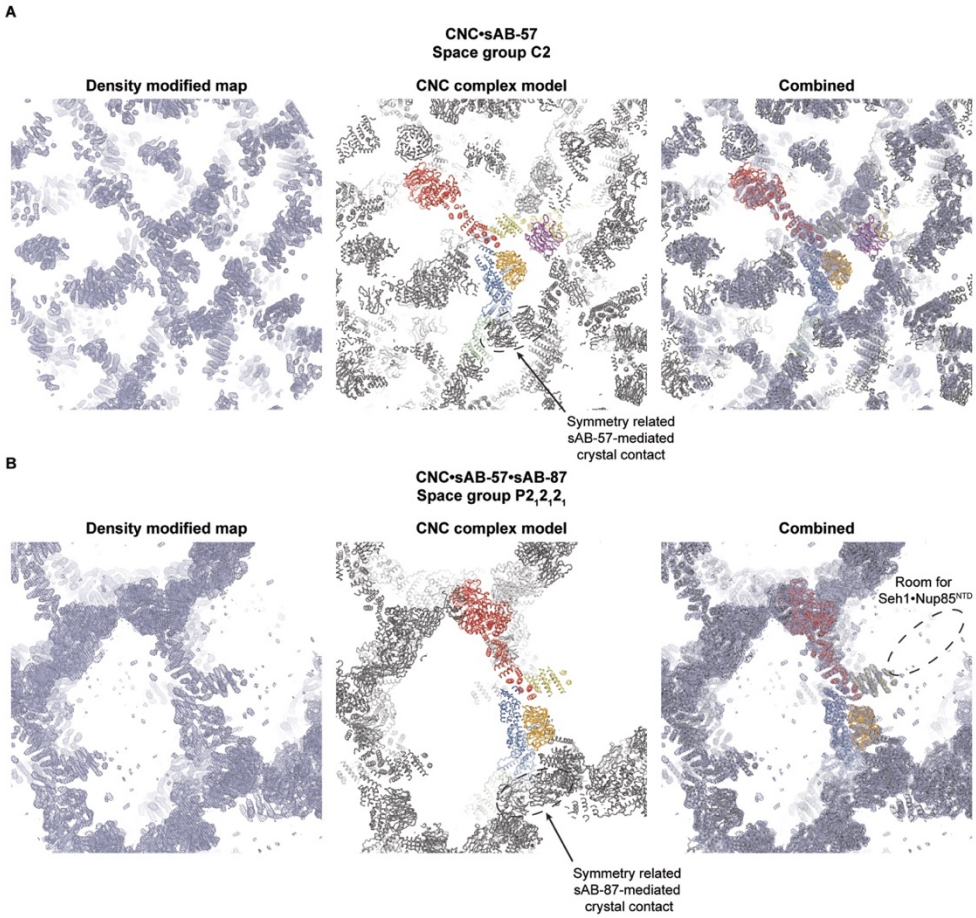
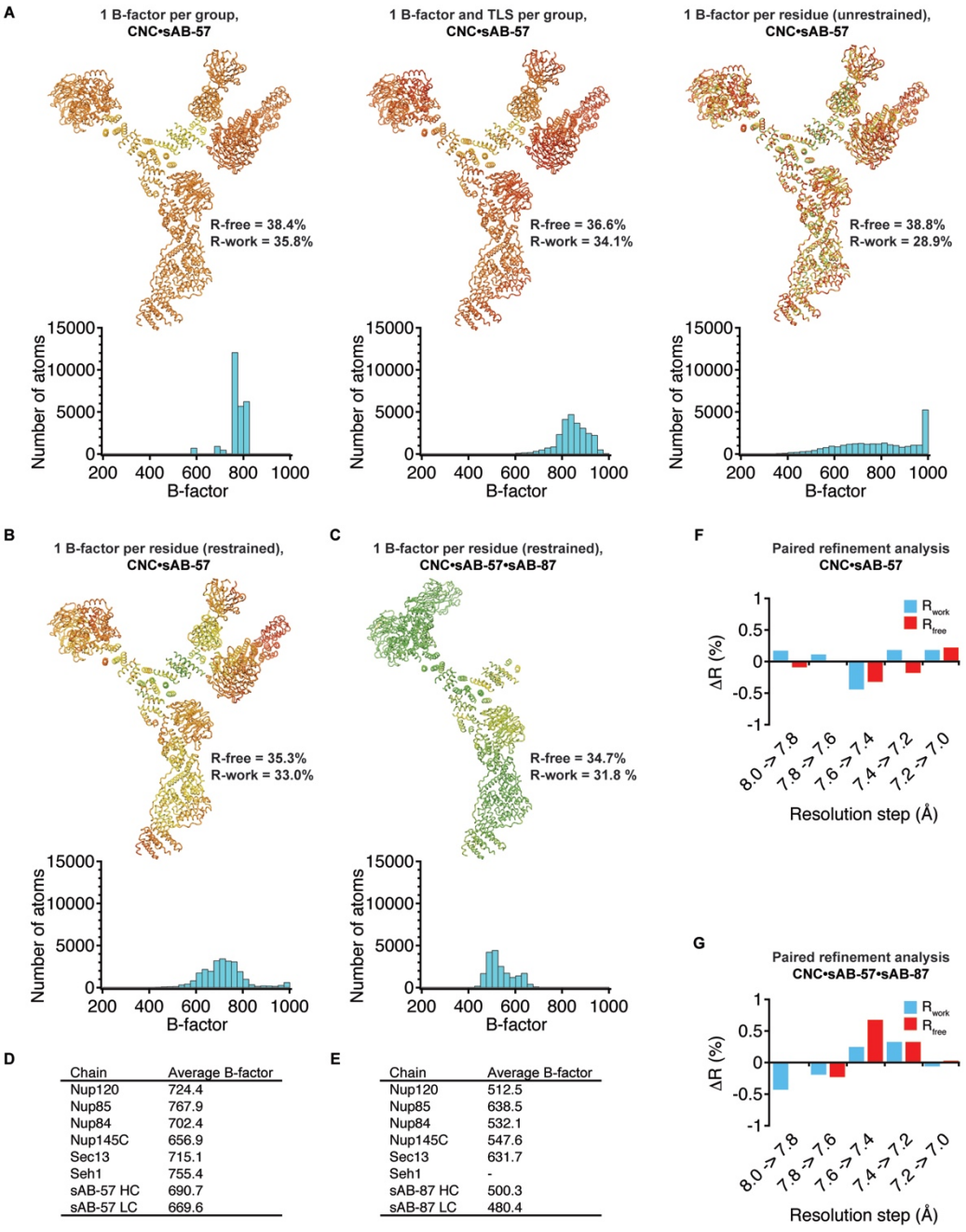


Fig. S4.
Crystal packing in crystals of CNC•sAB-57 and CNC•sAB-57•sAB-87. Representative views of the crystal packing for crystals of the CNC•sAB-57 complex (**A**) and CNC•sAB-57•sAB-87 complex (**B**). (left) Uncarved density-modified electron density contoured at 1 σ demonstrates the large solvent channels present in both crystals. (center) Ribbon representation of the asymmetric unit and surrounding symmetry mates colored gray highlights a major crystal contact made in both crystals by a synthetic antibody with the Nup145C•Nup84^{NTD} interface. (right) Combined view of both the electron density and unit cell.



Chain	Average B-factor
Nup120	724.4
Nup85	767.9
Nup84	702.4
Nup145C	656.9
Sec13	715.1
Seh1	755.4
sAB-57 HC	690.7
sAB-57 LC	669.6

Chain	Average B-factor
Nup120	512.5
Nup85	638.5
Nup84	532.1
Nup145C	547.6
Sec13	631.7
Seh1	-
sAB-87 HC	500.3
sAB-87 LC	480.4

Fig. S5.

Determination of optimal refinement strategy, resolution limits, and B-factor analysis. (**A**) Ribbon representations of the CNC•sAB-57 complex colored from blue to red by B-factor for various alternative refinement strategies with histograms of the B-factor distribution below. (**B**) Ribbon representation of the CNC•sAB-57 complex refined with the final refinement strategy of 1 B-factor per residue and colored on the same scale as in (A). (**C**) Ribbon representation of the CNC•sAB-57•sAB-87 complex refined with 1 B-

factor per residue and colored on the same scale as in (A). (**D**-**E**) Average B-factors per protein chain for the CNC•sAB-57 complex (**D**) or CNC•sAB-57•sAB-87 complex (**E**). (**F**-**G**). Paired refinement analysis of the resolution limit as described by Karplus and Diederichs (25) for the (**F**) CNC•sAB-57 complex or (**G**) CNC•sAB-57•sAB-87 complex. The improvement in R-factors gained for each 0.2 Å shell of data was assessed by re-calculating the R-factors in the lower resolution data after refinement with the higher resolution data.

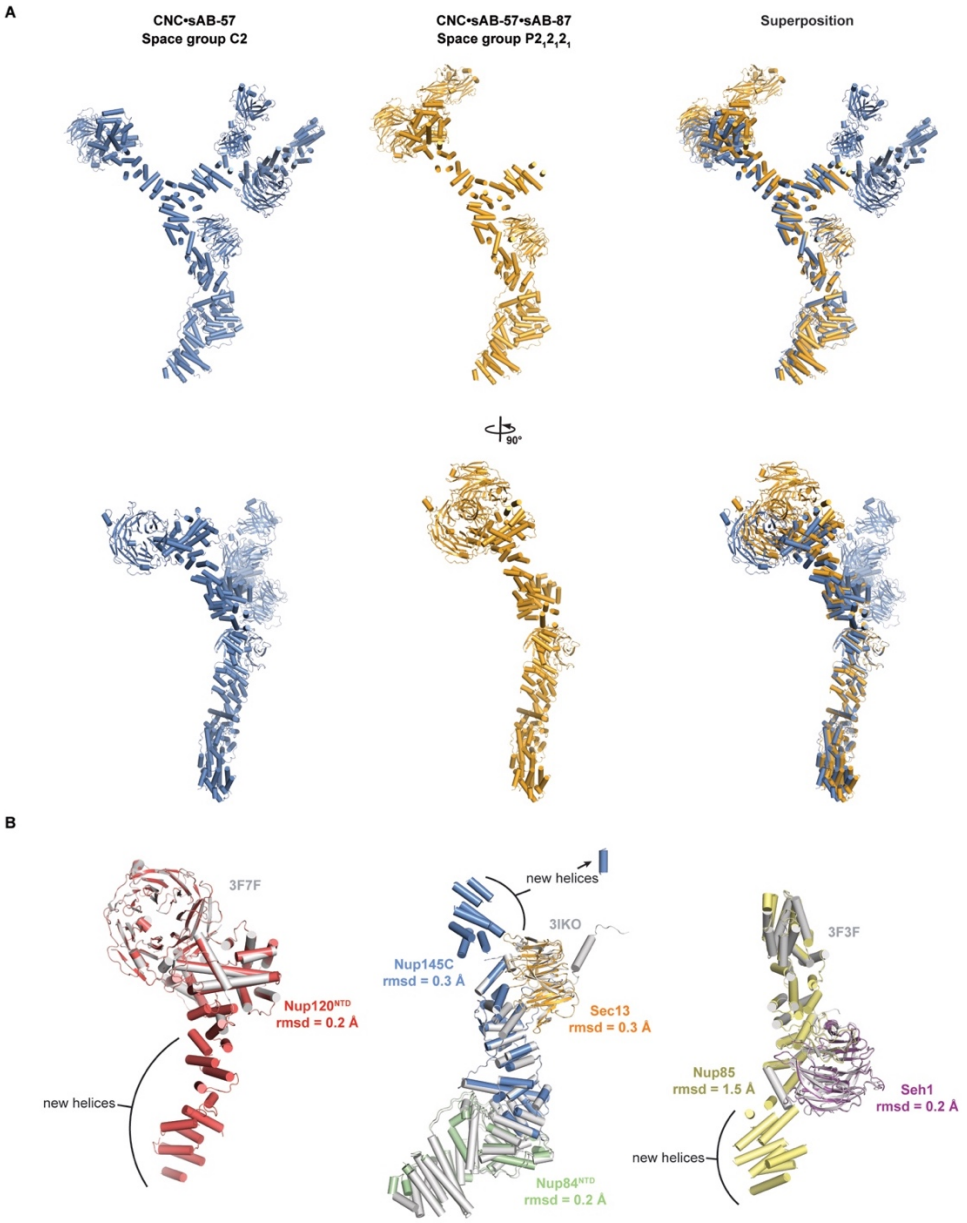


Fig. S6.

Comparison of the yeast CNC structures from different crystal forms. (**A**) The structures of the CNC•sAB-57 complex (blue) and CNC•sAB-57•sAB-87 complex (orange) are shown alone and superimposed over the central Sec13•Nup145C core. A view rotated by 90° is shown below. (**B**) Superposition of previously determined structures of the Nup120^{NTD}, the Sec13•Nup145C•Nup84^{NTD} hetero-trimer and the Seh1•Nup85 pair used for molecular replacement (PDB ID 3F7F, 3IKO, and 3F3F) (*6, 8, 9*) with their counterparts from the final crystallographic model. Cartoon representations are colored as in Fig. 1, whereas previous structures are colored in gray. Calculated root mean square displacements (rmsds) are indicated for each molecule. Nup85 has a large rmsd due to a large kink in the helical solenoid. Notably, crystal packing interactions of sAB-87 induce a slight rotation of the central triskelion.

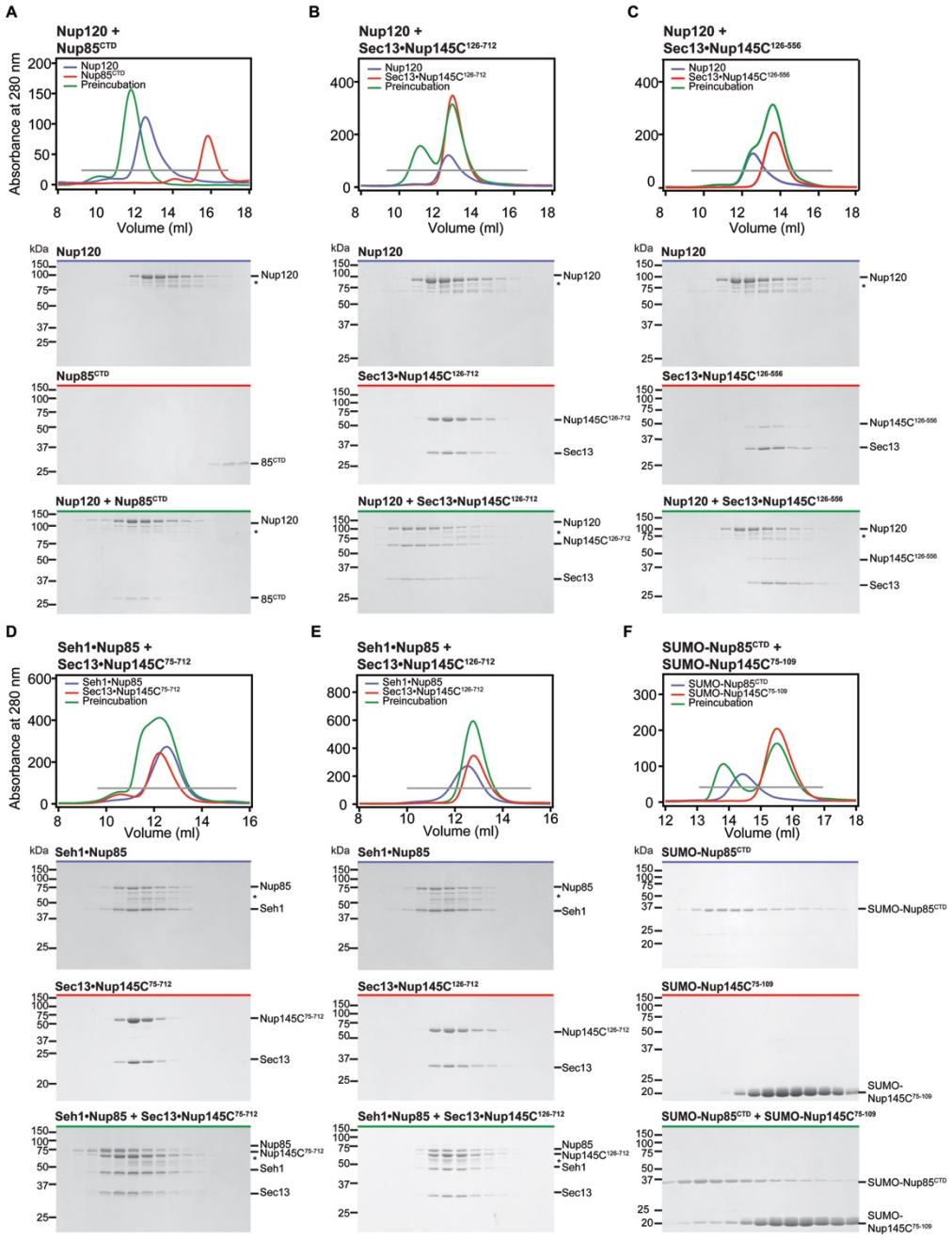


Fig. S7.

Biochemical characterization of the yeast CNC triskelion. (A-C) Nup120 interactions. Gel filtration profiles for Nup120 (A-C, blue), Nup85^{CTD} (**A**, red), Sec13•Nup145C$^{126-712}$ (**B**, red) and Sec13•Nup145C$^{126-556}$ (**C**, red) and after pre-incubation (green). **(D-E)** Seh1•Nup85 interactions. Gel filtration profiles for Seh1•Nup85 (D-E, blue), Sec13•Nup145C$^{75-712}$ (**D**, red), Sec13•Nup145C$^{126-712}$ (**E**, red) and after pre-incubation (green). **(F)** Interaction analysis between Nup85^{CTD} and Nup145C. Gel filtration profiles

for SUMO-Nup85^{CTD} (blue) and SUMO-Nup145C$^{75\text{-}109}$ alone (red) and after pre-incubation (green). Gray bars in the gel filtration profiles indicate the fractions resolved on the SDS-PAGE gels. Molecular mass standards and the positions of the proteins are indicated. Asterisks indicate degradation products. SDS-PAGE gels were stained with Coomassie brilliant blue. This size exclusion chromatography data is in agreement with previously published mass spectrometry data, which reported cross-links between Nup120^{K972} and Nup145C^{K672,K681,K694}, Nup120^{K943} and Nup145C^{K681}, and Nup120^{K972} and Nup85^{K772,K733,K734} (*12*).

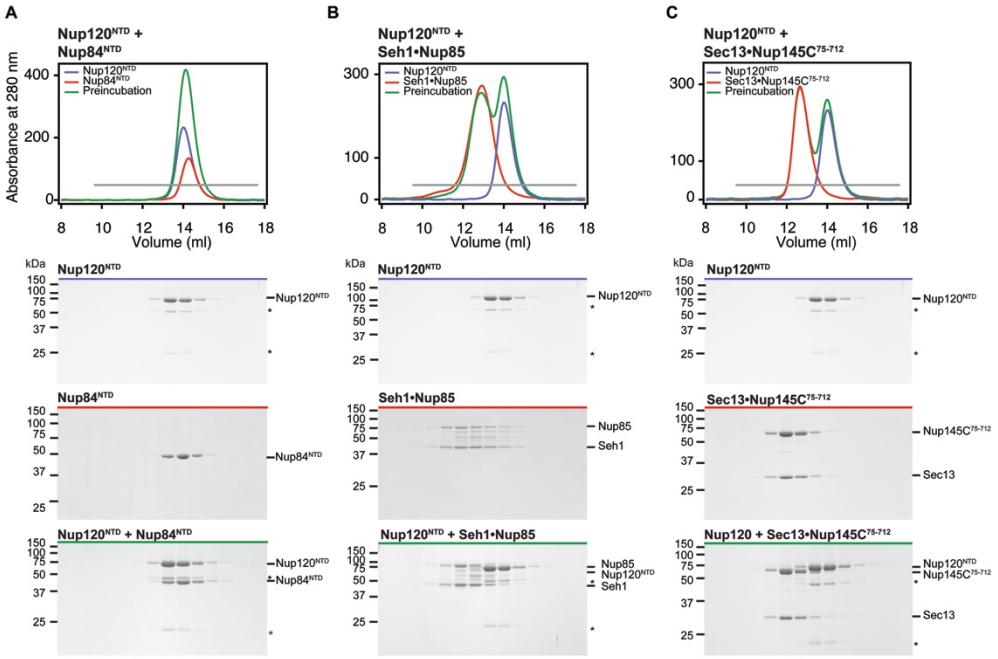


Fig. S8.
Nup120^{NTD} forms no interactions with Nup84^{NTD}, Seh1•Nup85, or Sec13•Nup145C$^{75-712}$. (A-C) Gel filtration profiles for the proteins alone: Nup120^{NTD} (A-C, blue), Nup84^{NTD} (**A**, red), Seh1•Nup85 (**B**, red) and Sec13•Nup145C$^{75-712}$ (**C**, red) and after pre-incubation of the different complexes (green). Gray bars in the gel filtration profiles indicate the fractions resolved on the SDS-PAGE gels. Molecular mass standards and the positions of the proteins are indicated. Asterisks indicate degradation products. SDS-PAGE gels were stained with Coomassie brilliant blue.

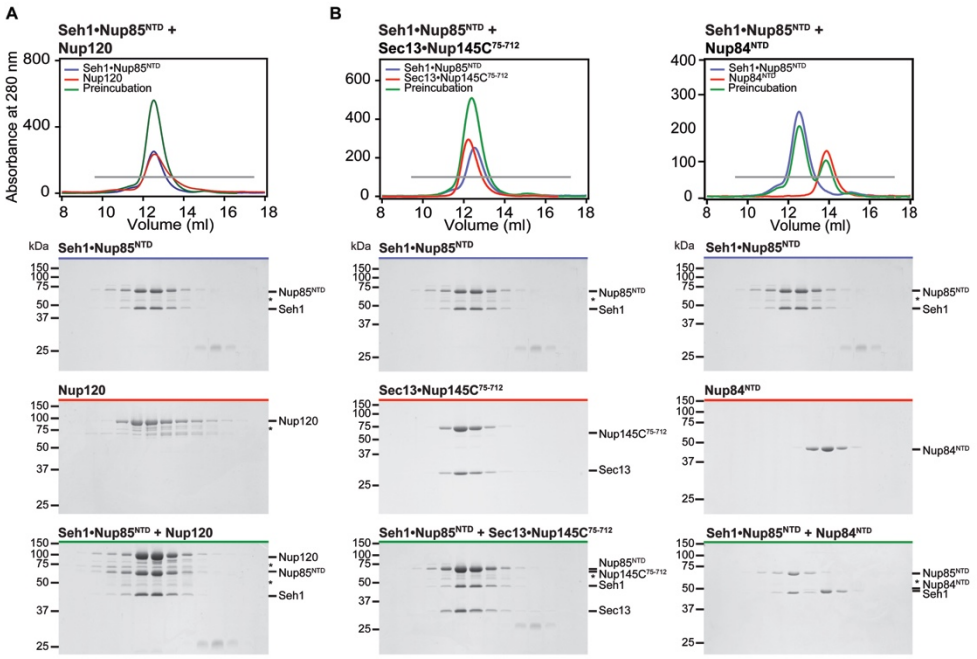


Fig. S9.
Seh1•Nup85^{NTD} forms no interactions with Nup84^{NTD}, Nup120 or Sec13•Nup145C$^{75-712}$. (A-C) Gel filtration profiles of Seh1•Nup85^{NTD} (A-C, blue), Nup120 (**A**, red), Sec13•Nup145C$^{75-712}$ (**B**, red), Nup84^{NTD} (C, red) and after pre-incubation (green). Gray bars in the gel filtration profiles indicate the fractions resolved on the SDS-PAGE gels. Molecular mass standards and the positions of the proteins are indicated. Asterisks indicate degradation products. SDS-PAGE gels were stained with Coomassie brilliant blue.

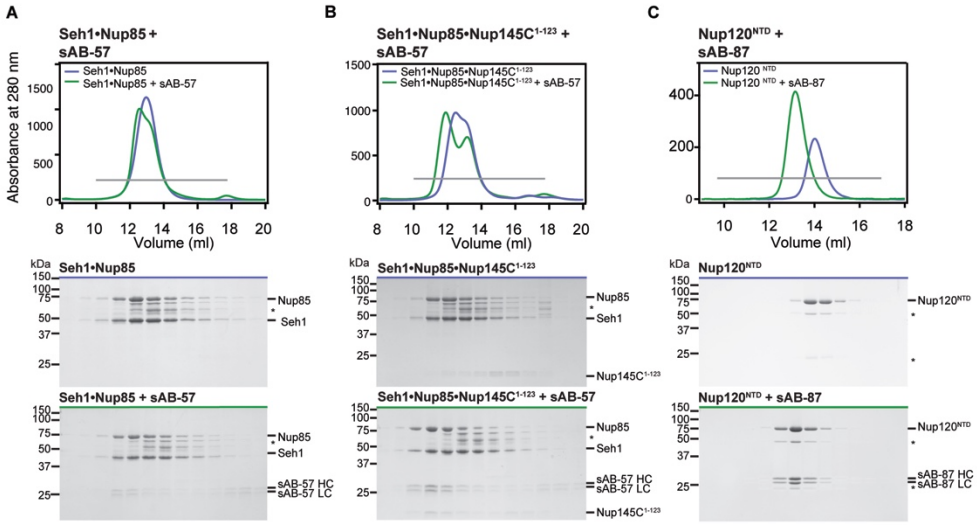


Fig. S10.

Synthetic antibody (sAB) interactions. (A) sAB-57 interacts with Seh1•Nup85. Gel filtration profiles for Seh1•Nup85 (blue), Seh1•Nup85 preincubated with sAB-57 (green). **(B)** sAB-57 interacts with Seh1•Nup85•Nup145C$^{1-123}$. Gel filtration profiles for Seh1•Nup85•Nup145C$^{1-123}$ (blue), Seh1•Nup85•Nup145C$^{1-123}$ preincubated with sAB-57 (green). Although sAB-57 can weakly interact with Seh1•Nup85 (A), the interaction is only stoichiometric in the presence of Nup145C$^{1-123}$ (B). **(C)** sAB-87 interacts with Nup120^{NTD}. Gel filtration profiles for Nup120^{NTD} (blue) and Nup120^{NTD} pre-incubated with sAB-87 (green). Notably, sAB-57 and sAB-87 non-specifically interact with the Superdex 200 resin in the tested buffer conditions and were thus not analyzed in isolation. Gray bars in the gel filtration profiles indicate the fractions resolved on the SDS-PAGE gels. Molecular mass standards and the positions of the proteins are indicated. Asterisks indicate degradation products. SDS-PAGE gels were stained with Coomassie brilliant blue.

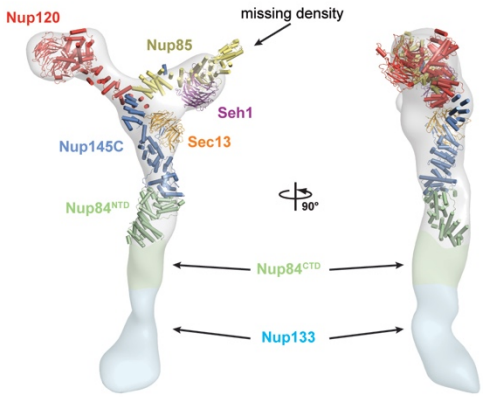


Fig. S11.
Comparison of the yeast coat nucleoporin complex crystal structure and its negative-stain EM reconstruction. Two views of the crystal structure, colored as in Fig. 1, superimposed on the negative-stain EM reconstruction (EMDB-5151 (*4*)) shown as a gray surface and contoured at 4.5 σ. Portions of the EM reconstruction are shaded green or cyan to indicate components of the complex that were not crystallized.

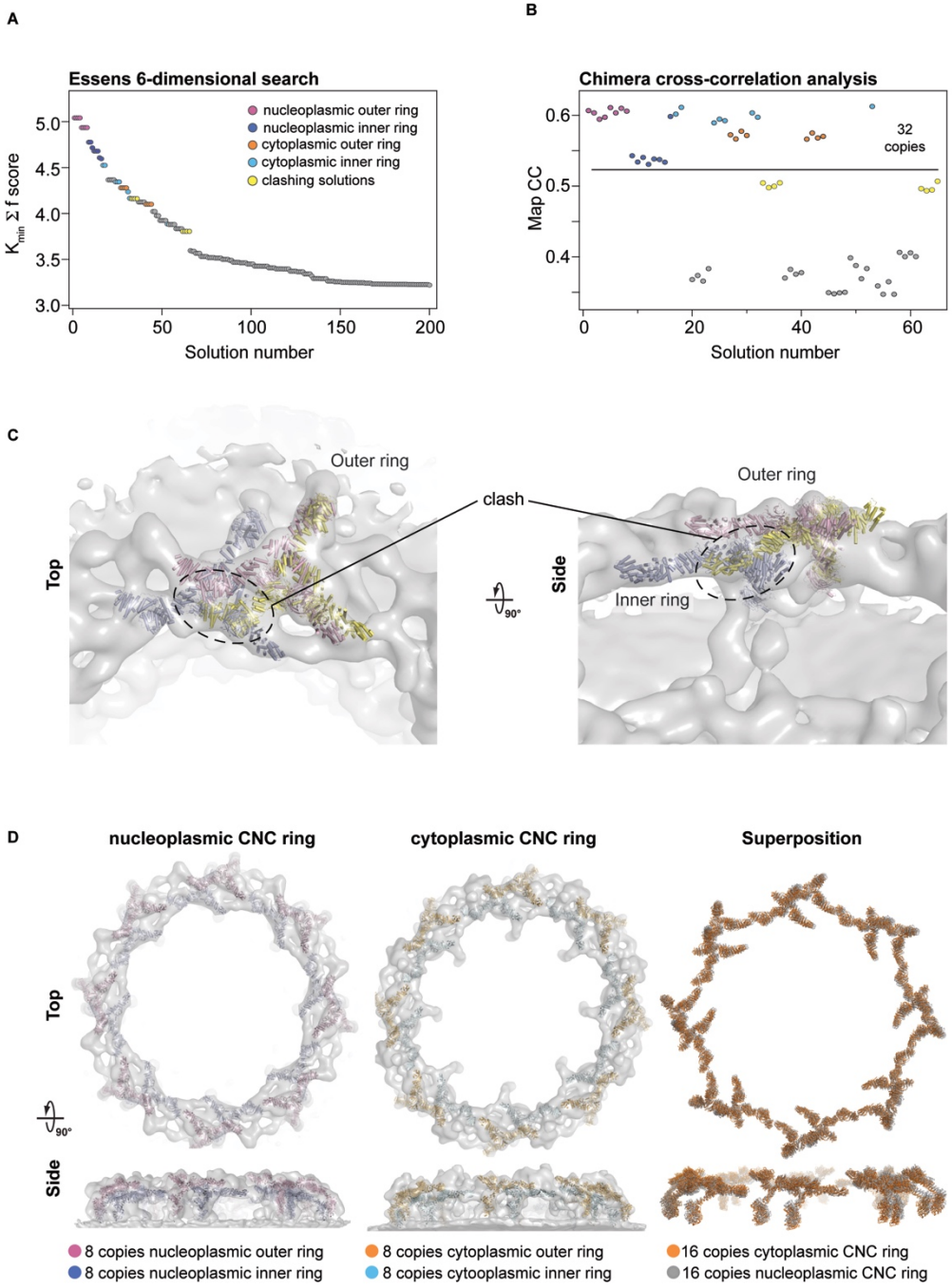


Fig. S12.
EM docking statistics. (**A**) The top 200 solutions (out of several thousands) plotted by ranked solution number and score. Top scoring solutions after refinement and rescoring with UCSF Chimera Fit in Map analysis are highlighted. (**B**) The top 65 solutions which showed a clear separation from the remaining solutions were refined and rescored with the UCSF Chimera Fit in Map tool. The highest 32 scoring solutions could be separated into

four groups related by eight-fold rotational symmetry and are colored accordingly. The next highest scoring group of solutions, which clash with the top 32 solutions, is colored in yellow. (**C**) Analysis of the next highest group of solutions. A member from each of the top scoring group of solutions is shown in cartoon representation, colored as in (A). The third solution clashes with the top two solutions, which are compatible with one another. (**D**) Comparison of the nucleoplasmic and cytoplasmic rings. Top and side views for the two faces of the NPC are shown with the members of each ring docked into the EM envelope. A superposition of the two rings without the EM density is shown on the right, which highlights the identical arrangement of CNCs despite no *a priori* information in the map restricting them to be the same.

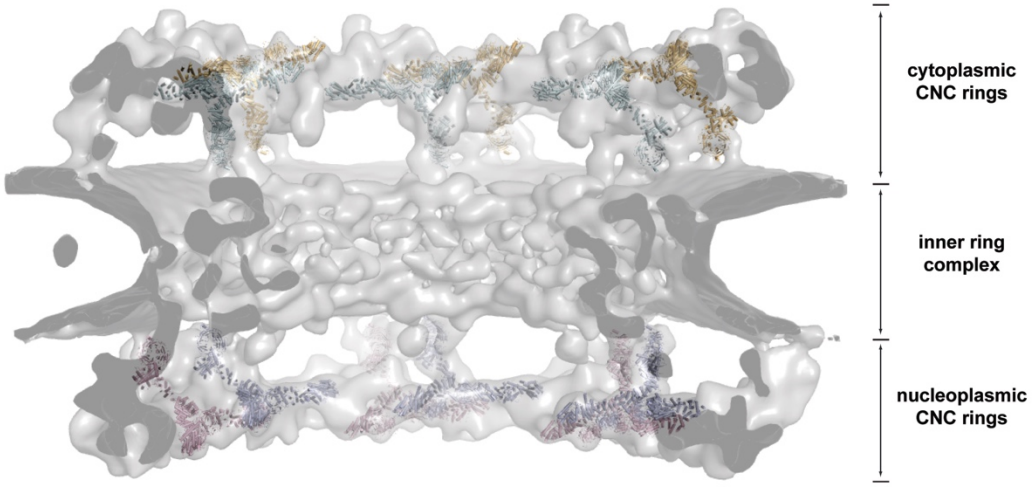


Fig. S13.
EM docking. A side view of the NPC is shown from within the central channel, highlighting the placement of 32 CNCs to the nuclear and cytoplasmic densities in the cryo-electron tomography reconstruction of the human NPC. The subunit organization of the inner ring remains unknown.

Table S1.
Bacterial expression constructs

Protein	Residues	Expression vector	Restriction sites 5', 3'	N-terminal overhang	Reference
Nup120♦, ‡	1-1037	pET3d	NdeI, NotI	MGSSHHHHHHSD	Lutzmann et al.
Nup145C	75-712	pETDuet1	BamHI, NotI	MGSSHHHHHHSQDP	This work
Sec13♦, ‡	1-297		NdeI, XhoI	None	
Nup145C	75-556	pETDuet1	BamHI, NotI	MGSSHHHHHHSQDP	This work
Sec13	1-297		NdeI, XhoI	None	
Nup145C	126-712	pETDuet1	BamHI, NotI	MGSSHHHHHHSQDP	This work
Sec13	1-297		NdeI, XhoI	None	
Nup145C	126-556	pETDuet1	BamHI, NotI	MGSSHHHHHHSQDP	Nagy et al.
Sec13	1-297		NdeI, XhoI	None	
Nup85	44-744	pETDuet1	NdeI, NotI	MGSSHHHHHHSQDP	This work
Seh1♦, ‡	1-349		NheI, XhoI	None	
Nup85	1-570	pETDuet1	NdeI, NotI	MGSSHHHHHHSQDP	This work
Seh1	1-349		NheI, XhoI	None	
$Nup85^{CTD}$	533-744	pET-MCN-SUMO	BamHI, NotI	S	This work
$Nup84^{NTD}$♦	1-451	pET28a-PreS	NdeI, NotI	GPH	Nagy et al.
Nup145C	74-109	pET-MCN-SUMO	BamHI, NotI	S	This work
Nup145C	1-123	pET28a -SUMO	BamHI, NotI	S	This work
$Nup120^{NTD}$‡	1-721	pGEX-4T1-TEV	BamHI, NotI	GAMGS	Seo et al.
$Nup84^{NTD}$‡	1-451	pET28a-PreS	NdeI, NotI	GPLMSGLNDIFEAQKI EWHEGSAGGSGHM	This work
sAB-57 LC	1-217				
sAB-57 HC♦	1-271	pSFV4	Nde1, Sal1	None	This work
sAB-57 LC	1-217				
sAB-87 HC♦	1-267	pSFV4	Nde1, Sal1	None	This work

♦ Constructs that were used for crystallization of the coat nucleoporin complex
‡ Constructs that were used for sAB selection

Table S2.
Crystallographic analysis

Data collection		*SeMet/K_2OsCl_6*	*Ta_6Br_{12}*
Protein	Yeast CNC•sAB-57	Yeast CNC• sAB-57•sAB-87	Yeast CNC• sAB-57•sAB-87
Synchrotron	APS[a]	APS[a]	APS[a]
Beamline	23-ID-D	23-ID-D	23-ID-D
Space group	C2	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	211.2, 186.4, 199.8	117.1, 180.0, 441.2	113.0, 175.8, 442.6
α, β, γ (°)	90.0, 100.9, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength	0.9794	0.9794	1.2548
Resolution (Å)	70.0 – 7.4	50.0 – 7.6	50.0 – 10.4
R_{meas} (%)[b]	10.1 (212.5)	10.8 (121.4)	14.9 (134.7)
$CC_{1/2}$	99.9 (51.8)	99.9 (60.6)	99.7 (80.3)
$< I / \sigma I >$[b]	15.3 (1.4)	13.5 (1.8)	6.4 (1.6)
Completeness (%)[b]	99.0 (94.9)	99.2 (97.7)	99.0 (100.0)
No. of observations	139,658	167,573	45,583
No. of unique reflections[b]	10,375 (1,578)	22,082 (3,431)	8,095 (1,062)
Redundancy[a]	13.5 (13.5)	7.0 (7.0)	5.6 (5.7)
Refinement			
Resolution (Å)	70.0 – 7.4	50.0 – 7.6	
No. of reflections	10,227	22,065	
No. of reflections test set	1,022 (10.0 %)	2,193 (9.9 %)	
R_{work} / R_{free}	33.0 / 35.3	31.8 / 34.7	
No. Protein atoms	26,147	19,824	
R.m.s deviations			
Bond lengths (Å)	0.006	0.005	
Bond angles (°)	1.2	1.1	
B-factors[c]	716.5	536.5	
Ramachandran plot[d]			
Favored (%)	91.0	91.9	
Additionally allowed (%)	7.3	6.4	
Outliers (%)	1.7	1.8	

[a]APS, Advanced Photon Source
[b]Highest-resolution shell is shown in parentheses
[c]B-factors include overall B-factor of the crystal
[d]As determined by MolProbity (*26*)

REFERENCES

1. Hoelz A, Debler EW, Blobel G. The structure of the nuclear pore complex. Annu Rev Biochem. 2011; 80:613–643.
2. Lutzmann M, Kunze R, Buerer A, Aebi U, Hurt E. Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins. EMBO J. 2002; 21:387–397.
3. Bui KH, et al. Integrated structural analysis of the human nuclear pore complex scaffold. Cell. 2013; 155:1233–1243.
4. Kampmann M, Blobel G. Three-dimensional structure and flexibility of a membrane-coating module of the nuclear pore complex. Nat Struct Mol Biol. 2009; 16:782–788.
5. Paduch M, et al. Generating conformation-specific synthetic antibodies to trap proteins in selected functional states. Methods. 2013; 60:3–14.
6. Debler EW, et al. A fence-like coat for the nuclear pore membrane. Mol Cell. 2008; 32:815–826.
7. Hsia KC, Stavropoulos P, Blobel G, Hoelz A. Architecture of a coat for the nuclear pore membrane. Cell. 2007; 131:1313–1326.
8. Nagy V, et al. Structure of a trimeric nucleoporin complex reveals alternate oligomerization states. Proc Natl Acad Sci USA. 2009; 106:17693–17698.
9. Seo HS, et al. Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. Proc Natl Acad Sci USA. 2009; 106:14281–14286.
10. Rizk SS, et al. Allosteric control of ligand-binding affinity using engineered conformation-specific effector proteins. Nat Struct Mol Biol. 2011; 18:437–442.
11. Thierbach K, et al. Protein interfaces of the conserved Nup84 complex from Chaetomium thermophilum shown by crosslinking mass spectrometry and electron

microscopy. Structure. 2013; 21:1672–1682.

12. Shi Y, et al. Structural Characterization by Cross-linking Reveals the Detailed Architecture of a Coatomer-related Heptameric Module from the Nuclear Pore Complex. Mol Cell Proteomics. 2014; 13:2927–2943.
13. Alber F, et al. The molecular architecture of the nuclear pore complex. Nature. 2007; 450:695–701.
14. Brohawn SG, Leksa NC, Spear ED, Rajashankar KR, Schwartz TU. Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. Science. 2008; 322:1369–1373.
15. Mitchell JM, Mansfeld J, Capitanio J, Kutay U, Wozniak RW. Pom121 links two essential subcomplexes of the nuclear pore complex core to the membrane. J Cell Biol. 2010; 191:505–521.
16. Harrison SC, Kirchhausen T. Structural biology: Conservation in vesicle coats. Nature. 2010; 466:1048–1049.
17. Debler EW, Hsia KC, Nagy V, Seo HS, Hoelz A. Characterization of the membrane-coating Nup84 complex: paradigm for the nuclear pore complex structure. Nucleus. 2010; 1:150–157.
18. Hoelz A, Nairn AC, Kuriyan J. Crystal structure of a tetradecameric assembly of the association domain of Ca2+/calmodulin-dependent kinase II. Mol Cell. 2003; 11:1241–1251.
19. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132.
20. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674.

21. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr. 2000; 56:965–972.

22. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132.

23. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221.

24. Joosten RP, Long F, Murshudov GN, Perrakis A. The PDB_REDO server for macromolecular structure model optimization. IUCrJ. 2014; 1:213–220.

25. Karplus PA, Diederichs K. Linking crystallographic model and data quality. Science. 2012; 336:1030–1033.

26. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21.

27. Pettersen EF, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25:1605–1612.

28. Kleywegt GJ, Jones TA. Template convolution to enhance or detect structural features in macromolecular electron-density maps. Acta Crystallogr D Biol Crystallogr. 1997; 53:179–185.

CHAPTER 2

EVIDENCE FOR AN EVOLUTIONARY RELATIONSHIP BETWEEN THE LARGE ADAPTOR NUCLEOPORIN NUP192 AND KARYOPHERINS

This chapter was adapted from:

Tobias Stuwe[#], **Daniel H. Lin**[#], Leslie N. Collins, Ed Hurt, André Hoelz (2014). Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins, ***Proc. Natl. Acad. Sci. U.S.A.***, 111(7):2530-2535.

ABSTRACT

Nucleocytoplasmic transport is facilitated by nuclear pore complexes (NPCs), which are massive proteinaceous transport channels embedded in the nuclear envelope. Nup192 is a major component of an adaptor nucleoporin subcomplex proposed to link the NPC coat with the central transport channel. Here, we present the structure of the ~110 kDa N-terminal domain (NTD) of Nup192 at 2.7-Å resolution. The structure reveals an open ring-shaped architecture composed of HEAT and ARM repeats. A comparison of different conformations indicates that the NTD consists of two rigid halves connected by a flexible hinge. Unexpectedly, the two halves of the ring are structurally related to karyopherin-α and β-karyopherin family members. Biochemically, we identify a conserved patch that binds an unstructured segment in Nup53 and show that a C-terminal tail region binds to a putative helical fragment in Nic96. The Nup53 segment that binds Nup192 is a classical nuclear localization-like sequence that interacts with karyopherin-α in a mutually exclusive and mechanistically distinct manner. The disruption of the Nup53 and Nic96 binding sites *in vivo* yields growth and mRNA export defects, revealing their critical role in proper NPC function. Surprisingly, both interactions are dispensable for NPC localization, suggesting that Nup192 possesses another nucleoporin interaction partner. These data indicate that the structured domains in the adaptor nucleoporin complex are held together by peptide interactions that resemble those found in karyopherin•cargo complexes and support the proposal that the adaptor nucleoporins arose from ancestral karyopherins.

INTRODUCTION

Nucleocytoplasmic transport is an essential process in which transport factors, called karyopherins, mediate nuclear transport of macromolecules larger than ~40 kDa (1-3). Karyopherins are classified into two families, α-karyopherins and β-karyopherins, which are composed of ARM and HEAT repeats, respectively (1, 3). Karyopherins facilitate nucleocytoplasmic transport of proteins by recognizing linear nuclear localization sequences (NLS) or nuclear export sequences (NES) in their cargo proteins. α-karyopherins are import adaptors that interact with β-karyopherins through an N-terminal importin-β binding domain (IBB) to coordinate cargo import through the nuclear pore complex (NPC) (1-4). While α-karyopherins are structurally rigid, β-karyopherins display remarkable conformational flexibility, which enables them to regulate cargo recognition (1, 3, 5).

The NPC is the sole gateway for bi-directional nucleocytoplasmic transport (6). Its transport channel is lined with intrinsically disordered phenylalanine-glycine (FG) repeats, which form a diffusion barrier and bind karyopherin•cargo complexes (1, 6). In yeast, NPCs are composed of 34 different proteins, termed nucleoporins (nups), that assemble in multiple copies into a ~60 MDa complex (6). Electron microscopy (EM) revealed that the NPC consists of a doughnut-shaped symmetric core, which is embedded in the nuclear envelope and decorated with filamentous structures on its cytoplasmic and nucleoplasmic faces. Cytoplasmic filaments extend into the cytoplasm and provide docking sites for karyopherins, the GTPase Ran, and the mRNA export machinery (1, 5, 6). On the nucleoplasmic face, a nuclear basket structure binds Ran and the transcription machinery and participates in chromatin organization (6).

The symmetric NPC core can be considered schematically as a series of concentric cylindrical layers, composed of integral membrane proteins of the pore membrane (POMs),

the coat forming Nup84 complex, the adaptor nucleoporin complex (ANC), and the central channel nups (6, 7). Structural studies provided key insights into the architecture and function of the coat and channel nups, as well as the interactions of the cytoplasmic filament nups with the mRNA export machinery (7-16).

An EM analysis of the coat-forming, heptameric Nup84 complex from *S. cerevisiae* revealed a ~400-Å long, Y-shaped architecture (17). Crystallographic studies established that the heptamer is composed of α-helical solenoids and β-propellers interacting through extensive hydrophobic interfaces (7-9, 12, 13, 15). The heptamer is architecturally similar to the membrane bending COP-I, COP-II, and clathrin coats, suggesting that the Nup84 complex forms a coat for the nuclear envelope (7-9, 18).

In contrast to the Nup84 complex, the interactions between the channel nups are mediated by ~250-residue α-helical regions, which adopt a range of alternative conformations and interactions including sliding α-helical domains, alternative assembly states, and changes in interaction partners (19, 20). The plasticity of these nups is believed to facilitate the transport of different sized cargoes by dilating and contracting the transport channel (19, 20).

Unlike the coat and channel nups, the interactions within the ANC are poorly understood, largely due to the poor behavior of the large adaptor nups in solution. However, the adaptor nups of the eukaryotic thermophile *C. thermophilum* possess improved biochemical robustness (21). The *C. thermophilum* ANC is composed of Nup192, Nup188, Nup170, Nic96, and Nup53, which are conserved throughout the eukaryotic kingdom (6, 21). Nic96 and Nup192 are the only essential members of the ANC in *S. cerevisiae* and are thought to function as the scaffolds onto which the adaptor layer assembles (21, 22). Negative-stain EM of Nup192 revealed a question mark-shaped architecture (21).

To gain structural and functional insight into the ANC, we determined the crystal structure of the N-terminal domain of *C. thermophilum* Nup192 at 2.7-Å resolution. Our structural analysis reveals that ctNup192^{NTD} forms an α-helical solenoid with a ~110 Å wide ring-shaped architecture with similarities to the β-karyopherin Cse1p and import adaptor karyopherin-α. Furthermore, we show that ctNup192^{NTD} possesses a binding site on its convex surface for an unstructured fragment of ctNup53, whereas a C-terminal tail fragment of ctNup192 binds to an α-helical region of ctNic96. Hence, ctNup192 is an interaction platform in the ANC with distinct binding sites. Disruption of either or both of these interactions results in growth and mRNA export defects *in vivo*, establishing their physiological importance. However, both interactions are dispensable for NPC localization, suggesting that Nup192 possesses at least one additional nup binding site sufficient to anchor Nup192 in the NPC. Based on our data, we propose that the interactions between the adaptor nups are governed by short linear motifs resembling those found in classical karyopherin•cargo complexes.

RESULTS

Structure determination

We identified a fragment of *C. thermophilum* Nup192 encompassing residues 1 to 958 that yielded well-behaved, soluble protein (ctNup192^{NTD}) (Fig. 1*A*). ctNup192^{NTD} crystallized in space group $P4_32_12$ with two molecules in the asymmetric unit. The structure was solved by single-wavelength anomalous dispersion using X-ray diffraction data from a seleno-L-methionine derivative (Fig. S1*A*, *B*). The final model was refined to 2.7-Å resolution with R_{work} and R_{free} values of 19.1 % and 23.1 %, respectively. For data collection and refinement statistics, see Table S1. As there was no dimerization in solution, we focused on the ctNup192^{NTD} monomer (Fig. S1*C*).

Structural overview

ctNup192^{NTD} folds into a ring-shaped solenoid composed of 42 α-helices (α1-α42) and a single β-hairpin (β1-β2) with a right-handed superhelical twist and dimensions of 110 Å × 90 Å × 40 Å (Fig. 1). The N- and C-terminal halves of the solenoid do not form direct contacts, resulting in an ~10-Å gap in the ring and an overall lock washer-like shape. ctNup192^{NTD} is divided into four structural segments: the N-terminal Head (residues 1 to 184), the HEAT repeat (residues 185 to 352), the small hinge (residues 353 to 415), and the C-terminal ARM repeat modules (residues 416 to 958) (Fig. 1).

The Head module consists of two helical pairs, α1-α2 and α7-α8, arranged in a HEAT-like topology interrupted by an α-helical insertion (α3-α6) containing the β-hairpin (β1-β2). The Head and HEAT modules are connected by a long, disordered loop between helices α8 and α9 and make hydrophobic contacts with HEAT repeat 1 (α9-α10) (Figs. 1*B* and S1*D*).

The HEAT module is composed of three HEAT repeats that each form one turn of a right-handed superhelix. The HEAT and ARM modules are bridged by a short hinge module composed of the hinge loop and two helices, α15 and α16. These helices are connected by a single glycine residue that is invariant across fungi, suggesting that this feature is evolutionarily conserved (Fig. S2). Helix α15 caps the end of the HEAT module superhelix and helix α16 initiates the ARM module superhelix. The hinge loop is an ordered, 31-residue connector, which forms contacts along the entire concave surface of the ring. The ARM module forms a right-handed superhelix composed of non-canonical ARM repeats. ARM repeat 6 (α32-α35) is a degenerate ARM repeat in which the α31-α32 loop replaces the first helix (Figs. 1 and S1*D*).

Structural similarity to karyopherins

The unusual combination of HEAT and ARM repeats in ctNup192^{NTD} prompted us to look for proteins with a similar architecture. In structural homology searches with the DALI server, we found that yeast karyopherin-α (Kap-α) was most similar to the ARM module of ctNup192^{NTD} (23). The structure of yeast Kap-α superposes with the ARM module with a root-mean-square deviation (RMSD) of 5.7 Å over 303 Cα atoms (Fig. 2). Both proteins have an identical topology, but the α31-α32 loop perturbs the curvature of the ARM module superhelix, generating a greater curvature in ctNup192^{NTD}. Intriguingly, the hinge and α31-α32 loops overlap with the Kap-α recognition sites for a bipartite NLS (Fig. 2). However, the conserved tryptophans that line the concave surface of Kap-α and bind the NLS peptide are absent in ctNup192^{NTD}. Likewise, a DALI search on the HEAT module of ctNup192^{NTD} found the export β-karyopherin Cse1p to be structurally similar. The HEAT module superposed with three N-terminal HEAT repeats of Cse1p with a RMSD of 5.2 Å over 117

Cα atoms (Fig. S3*A*). Thus, ctNup192^{NTD} possesses a surprising architectural similarity to members of both karyopherin families, wherein the C-terminal ARM module is structurally analogous to Kap-α and the N-terminal HEAT module is similar to β-karyopherins.

Structural analysis

Given the unusual topology of ctNup192^{NTD}, we further analyzed the individual ARM and HEAT repeats. Sequence and structural alignments revealed that while the ARM and HEAT repeats of ctNup192^{NTD} possess the hallmark sequence characteristics for their respective motifs, they are also more divergent in helical length and position (Fig. S3*B*, *C*). Superposition of the two molecules in the asymmetric unit and the structure of *S. cerevisiae* Nup192^{NTD} revealed conformational changes in the entire ring that open and close the gap between the two ends of the ring by ~10 Å (Fig. S4*A*) (24). These conformational changes are mostly the result of rigid body rearrangements around the hinge module (Fig. S4*B*, *C*). For a detailed analysis, see Supporting Text.

Surface properties

In order to identify functionally important surfaces in ctNup192^{NTD}, we created surface maps depicting evolutionary conservation and electrostatic potential (Figs. 3 and S2). An immediately apparent feature is a conserved hydrophobic groove on the convex surface of the ring, next to the hinge module (Fig. 3*A*, *B*). In CRM1, the availability of the hydrophobic NES binding cleft is modulated by a RanGTP-dependent rearrangement of an acidic loop (25). The hydrophobic groove observed in ctNup192^{NTD} is located in a similar position on the outside of the ring, suggesting that the interaction with another nup at this site is regulated similarly.

Additionally, we could identify a conserved, charged surface patch that spans the α-helical insertion of the Head module and the front face of the HEAT module (Fig. 3). A third, C-terminal surface also displays a high degree of conservation, but it is formed by the last α-helix in the truncated construct, suggesting that it is normally buried in the NTD-CTD interface of full-length ctNup192 (Fig. 3*B*). Finally, we identified a surface pocket with a highly negative electrostatic potential on the back face of ctNup192^{NTD} (Fig. 3*C*). While this pocket is not strictly conserved in sequence, its negative character is maintained in the structure of scNup192^{NTD} (24). Furthermore, this acidic pocket is immediately adjacent to the hydrophobic groove, suggesting that these two features form a composite binding site for another nup. Given their distinct chemical natures, all of the identified surfaces represent likely protein-protein interaction sites.

Biochemical analysis

Nup192 was originally identified as the *S. cerevisiae* homologue of vertebrate Nup205 and subsequently as an interaction partner of Nic96 (22, 26). Recently, ctNup192 was found to interact with both a fragment of ctNic96 and a region of ctNup53 (21). Based on these results, we tested whether an N-terminal fragment of ctNup53 (ctNup53^{N}, residues 1 to 90) and an N-terminal, α-helical segment of ctNic96 (ctNic96^{H2}, residues 262 to 301) form a complex with ctNup192^{NTD} (Fig. 4*A*). In size-exclusion chromatography (SEC) interaction experiments, ctNup192^{NTD} formed a stoichiometric complex with ctNup53^{N}, but failed to interact with ctNic96^{H2} (Figs. 4*B, C* and S5*A*-*C*). In contrast, the ctNup192 C-terminal domain (ctNup192^{CTD}, residues 976 to 1756) is capable of forming a stoichiometric complex with ctNic96^{H2}, but fails to form a complex with ctNup53^{N} (Figs. 4*D, E*, and S5*D*, *E*). Further

mapping revealed that only the C-terminal 340 residues of ctNup192, ctNup192^{TAIL}, mediate the interaction with ctNic96^{H2} (Fig. S5*F*).

ctNup53^{N} contains a 37-residue region (residues 31 to 67) which resembles a classical bipartite NLS with the consensus sequence KR(X)$_{10-12}$K(K/R)X(K/R), but with deviations wherein lysine and arginine residues are interchanged (Fig. 5*A*). Given the structural similarity of ctNup192^{NTD} to Kap-α, we tested whether ctNup53^{N} is also capable of interacting with Kap-α Indeed, ctNup53^{N} and Kap-α form a stoichiometric complex on a gel filtration column (Figs. 4*F* and S5*G*). We next tested whether ctNup53^{N} could simultaneously interact with both ctNup192^{NTD} and Kap-α by incubating the preformed Kap-α•ctNup53^{N} pair with ctNup192^{NTD}. However, in our SEC interaction experiments, ctNup53^{N} binding to ctNup192^{NTD} and Kap-α is mutually exclusive (Figs. 4*G* and S5*H*). The unexpected finding that ctNup192^{NTD} and Kap-α both interact with the same ctNup53 fragment further supports an evolutionary connection between karyopherins and adaptor nups

In order to gain additional insight into the molecular details of the ctNup53^{N} interactions, we generated a minimal fragment comprising residues 31 to 67, ctNup53$^{31-67}$, and purified nine ctNup53$^{31-67}$ alanine mutants of conserved residues. In SEC interaction experiments, only mutations of basic residues at consensus bipartite NLS positions (R39, K40, R53, and R54) impaired the interaction with Kap-α (Figs. 5*A* and S6*A*). Additionally, two of the mutations affecting Kap-α binding, R53A and R54A, also weakened the interaction with ctNup192^{NTD} (Figs. 5*A* and S6*B*), providing a molecular explanation for the exclusivity of these interactions. Strikingly, a single mutation, F48A, completely abolished the ctNup53$^{31-67}$-ctNup192^{NTD} interaction, with no effect on Kap-α binding (Figs. 5*A, B* and S6*B*). Thus, the ctNup53$^{31-67}$-Kap-α interaction appears to be mechanistically similar to

classical NLS binding, whereas the ctNup53$^{31-67}$-ctNup192^{NTD} interaction is mediated by a distinct binding mode.

We next determined whether any of the potential protein-protein interaction sites on the ctNup192^{NTD} surface mediate binding to ctNup53$^{31-67}$. We purified 33 single alanine mutants of conserved residues distributed throughout the identified surfaces, as well as a variant of ctNup192^{NTD} that lacked the Head module (Fig. 5). All tested mutants were properly folded, as they were indistinguishable in their behavior from wildtype ctNup192^{NTD}. Whereas deletion of the Head module and mutations in the potential NTD-CTD interface and α31-α32 loop had no effect on ctNup53$^{31-67}$ binding, there was a strong clustering of mutations in the acidic pocket and hydrophobic groove that reduced (E295A, E335A, D431A, D488A, and D439A) or completely abolished (L441A, W499A) ctNup53$^{31-67}$ binding (Figs. 5*B* and S7*A*). These mutations are on the convex surface of ARM1 and ARM2 or immediately adjacent to it (Fig. 5*C*). L441 and W499, which are essential for ctNup53$^{31-67}$ binding, form a deep hydrophobic pocket in the conserved hydrophobic groove (Fig. 5*C*). These results establish that ctNup192^{NTD} binds to ctNup53$^{31-67}$ with a combined surface that stretches from the hydrophobic groove formed by the convex surfaces of ARM1 and ARM2 to the immediately adjacent acidic pocket. The mechanism of the Nup192-Nup53 interaction is evolutionarily conserved, as wildtype *S. cerevisiae* scNup192^{NTD} and scNup53^{N} form stoichiometric complexes which are abolished by the corresponding point mutations, scNup192$^{NTD, W513A}$ and scNup53$^{N, F124A}$ (Figs. S6*C* and S7*B*).

Using isothermal titration calorimetry (ITC), we found that the ctNup192^{NTD}-ctNup53$^{31-67}$ interaction has a dissociation constant of ~0.1 μM and that single alanine mutations, ctNup53$^{31-67, F48A}$ and ctNup192$^{NTD, W499A}$, reduce binding to undetectable levels, consistent with our SEC interaction experiments (Fig. S7*C*-*E*). Surprisingly, the interaction

between ctNup53$^{31-67}$ and Kap-α is 250-fold weaker (~25 μM) than the interaction between ctNup53$^{31-67}$ and ctNup192^{NTD} (Fig. S7*F*), likely a result of a deviating NLS sequence observed in ctNup53$^{31-67}$. The ITC results are summarized in Fig. S7*G*.

Given the similarities in overall shape and architecture between Nup192 and the other large adaptor nup Nup188 (6), we tested whether ctNup188^{NTD} also interacts with ctNup53$^{31-67}$, but detected no binding in a SEC interaction experiment (Fig. S8*A*). A comparison of the ctNup192^{NTD} and *M. thermophila* Nup188^{NTD} structures provides a structural explanation for this result (Fig. S8*B, C*) (27). While the two proteins share the same topology and organization, two features are unique to mtNup188^{NTD}: an N-terminal clamp module that binds the C-terminus of mtNup188^{NTD} and an SH3-like insertion in ARM repeat 3 between helices α23 and α24, protruding out of the convex surface (Fig. S8*B*). Furthermore, the ctNup53$^{31-67}$ binding site is not conserved in mtNup188^{NTD} and possesses the following significant alterations: (1) the bulky hydrophobic residues that form the deep pocket in ctNup192^{NTD} are absent in mtNup188^{NTD} and instead the helices pack more closely together, and (2), helix α18, which forms a large part of the conserved hydrophobic groove, is absent in mtNup188^{NTD} and this part of the surface is instead occupied by the SH3-like insertion, thus restricting access to the convex surface (Fig. S8*C*).

Altogether, these results show that Nup192 interacts with two other adaptor nups, Nup53 and Nic96, via spatially separated binding sites at opposite ends of the question mark-shaped molecule, consistent with its ANC-scaffolding role. Despite their structural and evolutionary relationship, the large adaptor nups Nup192 and Nup188 fulfill functionally distinct roles in the NPC, as the interaction with Nup53 is not replicated in Nup188.

Functional analysis

Nup192 is essential for viability in *S. cerevisiae*, but its roles in NPC structure and function are poorly understood (22). To identify the roles of Nup192 and its binding partners *in vivo*, we determined the viability of *S. cerevisiae nup192*Δ deletion strains complemented by various GFP-tagged Nup192 constructs. Neither Nup192^{NTD} nor Nup192^{CTD} alone were sufficient to overcome the lethality of the Nup192 knock-out (Fig. 6*A*). In contrast, Nup192 variants that abolish the interaction with Nup53 (Nup192$^{FL,\ W513A}$) or Nic96 (Nup192$^{\Delta TAIL}$) were sufficient to restore viability (Fig. 6*A*). Viability correlated with targeting of the GFP-Nup192 fusion constructs to the nuclear envelope. While GFP-Nup192^{NTD} and GFP-Nup192^{CTD} were diffusely localized throughout the entire cell, GFP-Nup192$^{FL,\ W513A}$ and GFP-Nup192$^{\Delta TAIL}$ variants displayed robust nuclear rim staining consistent with NPC incorporation (Fig. 6*B, C*). Despite rescuing lethality, individual disruption of the Nup53 or Nic96 interactions yielded mild and severe growth defects, respectively (Fig. 6*D*). Consistent with these results, complementation of a *nup53*Δ*nup59*Δ strain with the GFP-Nup53^{F124A} mutant, which is defective in Nup192 binding, exhibited no observable growth defect (Fig. 6*D*).

We next assayed mRNA export by fluorescence in situ hybridization of an Alexa647-labeled dT_{50} oligonucleotide as a probe for NPC function. In agreement with the observed growth phenotypes, disruption of Nic96 binding (Nup192$^{\Delta TAIL}$) caused a substantial mRNA export defect as ~53 % of cells displayed nuclear mRNA retention compared to ~7 % of cells containing full-length Nup192. Loss of Nup53 binding (Nup192$^{FL,\ W513A}$) had only a mild effect on mRNA export with ~10 % of cells displaying a defect (Fig. 6*E*). Interestingly, the loss of both the Nic96 and Nup53 interactions failed to have an additive effect, as the Nup192$^{\Delta TAIL,\ W513A}$ variant exhibited a less pronounced growth phenotype than Nup192$^{\Delta TAIL}$, despite still possessing a major mRNA export defect with ~49 % of cells displaying nuclear

mRNA retention (Fig. 6*D*, *E*). Unexpectedly, GFP-Nup192$^{\Delta TAIL,\ W513A}$ still demonstrated robust rim staining, indicating that both interactions are dispensable for NPC incorporation. These results confirm the physiological importance of the interactions within the ANC interaction network that we characterized biochemically, and further suggest a yet to be identified nup interaction that is sufficient to anchor Nup192 in the NPC.

DISCUSSION

We determined the structure of the N-terminal domain of the large adaptor nucleoporin Nup192 from *C. thermophilum*. The structure revealed a ring-shaped architecture composed of an unusual combination of HEAT and ARM repeats that display unexpected structural similarities with karyopherins. Comparison of distinct conformational states of ctNup192^{NTD} revealed a hinge module that bridges the two halves of the ring and confers conformational plasticity. Furthermore, ctNup192 binds to linear sequence motifs in two other adaptor nups, ctNup53 and ctNic96, in a manner that resembles known karyopherin•cargo complex interactions (1-3, 5). We also found that the same segment in ctNup53 interacts with both ctNup192^{NTD} and Kap-α in a mutually exclusive fashion, utilizing partially overlapping, interdigitated binding motifs. Together, these data suggest an evolutionary connection between the karyopherins and the large, all-helical, adaptor nucleoporin Nup192.

Our interaction analyses facilitate the construction of a high-resolution biochemical map of the adaptor nup interaction network. Mutational analysis of the ctNup192^{NTD} surface revealed that ctNup53 binds at the midpoint of the convex surface of the ring in a combined interface that includes a conserved hydrophobic groove and an acidic pocket. These findings suggest that ctNup53 binding to ctNup192^{NTD} is mechanistically distinct from the classical bipartite NLS-Kap-α interaction, which is mediated by an array of tryptophans and acidic residues lining the concave Kap-α surface (28). Our finding that a single hydrophobic residue in ctNup53 is essential for ctNup192^{NTD} binding but dispensable for the Kap-α interaction further supports this conclusion. Whether ctNup53 indeed binds to Kap-α in a classical fashion or possesses a distinct interaction mode, as previously shown for other nups, awaits further structural characterization (29).

We also found that the ctNic96^{H2} interaction site maps to the C-terminal tail of ctNup192. By docking the structure of ctNup192^{NTD} into a negative-stain EM envelope of full-length ctNup192 (21), we establish that the ctNup53 binding site is located at the top of the question mark-shaped molecule, which is the furthest possible point from the C-terminal tail that binds ctNic96 – a distance of ~150 Å (Fig. 7). Our *in vivo* localization analysis further suggests that there is at least one additional nup interaction that is sufficient to anchor Nup192 in the NPC, as the interactions with Nup53 and Nic96 are dispensable for its localization. Nevertheless, the observed growth and mRNA export defects establish both identified interactions as important for proper NPC function.

We propose that the interactions between the structured domains in the ANC are primarily mediated by short linear motifs, which we suggest to term "Nucleoporin Anchor Sequences" (NAS). Such a peptide interaction network would be in stark contrast to the three-dimensional interfaces previously identified in the Nup84 complex (7-9, 12, 15). By tethering large, distortable adaptor nup solenoids with predominantly unstructured "linker" nups, the adaptor layer would be able to dynamically cushion the proposed dilations of the central transport channel during cargo translocation. Furthermore, the interactions between the linker nups and the large structural nups may allow for a hierarchical NPC assembly pathway resembling the assembly and disassembly of transport factor•cargo complexes. An intriguing possibility is that the mutually exclusive interactions of Nup53 with Kap-α and Nup192 may play a role in the regulation of NPC assembly or cargo transport, possibly in a manner similar to the competitive interactions in the Nup53-Nup170-Kap121 network (30, 31). Finally, such a NAS interaction network is likely to also facilitate the oligomerization of the NPC subcomplexes in the assembled NPC, a possibility exemplified by our previous observation that an unstructured segment in Nup133 mediates the head-to-tail arrangement of the Nup84

complex (13). Future work is necessary to identify all interactions within the ANC, understand its overall organization, and determine its roles in NPC function.

ACKNOWLEDGEMENTS

We thank members of the Hoelz laboratory, Alina Patke, and Yunji Wu for critical reading of the manuscript, David King for mass spectrometry analysis, Jens Kaiser and the scientific staff of SSRL Beamline 12-2 for their support with X-ray diffraction measurements, the University of Colorado Biophysics Core for assistance with ITC measurements, and Elena Conti for material. We acknowledge the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program for their support of the Molecular Observatory at the California Institute of Technology. SSRL is supported by the Department of Energy and by the National Institutes of Health (NIH). DHL is supported by a NIH Research Service Award (5 T32 GM07616). AH was supported by the Albert Wyrick V Scholar Award of the V Foundation for Cancer Research, the 54th Mallinckrodt Scholar Award of the Edward Mallinckrodt, Jr. Foundation, and a Kimmel Scholar Award of the Sidney Kimmel Foundation for Cancer Research.

REFERENCES

1. Cook A, Bono F, Jinek M, & Conti E (2007) Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* 76:647-671.
2. Hoelz A & Blobel G (2004) Cell biology: popping out of the nucleus. *Nature* 432:815-816.
3. Xu D, Farmer A, & Chook YM (2010) Recognition of nuclear targeting signals by Karyopherin-beta proteins. *Curr Opin Struct Biol* 20:782-790.
4. Lott K & Cingolani G (2011) The importin beta binding domain as a master regulator of nucleocytoplasmic transport. *Biochim Biophys Acta* 1813:1578-1592.
5. Debler EW, Blobel G, & Hoelz A (2009) Nuclear transport comes full circle. *Nat Struct Mol Biol* 16:457-459.
6. Hoelz A, Debler EW, & Blobel G (2011) The structure of the nuclear pore complex. *Annu Rev Biochem* 80:613-643.
7. Hsia KC, Stavropoulos P, Blobel G, & Hoelz A (2007) Architecture of a coat for the nuclear pore membrane. *Cell* 131:1313-1326.
8. Brohawn SG, Leksa NC, Spear ED, Rajashankar KR, & Schwartz TU (2008) Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. *Science* 322:1369-1373.
9. Debler EW*, et al.* (2008) A fence-like coat for the nuclear pore membrane. *Mol Cell* 32:815-826.
10. Kassube SA*, et al.* (2012) Crystal structure of the N-terminal domain of Nup358/RanBP2. *J Mol Biol* 423:752-765.
11. Lin DH, Zimmermann S, Stuwe T, Stuwe E, & Hoelz A (2013) Structural and Functional Analysis of the C-Terminal Domain of Nup358/RanBP2. *J Mol Biol* 425:1318-1329.

12. Nagy V*, et al.* (2009) Structure of a trimeric nucleoporin complex reveals alternate oligomerization states. *Proc Natl Acad Sci U S A* 106:17693-17698.

13. Seo HS*, et al.* (2009) Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. *Proc Natl Acad Sci U S A* 106:14281-14286.

14. Stuwe T, von Borzyskowski LS, Davenport AM, & Hoelz A (2012) Molecular basis for the anchoring of proto-oncoprotein Nup98 to the cytoplasmic face of the nuclear pore complex. *J Mol Biol* 419:330-346.

15. Whittle JR & Schwartz TU (2009) Architectural nucleoporins Nup157/170 and Nup133 are structurally related and descend from a second ancestral element. *J Biol Chem* 284:28442-28452.

16. Yoshida K, Seo HS, Debler EW, Blobel G, & Hoelz A (2011) Structural and functional analysis of an essential nucleoporin heterotrimer on the cytoplasmic face of the nuclear pore complex. *Proc Natl Acad Sci U S A* 108:16571-16576.

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

18. Debler EW, Hsia KC, Nagy V, Seo HS, & Hoelz A (2010) Characterization of the membrane-coating Nup84 complex: paradigm for the nuclear pore complex structure. *Nucleus* 1:150-157.

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

20. Solmaz SR, Blobel G, & Melcak I (2013) Ring cycle for dilating and constricting the nuclear pore. *Proc Natl Acad Sci U S A* 110:5858-5863.

21. Amlacher S*, et al.* (2011) Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell* 146:277-289.

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

23. Holm L & Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38:W545-549.

24. Sampathkumar P*, et al.* (2013) Structure, dynamics, evolution, and function of a major scaffold component in the nuclear pore complex. *Structure* 21:560-571.

25. Monecke T*, et al.* (2013) Structural basis for cooperativity of CRM1 export complex formation. *Proc Natl Acad Sci U S A* 110:960-965.

26. Grandi P*, et al.* (1997) Nup93, a vertebrate homologue of yeast Nic96p, forms a complex with a novel 205-kDa protein and is required for correct nuclear pore assembly. *Mol Biol Cell* 8:2017-2038.

27. Andersen KR*, et al.* (2013) Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* 2:e00745.

28. Conti E & Kuriyan J (2000) Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure* 8:329-338.

29. Matsuura Y & Stewart M (2005) Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. *EMBO J* 24:3681-3689.

30. Lusk CP, Makhnevych T, Marelli M, Aitchison JD, & Wozniak RW (2002) Karyopherins in nuclear pore biogenesis: a role for Kap121p in the assembly of Nup53p into nuclear pore complexes. *J Cell Biol* 159:267-278.

31. Kobayashi J & Matsuura Y (2013) Structural basis for cell-cycle-dependent nuclear import mediated by the karyopherin Kap121p. *J Mol Biol* 425:1852-1868.

32. Hoelz A, Nairn AC, & Kuriyan J (2003) Crystal structure of a tetradecameric assembly of the association domain of Ca2+/calmodulin-dependent kinase II. *Mol Cell* 11:1241-1251.

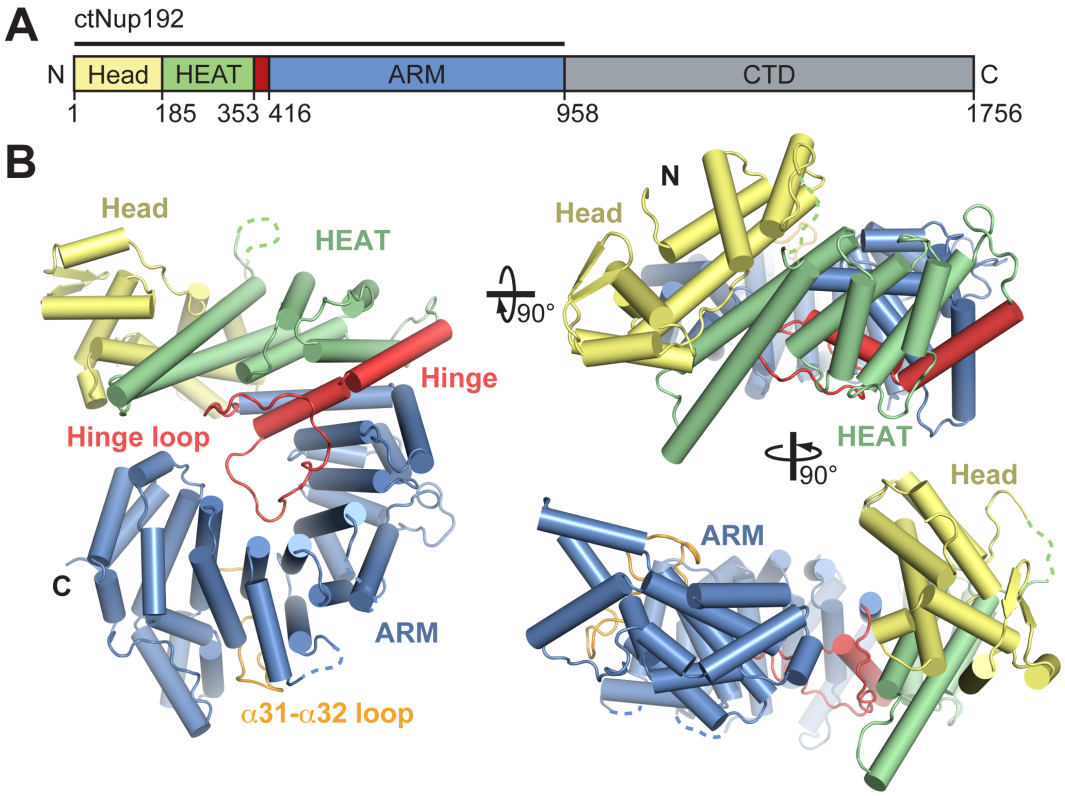


Fig. 1.
Structural overview of the N-terminal domain of ctNup192. (*A*) Domain structure of *C. thermophilum* Nup192 (crystallized fragment indicated by a black bar). The N-terminal Head module (yellow), the HEAT module (green), the hinge module (red), the ARM module (blue), and the C-terminal domain (gray) are indicated. (*B*) Structure of ctNup192^{NTD} shown in cartoon representation using coloring scheme as in (*A*). The hinge and α31-α32 loops, which line the inner arch of the ring are colored in red and orange, respectively. Dashed lines indicate disordered loops.

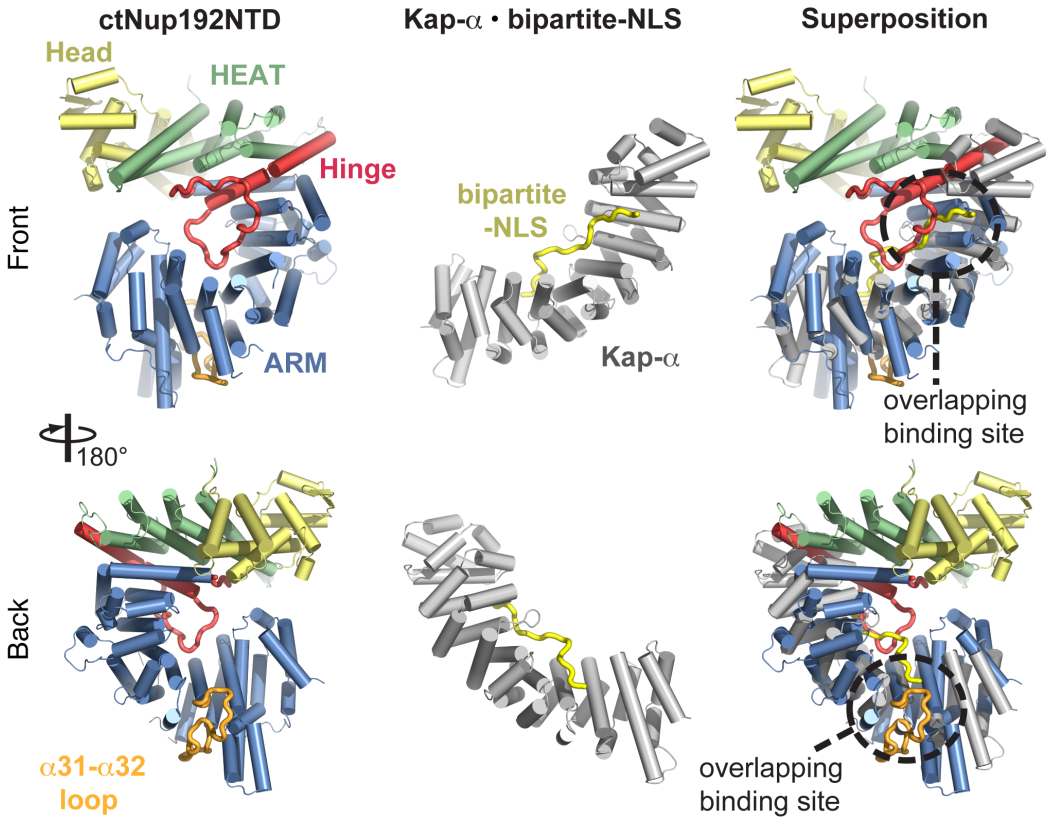


Fig. 2.

Structural similarity of the ctNup192^{NTD} ARM module to Kap-α. Overview of ctNup192^{NTD}, colored as in Fig. 1*B*, the Kap-α•bipartite-NLS complex (gray and yellow) (PDB 1EE5) (28), and their superposition on the right.

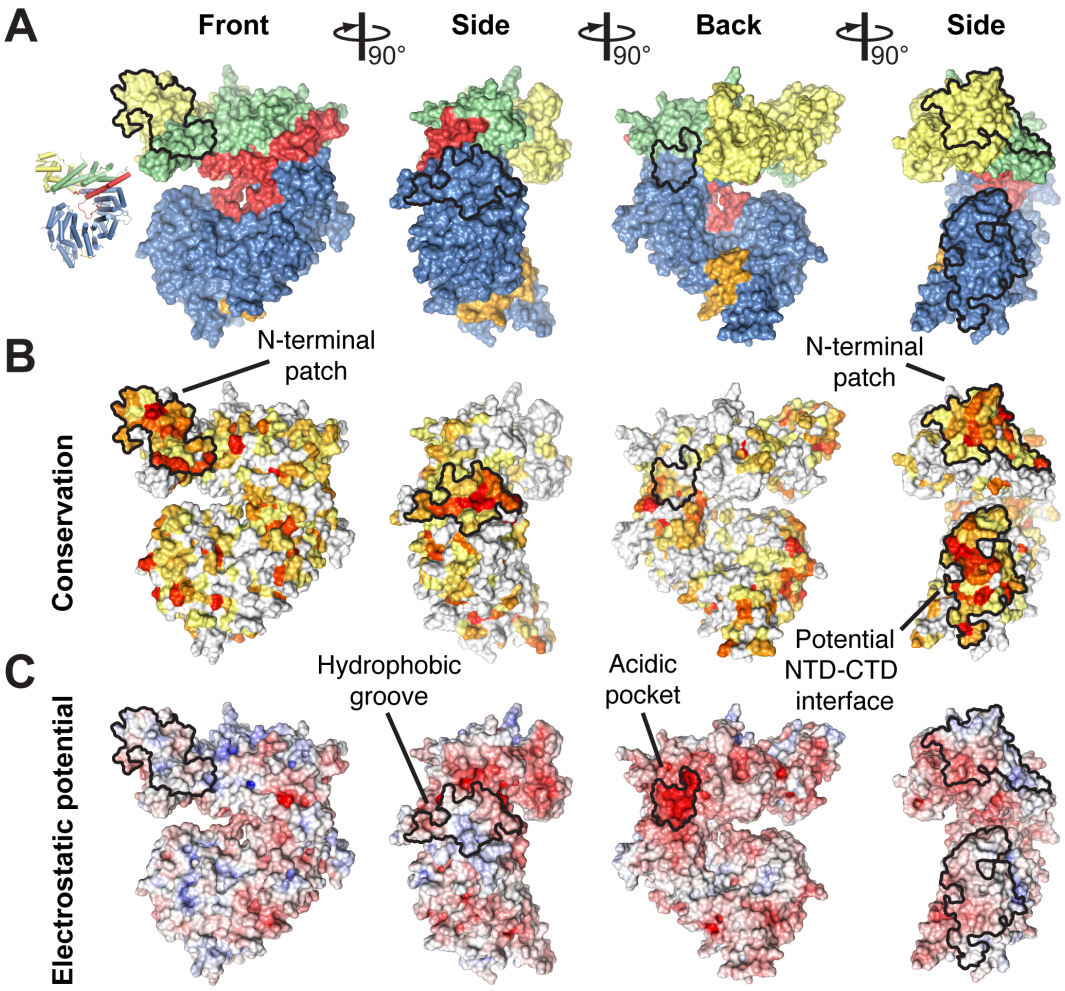


Fig. 3.
Surface properties of ctNup192^{NTD}. (*A*) Surface representation of ctNup192^{NTD} in four different orientations colored as in Fig. 1*B*. (*B*) Surface representation of ctNup192^{NTD} colored according to conservation within 7 fungal Nup192 sequences (see also, Fig. S2). Sequence conservation is shaded from white (< 45 % similarity) to yellow (45 % similarity) to red (100 % identity). (*C*) Surface representation colored according to electrostatic potential from -10 k_BT/e (red) to +10 k_BT/e (blue).

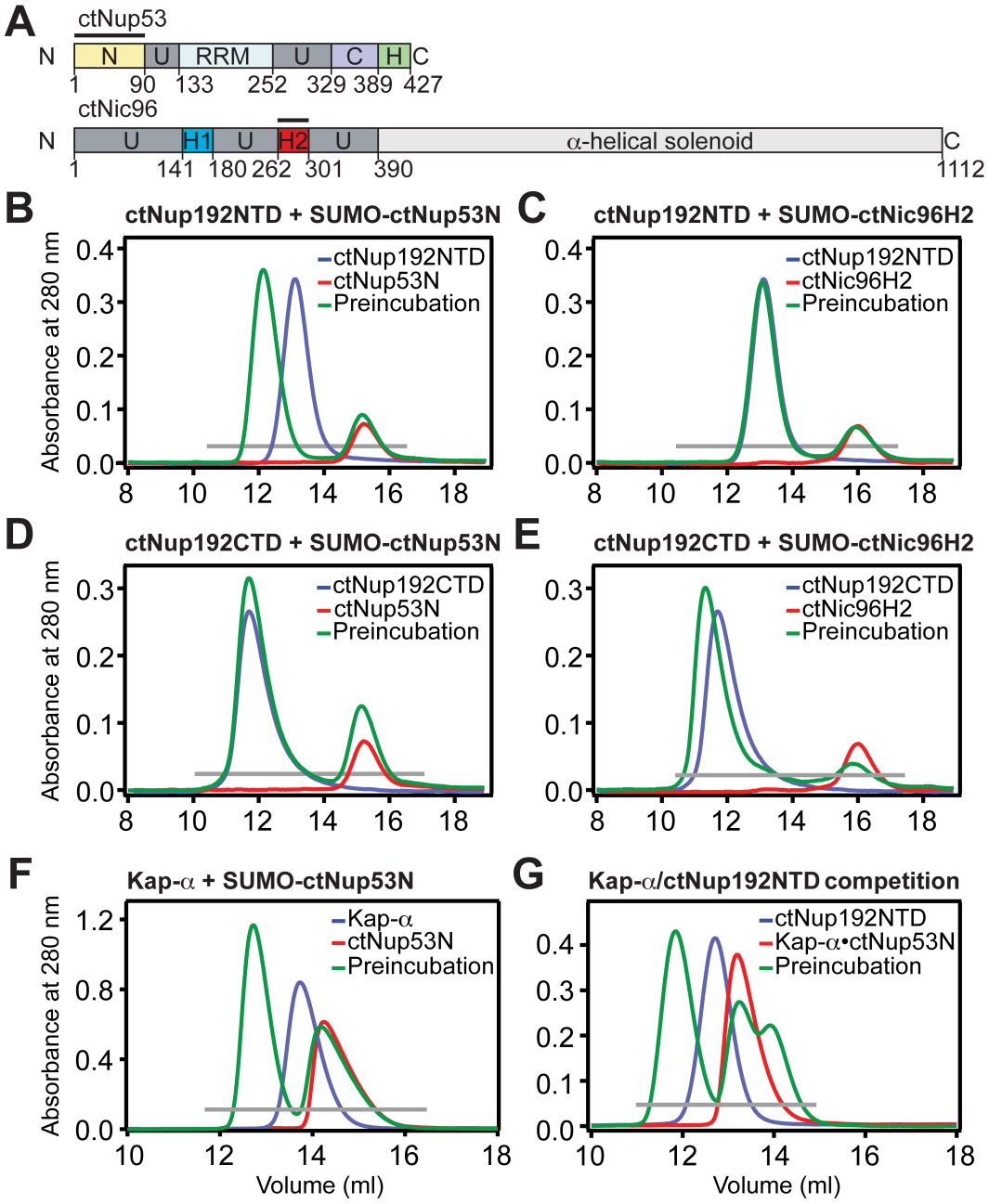


Fig. 4.

Biochemical interaction analysis. (A) Domain organization of ctNup53 and ctNic96. Black bars highlight fragments used for interaction analysis. (*B, C*) SEC analysis of ctNup192^{NTD} with (*B*) ctNup53^{N} or (*C*) ctNic96^{H2}. (*D, E*) SEC analysis of ctNup192^{CTD} with (*D*) ctNup53^{N} or (*E*) ctNic96^{H2}. (F) SEC analysis of Kap-α with ctNup53^{N}. (G) SEC analysis of competitive binding between ctNup53^{N} and Kap-α and ctNup192^{NTD}. Gray bars and colored lines designate the analyzed fractions in the respective SDS-PAGE gels stained with Coomassie brilliant blue (Fig. S5*B-E*, *G*, *H*).

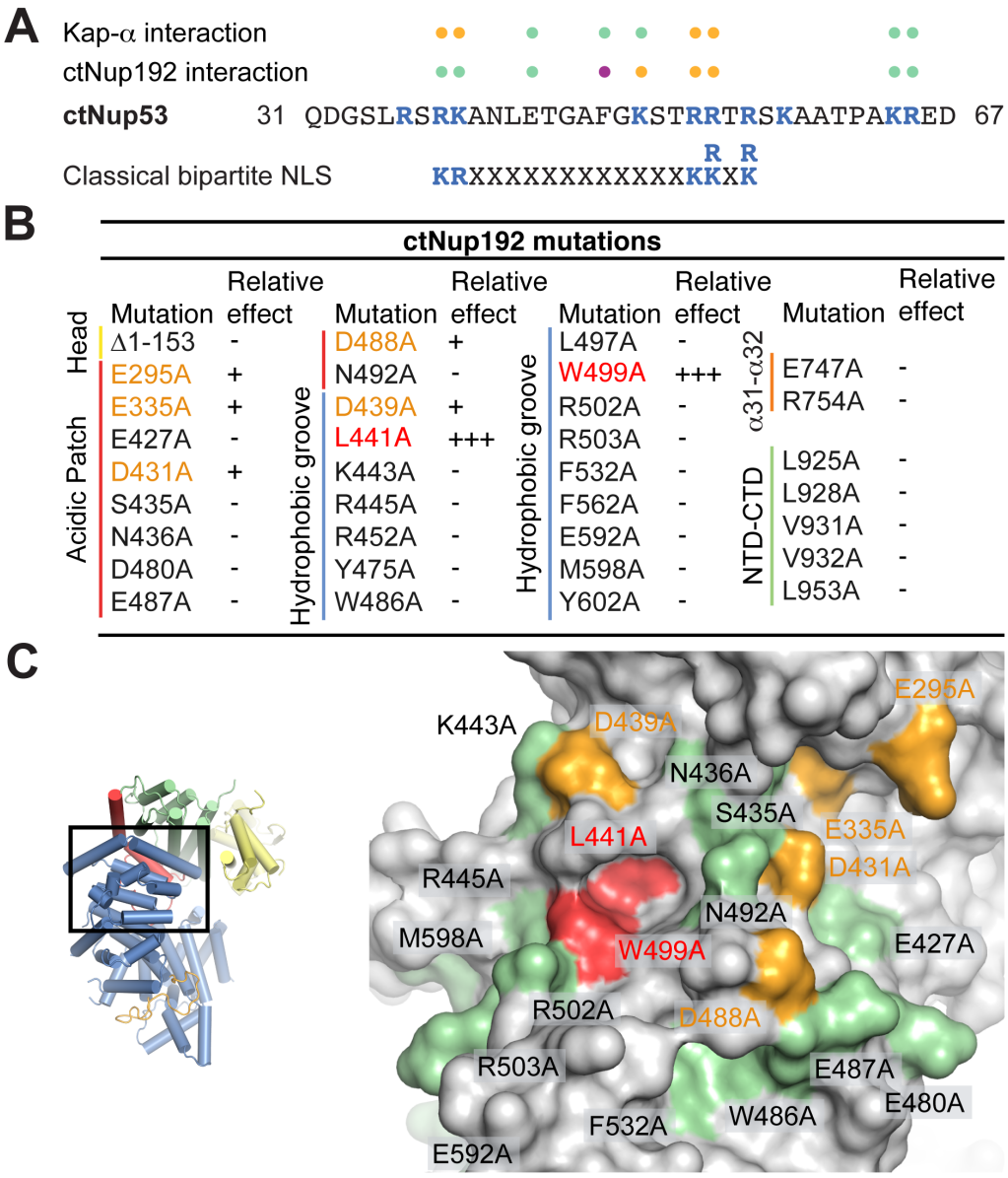


Fig. 5.

Mutational analysis of the ctNup53 interactions with ctNup192 and Kap-α. (*A*) Sequence comparison of ctNup53$^{31\text{-}67}$ with a consensus classical bipartite NLS and summary of the effects of ctNup53$^{31\text{-}67}$ mutations on ctNup192^{NTD} and Kap-α binding. Basic residues are highlighted in blue, dots indicate mutations that cause no effect (green), reduced binding (orange), or complete disruption (purple). See also Fig. S6*A*, *B*. (*B*) Mutational analysis of the ctNup192^{NTD}-ctNup53$^{31\text{-}67}$ interaction. Mutations that have no effect (-, black), reduce binding (+, orange), or abolish the interaction (+++, red) are indicated. See also Fig. S7*A*. (*C*) Surface mapping of ctNup192^{NTD} mutagenesis results. As reference, a cartoon representation of ctNup192^{NTD}, colored as in Fig. 1*B* is shown on the left. The locations of mutations on the ctNup192^{NTD} surface are labeled and colored in green, orange, and red to indicate no effect, reduced binding, or complete disruption of ctNup53$^{31\text{-}67}$ binding, respectively.

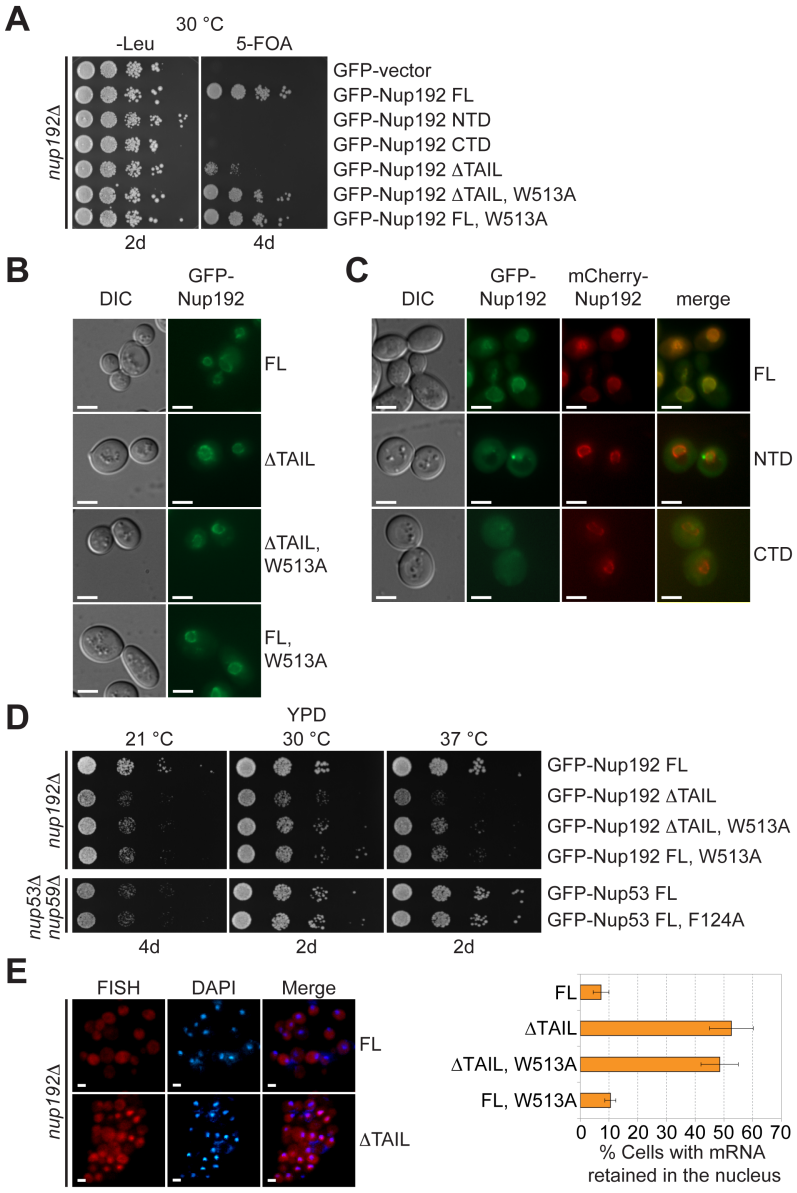


Fig. 6.

***In vivo* analyses of Nup192 and Nup53 mutants in *S. cerevisiae*.** (*A*) The *nup192Δ* shuffle strain containing a mCherry-Nup192 cover plasmid (URA3) was transformed with control or GFP-Nup192 variants (Leu). Growth was followed on SDC-Leu and 5-FOA/SDC plates for the indicated times and temperatures. (*B,C*) *In vivo* localization of GFP-Nup192 variants in the *nup192Δ* strain visualized by fluorescence and differential interference contrast (DIC) microscopy. Cells in (*C*) still carry the mCherry-Nup192 full-length plasmid. (*D*) Growth analysis of *nup192Δ* and *nup53Δnup59Δ* strains. Cells were spotted on YPD plates and grown for 2-4 days at the indicated temperatures. (*E*) mRNA export assay of GFP-Nup192 variants. Representative images of wildtype GFP-Nup192 (top) and GFP-Nup192$^{\Delta TAIL}$ (bottom) complemented *nup192Δ* cells are shown. Quantification of nuclear poly(A) mRNA retention is shown on the right. Error bars indicate standard deviations derived from six independent images, each containing ~100 cells. Scale bars are 5 μm.

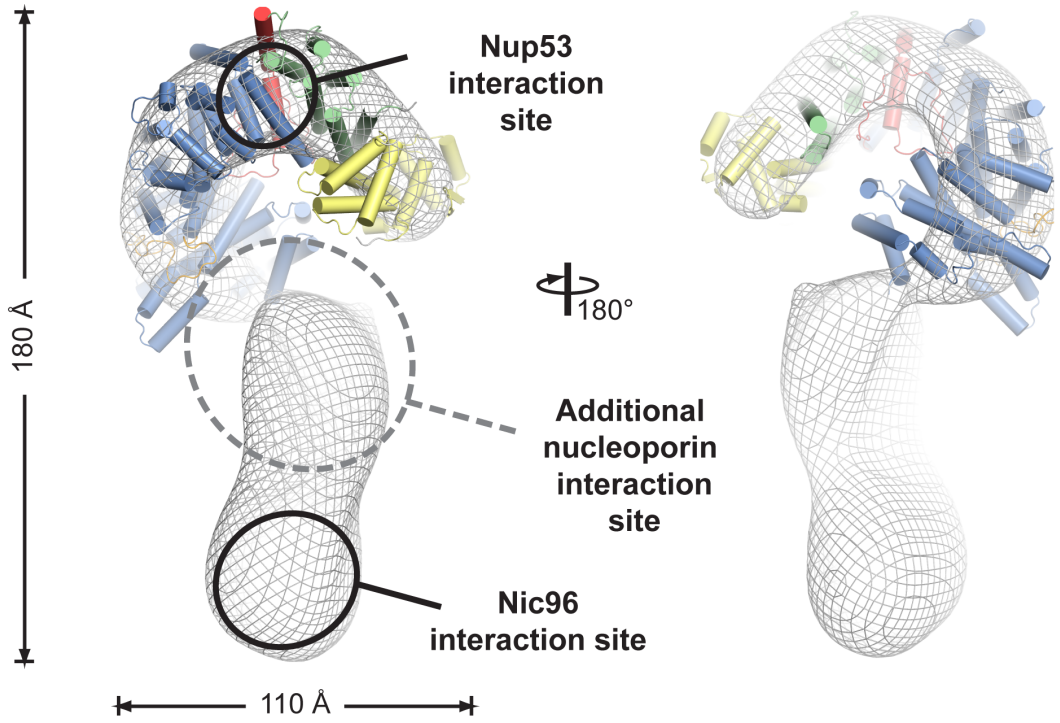


Fig. 7.
Model for the ctNup192 interaction network. The structure of ctNup192^{NTD} is docked into the EM envelope of full-length ctNup192 (21). General location of the interaction sites for Nup53$^{31-67}$ (top) and Nic96^{H2} (bottom) on Nup192 are shown. The location of a yet to be identified nup interaction site is indicated.

SUPPORTING TEXT

ARM and HEAT repeat analysis

To further investigate the similarities between the ctNup192^{NTD} ARM repeat module and Kap-α, we superposed the individual ARM repeats of ctNup192^{NTD} and compared them with a canonical ARM repeat of Kap-α (Fig. S3*B*). ARM repeats are α-helical sequence motifs consisting of three helices, termed H1, H2, and H3, which are arranged in a triangular pattern that constitutes one turn of a right-handed superhelix (1). All ARM motifs in ctNup192^{NTD} superposed well with the canonical Kap-α ARM repeat with RMSD values ranging from 1.1 to 4.2 Å. However, the superposition also revealed that the ARM repeats of ctNup192^{NTD} are far more irregular than the ARM repeats of Kap-α. Whereas the ARM repeats of Kap-α are all ~40 residues in length and can be superposed with very little variation, the ARM repeats of ctNup192^{NTD} occasionally contain long loop decorations or slightly shorter helices, resulting in ARM repeats that range in length from 36 to 83 residues. Based on the structural superposition, we generated a sequence alignment of the ARM repeats and compared it to a recently determined consensus sequence (1). There are thirteen positions in canonical ARM repeats where hydrophobic residues are greatly preferred over hydrophilic residues (> 90 %): one in H1, five in H2, and seven in H3. While the positions in H1 and H3 are mainly conserved in the ctNup192^{NTD} ARM repeats, H2 is much more divergent, with greater variance in helical length, position, and sequence (Fig. S3*B*). As a result of these deviations, the angles between the helices also vary more than in canonical ARM repeat proteins.

Similarly, the HEAT module could generally be superposed with other HEAT repeat containing proteins such as CRM1 and a structure-based sequence alignment reveals that the consensus hydrophobic positions in helices αA and αB are conserved, as identified previously (1). There are nine positions in canonical HEAT repeats where hydrophobic

residues are greatly preferred, and we found that these positions were largely conserved in the HEAT repeats of ctNup192^{NTD} (Fig. S3*C*). Whereas HEAT repeats 2 and 3 have relatively normal helical lengths, HEAT repeat 1 is unusual in that its helices are 34 and 28 residues long, compared to the 13 and 17 residues observed in canonical HEAT repeats. This feature is evolutionarily conserved in fungal Nup192 proteins (Fig. S2). The N-terminal half of helix α9 and the C-terminal half of α10 participate in the HEAT superhelix, while the rest of the helices protrude from the structure (Fig. 1*B*).

Conformational plasticity

Many extended α-helical solenoids, including members of the karyopherin-α family, exhibit extensive conformational flexibility (2-5). When we performed a structural superposition of the two ctNup192^{NTD} molecules in the asymmetric unit, we observed two different conformations (Fig. S4). The two molecules can be superposed with a RMSD of 1.2 Å over 778 Cα atoms, but the N-terminal and C-terminal halves can be superposed separately with RMSD values of 0.3 Å over 259 Cα atoms and 0.6 Å over 403 Cα atoms, respectively. The gap between the N-terminal HEAT module and the C-terminal ARM module in the two ctNup192^{NTD} structures differs by ~4 Å as a result of a rigid body rotation of the N-terminal Head and HEAT modules away from the C-terminal ARM module (Fig. S4*A*). This conformational change is mediated by the hinge module, which includes helices α15 and α16 and the long hinge loop that follows these helices and caps the HEAT and ARM modules (Fig. S4*C*).

Further conformational changes were apparent when the structures of ctNup192^{NTD} were compared to the recently determined structure of *S. cerevisiae* Nup192^{NTD} (6). scNup192^{NTD} possesses a similar overall architecture and can be superposed onto

ctNup192^{NTD} with a RMSD of 3.5 Å over 646 Cα atoms. However, the N-terminal half of the molecule is rotated away from the C-terminal half, resulting in a gap between the two halves that opens an additional ~5 Å to a total distance of ~18 Å (Fig. S4*A*). This is most apparent when the N- and C-terminal halves of the ring are superposed separately, which results in substantially lower RMSD values of 2.9 Å over 232 Cα atoms and 2.3 Å over 375 Cα atoms, respectively. Moreover, the hinge axis is not parallel to the equatorial plane of the Nup192^{NTD} ring, and facilitates not only an increase of the ring gap, but also a rotation of the Head and HEAT modules by ~26º out of the equatorial plane of the ring (Fig. S4*A*). As such, the observed conformational changes are more similar to the opening of a lock-washer than the opening of a clamp. Again, the hinge module mediates these conformational changes, but surprisingly, the scNup192^{NTD} hinge loop adopts a substantially different conformation and no longer contacts the C-terminal three ARM repeats (Fig. S4*D*).

The conformational changes of the hinge loop are very similar to the observed relocation of the acidic loop in the export β-karyopherin CRM1 (4, 5, 7). Like the hinge loop of ctNup192^{NTD}, the acidic loop of CRM1 displays species dependent variation in sequence and length. Furthermore, the CRM1 acidic loop occupies a similar position within the CRM1 ring as the hinge loop does in the ctNup192^{NTD} ring and also makes extensive contacts within the concave surface (4, 5, 7). While the conformational changes observed here for ctNup192^{NTD} are not as dramatic, it is conceivable that they nevertheless play an important role in regulating the interactions with other adaptor nucleoporins.

Together, these observations provide further evidence for an evolutionary relationship between Nup192 and the flexible β-karyopherins.

Methods

Protein Expression and Purification. DNA fragments encoding *C. thermophilium* Nup192^{NTD} (residues 1 to 958), Nup192^{CTD} (residues 976 to 1756), Nup192$^{\Delta HEAD}$ (residues 153-958), and Nup192^{TAIL} (residues 1416-1756) were amplified by PCR and cloned into a modified pET28a vector, which contains an N-terminal hexahistidine tag followed by a PreScission protease cleavage site, using NdeI and NotI restriction sites (8). DNA fragments encoding residues 1 to 90 and 31 to 67 of ctNup53 and residues 262 to 301 of ctNic96 were cloned into modified pET28a or pET-MCN vectors containing an N-terminal hexahistidine-SUMO tag, using BamHI and NotI restriction sites (9, 10). A DNA fragment encoding *C. thermophilum* Nup188^{NTD} (residues 1 to 1134) was cloned into the modified pET28a vector with an N-terminal hexahistidine-SUMO tag, using AseI and BamHI restriction sites. The *S. cerevisiae* Kap-α expression construct was a kind gift from Elena Conti (11). The details of the bacterial expression constructs are listed in Table S2.

All proteins were expressed in *E. coli* BL21-CodonPlus (DE3)-RIL cells (Stratagene) in Terrific Broth media. Seleno-L-methionine-labeled protein was produced in a synthetic medium that suppresses methionine biosynthesis, following standard protocols (12). ctNup192 and ctNup188 fragment expression was induced at an OD_{600} of 0.6 with 0.5 mM IPTG at 37 °C for 3 hours. Expression of ctNup53 and ctNic96 fragments was induced at an OD_{600} of 0.8 with 0.5 mM IPTG at 18 °C for 16 hours. Cells were harvested by centrifugation and resuspended in a buffer containing 20 mM TRIS, pH 8.0, 500 mM NaCl, 15 mM imidazole, 4 mM β-mercaptoethanol (β-ME), and complete EDTA-free protease inhibitor cocktail (Roche).

For purification of all proteins, the cells were lysed with a cell disruptor (Avestin) and DNAse I (Roche) was added to the lysate before centrifugation at 30,000 x g for 1 hour.

The supernatant was filtered through a 0.45 μm filter (Millipore) and loaded onto a Ni-NTA column (GE Healthcare) equilibrated in buffer A (20 mM TRIS, pH 8.0, 500 mM NaCl, 15 mM imidazole, and 4 mM β-ME). Protein was eluted with a linear gradient of buffer B (20 mM TRIS, pH 8.0, 500 mM NaCl, 500 mM imidazole, and 4 mM β-ME). Protein-containing fractions were pooled and incubated overnight with PreScission or ULP1 protease at 4 °C while dialyzing against buffer A. Digested protein was loaded onto a Mono Q 10/100 GL ion-exchange column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM DTT. Protein was eluted using a linear gradient of a buffer containing 20 mM TRIS, pH 8.0, 2.0 M NaCl, and 5 mM DTT, concentrated in a centrifugal filter (Millipore), and loaded on a HiLoad Superdex 200 16/60 gel filtration column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM DTT. Protein-containing fractions were pooled and concentrated to 20 mg/ml for crystallization or biochemical studies.

Fractions from Ni-NTA elution containing ctNup53 or ctNic96 SUMO-fusion proteins were dialyzed against a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM DTT and loaded onto a Mono Q 10/100 GL column, eluted with a linear gradient of a buffer containing 20 mM TRIS, pH 8.0, 2.0 M NaCl, and 5 mM DTT, concentrated in a centrifugal filter (Millipore), and loaded onto a Superdex 200 10/300 GL gel filtration column (GE Healthcare). Protein containing-fractions were pooled and concentrated for biochemical studies.

ctNup53 and ctNup192 mutants were generated by QuikChange mutagenesis, confirmed by DNA sequencing, and expressed and purified as the wildtype proteins. The ctNup192^{TAIL}•ctNic96^{H2} complex was generated by co-expression of the two proteins and

purified with the same protocol as ctNup192^{TAIL}. *S. cerevisiae* Kap-α was expressed and purified as previously described (11).

Protein Crystallization and Data Collection. Protein crystallization was carried out at 21 °C in hanging drops consisting of 1.0 μl protein solution and 1.0 μl reservoir solution. Crystals appeared in the tetragonal space group $P4_32_12$ with two molecules in the asymmetric unit. These crystals were improved by microseeding, which produced crystals that grew to maximum dimensions of ~100×100×300 μm^3 in 1 week. Crystals used for diffraction experiments were grown in 0.1 M MES, pH 5.7, 0.6 M $MgCl_2$, and 5 % (w/v) PEG 4000 with a protein concentration of 20 mg/ml. Seleno-L-methionine labeled crystals were grown in identical conditions. Native crystals we derivatized in the crystallization drop by adding 0.1 μl of a saturated $[Ta_6Br_{12}]^{2+}$ cluster solution, followed by 16 hour incubation prior to freezing. Crystals were cryoprotected by gradually supplementing the drop in 2 % steps with 24 % ethylene glycol and flash frozen in liquid nitrogen. X-ray diffraction data was collected at 100 K at Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL).

Structure Determination and Refinement. X-ray diffraction data was processed with the HKL2000 denzo/scalepack package and XDS (13, 14). Initial phases were calculated in Phaser using single anomalous dispersion X-ray diffraction data obtained from a Ta_6Br_{12} cluster derivative. These phases were used to locate 59 selenium atoms in anomalous X-ray diffraction data obtained from a seleno-L-methionine-labeled crystal (15). Solvent flattening and NCS averaging was performed in Resolve to improve phases of the seleno-L-methionine-labeled derivative (16, 17). The experimental map was of excellent quality and allowed for unambiguous placement of all helices and sequence assignment, aided by the

positions of the selenium atoms (Fig. S1). Iterative rounds of model building and refinement were performed with COOT and PHENIX (16, 18). Initial rounds of refinement were performed with NCS restraints and individual isotropic B-factor refinement. Final refinement rounds were performed without NCS restrains, with hydrogen atoms as riding atoms, and with TLS groups, identified by TLSMD (19). The final model was refined to 2.7 Å resolution with R_{work} and R_{free} values of 19.1 % and 23.1 %, respectively. No density was observed for residues 174 to 180, 569 to 589, and 680 to 698 and for residues 64 to 66, 170 to 181, 537 to 547, 567 to 587, 678 to 698, 804 to 820, and 894 to 916 for the first and second molecule in the asymmetric unit, respectively. These residues are presumed to be disordered and have been omitted from the final model. The stereochemical quality was assessed with PROCHECK and MolProbity and there were no Ramachandran outliers detected by either program (20, 21). For details of the data collection and refinement statistics see Table S1.

Multiangle Light Scattering. Purified ctNup192^{NTD} was characterized by multiangle light scattering following size-exclusion chromatography (22). 750 μg of ctNup192^{NTD} was injected onto a Superdex 200 10/300 GL gel filtration chromatography column equilibrated in a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM DTT. The chromatography system was connected in series with an 18-angle light-scattering detector (DAWN HELEOS II, Wyatt Technology), a dynamic light-scattering detector (DynaPro Nanostar, Wyatt Technology), and a refractive index detector (Optilab t-rEX, Wyatt Technology). Data were collected every 1 s at a flow rate of 0.5 ml/min at 25 °C. Data analysis was carried out using the program ASTRA 6, yielding the molar mass and mass distribution (polydispersity) of the sample.

Analytical Size-Exclusion Chromatography. Protein interaction experiments were carried out on a Superdex 200 10/300 GL gel filtration column equilibrated in a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM DTT. Three-fold molar excess of N-terminal SUMO-fused ctNic96^{H2} or ctNup53^{N} was mixed with ctNup192^{NTD} or ctNup192^{CTD} and incubated for 30 min on ice. In the case of the interaction analysis for ctNup53^{N} with Kap-α four-fold molar excess of the N-terminal SUMO-fused ctNup53^{N} was mixed with Kap-α and incubated for 30 min on ice. Complex formation was monitored by injection of the pre-incubated proteins or the individual components onto the gel filtration column. The ctNup53$^{31-67}$ competition experiment was performed by pre-incubating ctNup192^{NTD} with a purified, stoichiometric complex of Kap-α•ctNup53^{N}. Interaction tests using ctNup53$^{31-67}$ and ctNup192^{NTD} variants were performed similarly. To assay the interaction between ctNup192^{TAIL} and ctNic96^{H2}, equal amounts of ctNup192^{TAIL} or the purified ctNup192^{TAIL}•ctNic96^{H2} heterodimer were injected onto the gel filtration column. All proteins were analyzed under identical buffer conditions and complex formation was confirmed by SDS-PAGE of the protein-containing fractions, followed by Coomassie brilliant blue staining.

Isothermal Titration Calorimetry. ITC measurements were performed at 21 °C using a VP-ITC calorimeter (GE Healthcare) and consisted of 30 injections of 10 µl with a 180 s spacing. Reference power was 10 µcal/s for titrations with ctNup192^{NTD} and 20 µcal/s for titrations with Kap-α. For titrations with ctNup192^{NTD} variants, 200 µM ctNup53$^{31-67}$ was injected into 10 µM ctNup192^{NTD}. For titrations with Kap-α, 1.5 mM ctNup53$^{31-67}$ was injected into 150 µM Kap-α. Titrations using wildtype proteins were performed in triplicate.

Heat from dilution was subtracted for baseline correction. All data was analyzed using Origin 7.0 software with MicroCal add-ons.

Yeast Strains. The open reading frame of Nup192 in the *S. cerevisiae* haploid strain BY4741 was replaced with the HIS3 cassette by homologous recombination as previously described (23). Due to the lethality of the *NUP192* knockout, the BY4741 strain was complemented with a pRS416 construct carrying full-length *S. cerevisiae NUP192* with an N-terminal mCherry tag under the control of the *NOP1* promoter. Subsequently, pRS415-GFP constructs carrying various Nup192 variants were introduced. The transformants were selected twice on SDC-Leu plates containing 5-fluoroorotic acid (5-FOA) (Bio Gold) to ensure the loss of the pRS416-mCherry-*NUP192* construct prior to analysis. The details of the yeast expression constructs are listed in Table S2.

The strain carrying the Nup53 plasmids in a double-deletion background was generated as follows. The *NUP53* deletion was introduced into BY4741 *nup59*::kanMX4 (Open Biosystems) and covered with pRS416-mCherry-*NUP53*, resulting in the strain *nup53Δnup59Δ* (*MATα his3Δ1 leu2Δ0 ura3Δ0 nup59*::kanMX4 *nup53*::HIS3 pRS416-mCherry-*NUP53*). This strain was transformed with the plasmids pRS415-GFP-*NUP53* or pRS415-GFP-*nup53*^{F124A} and transformants were selected twice on SDC-Leu plates containing 5-FOA.

Yeast Analyses. For viability analysis, *S. cerevisiae* strains carrying GFP-Nup192 variants were grown at 30 °C to mid-log phase in SDC-Leu media and diluted to 10 million cells/mL. This stock was used to generate a 10-fold dilution series, of which 5 μL were spotted on

SDC-Leu and 5-FOA/SDC-Leu plates and grown at 30 °C for 2-4 days, respectively. For growth analysis of the shuffled strains, the same dilutions were prepared, spotted on YPD plates, and grown at 21 °C, 30 °C, and 37 °C for 2-4 days. For localization analysis, live cells were analyzed using a Carl Zeiss Observer Z.1 equipped with a Hamamatsu camera C10600 Orca-R^2.

FISH mRNA Export Assay. Liquid cultures of single-deletion yeast strains carrying GFP-fusion proteins of Nup192 were grown overnight at 30 °C in SDC-Leu media to an OD_{600} of 0.4 and subsequently shifted to 37 °C for 4 h before fixation in formaldehyde. These cells were then analyzed by FISH using an Alexa-647-labeled 50-mer oligo dT probe as previously described (24, 25). The statistical analysis was carried out using 6 independent images with at least 100 cells each.

EM Docking. The structure of ctNup192^{NTD} was manually placed into the EM envelope of full-length ctNup192, taking advantage of the published localization of the N-terminus, as determined by DID-Dyn2 labeling (26). This initial placement was then refined against the EM envelope using the rigid body refinement routine in Molrep (27).

Illustration and Figures. Sequence alignments were generated using ClustalX and colored with Alscript (28, 29). Structural figures were generated using PyMOL (www.pymol.org) and the electrostatic potential was calculated with the Adaptive Poisson-Boltzmann Solver (APBS) (30).

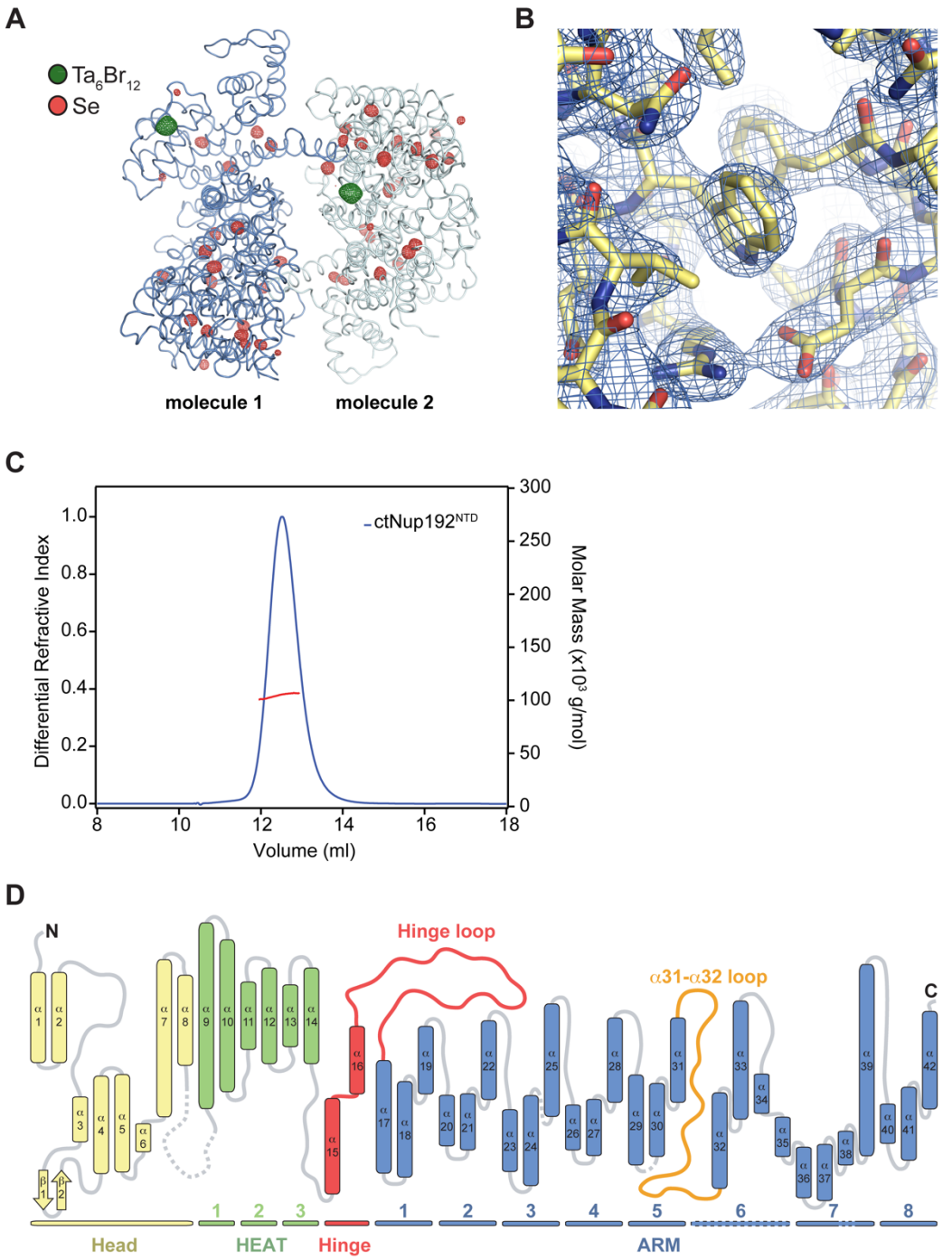


Fig. S1.
Structure determination and oligomeric state analysis of ctNup192^{NTD}. (*A*) The two molecules in the asymmetric unit are shown in a ribbon representation and colored in blue and light blue. Anomalous difference Fourier maps of the Ta_6Br_{12} cluster (green) and seleno-L-methionine labeled protein (red) derivatives are contoured at 10.0 σ and 5.0 σ, respectively. (*B*) Representative final $2|F_O|-|F_C|$ density map contoured at 1.0 σ. (*C*) SEC-

MALS analysis of ctNup192^{NTD}. The normalized differential refractive index (blue) is plotted against the elution volumes from a Superdex 200 10/300 GL gel filtration column and overlaid with the experimental molecular mass for the peak fractions (red). The determined molecular mass for ctNup192^{NTD} is 105.0 kDa (theoretical 108.8 kDa). (*D*) Schematic representation of the α-helical motifs identified in ctNup192^{NTD}. The Head module includes helices α1-α8 and a α-hairpin composed of β-strands β1-β2. The HEAT module is composed of three HEAT repeats, helices α9-α14, and is connected to the ARM module via the hinge module which contains helices α15 and α16. The ARM module contains 8 turns of a right-handed superhelix composed of helices α17-42. HEAT and ARM repeats are numbered and highlighted below the helices with dashed lines indicating deviations from canonical ARM repeats.

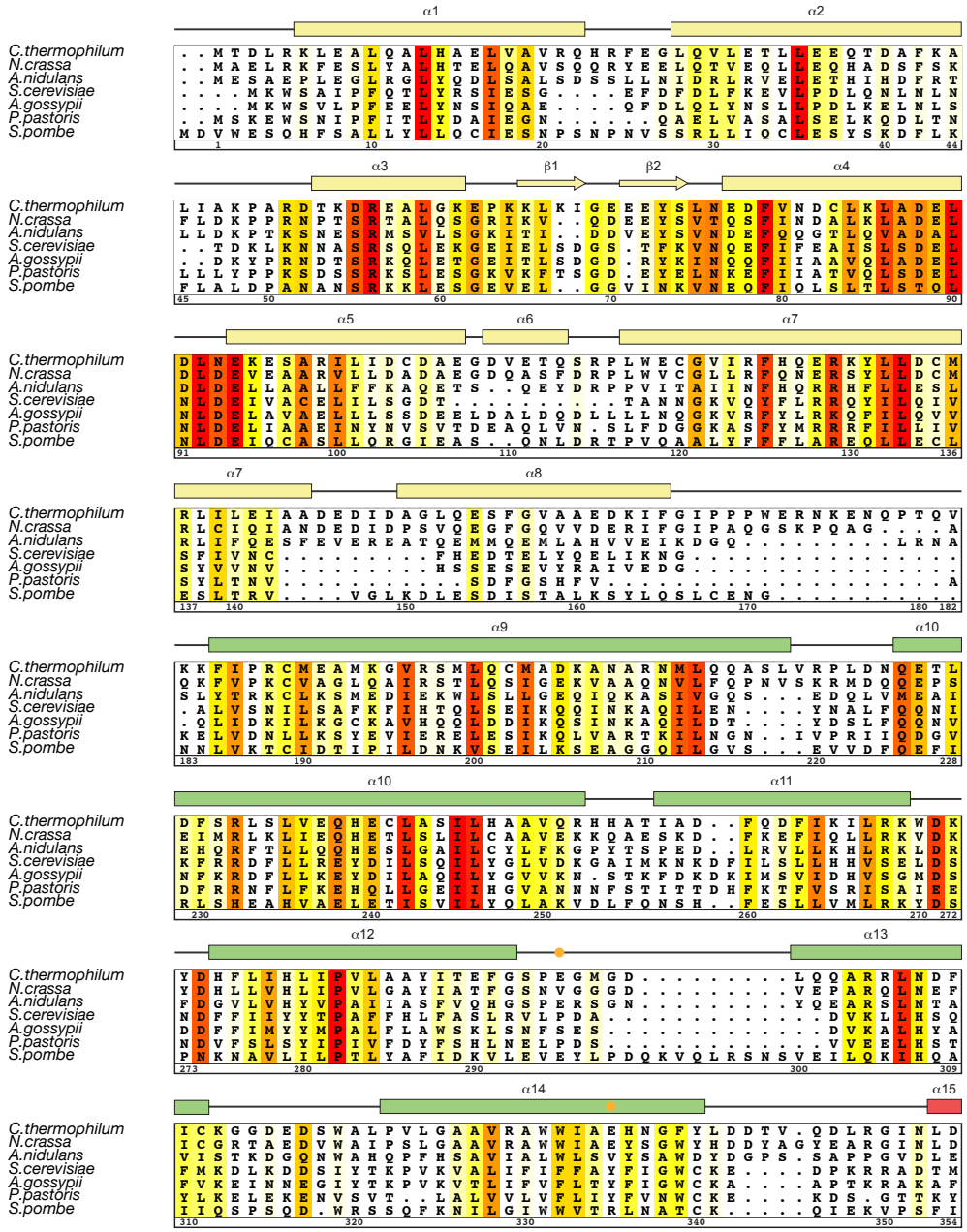


Fig. S2.

Multi-species sequence alignment of Nup192^{NTD}. Seven diverse fungal species were aligned and colored by sequence similarity according to the Blosum62 matrix from white (less than 45 % similarity) to yellow (45 % similarity) to red (100 % identity). The numbering is according to *C. thermophilum* Nup192^{NTD}. Secondary structure is indicated above the sequence as rectangles (α-helices), arrows (β-strands), and lines (unstructured regions). Secondary structure elements are colored according to Fig. 1. Dots in the secondary structure plot indicate residues that reduce (orange) or completely disrupt (red) the ctNup53 interaction upon mutation to alanine. An asterisk indicates the position of the invariant glycine 371 between hinge helices α15 and α16.

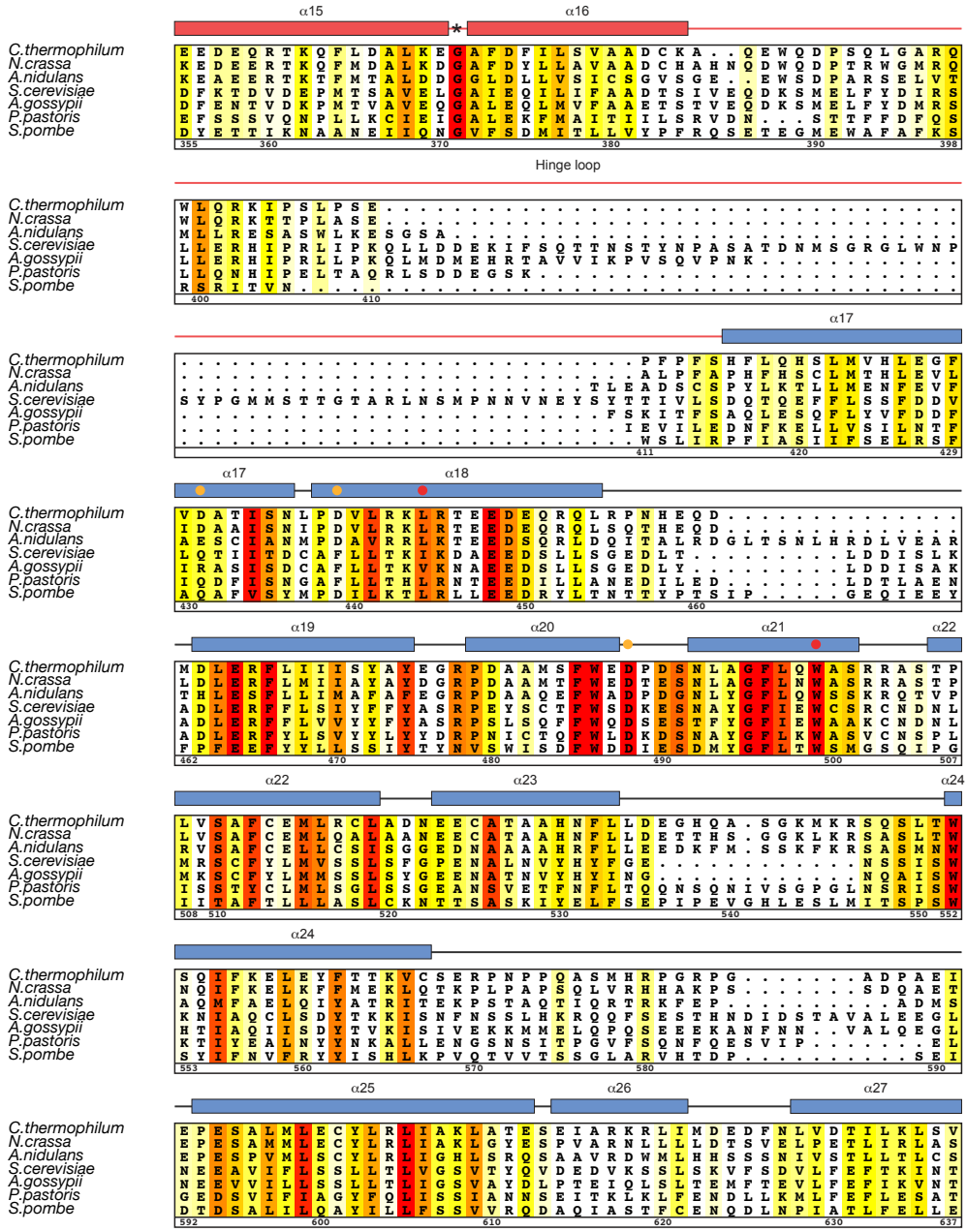


Fig. S2 continued.

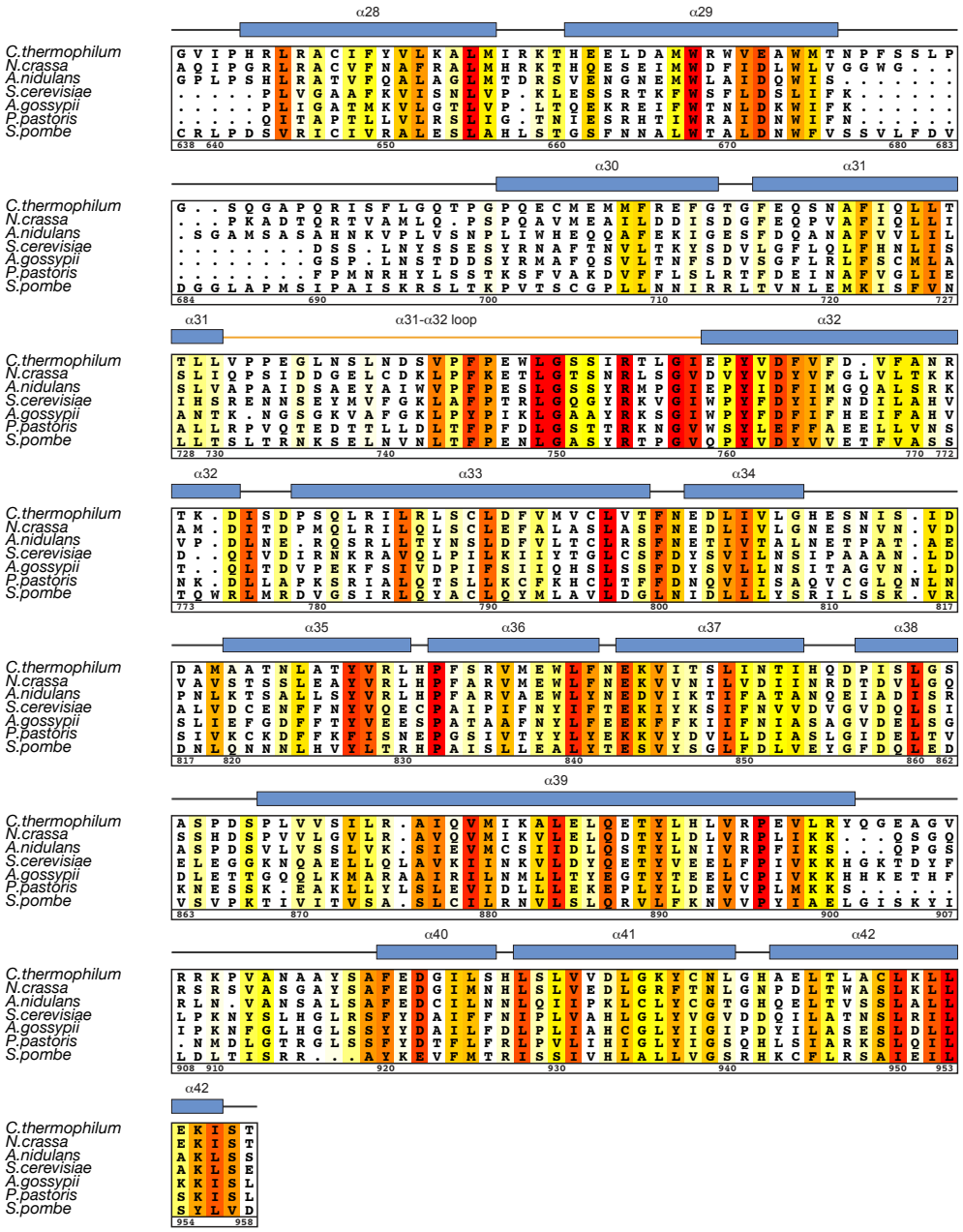


Fig. S2 continued.

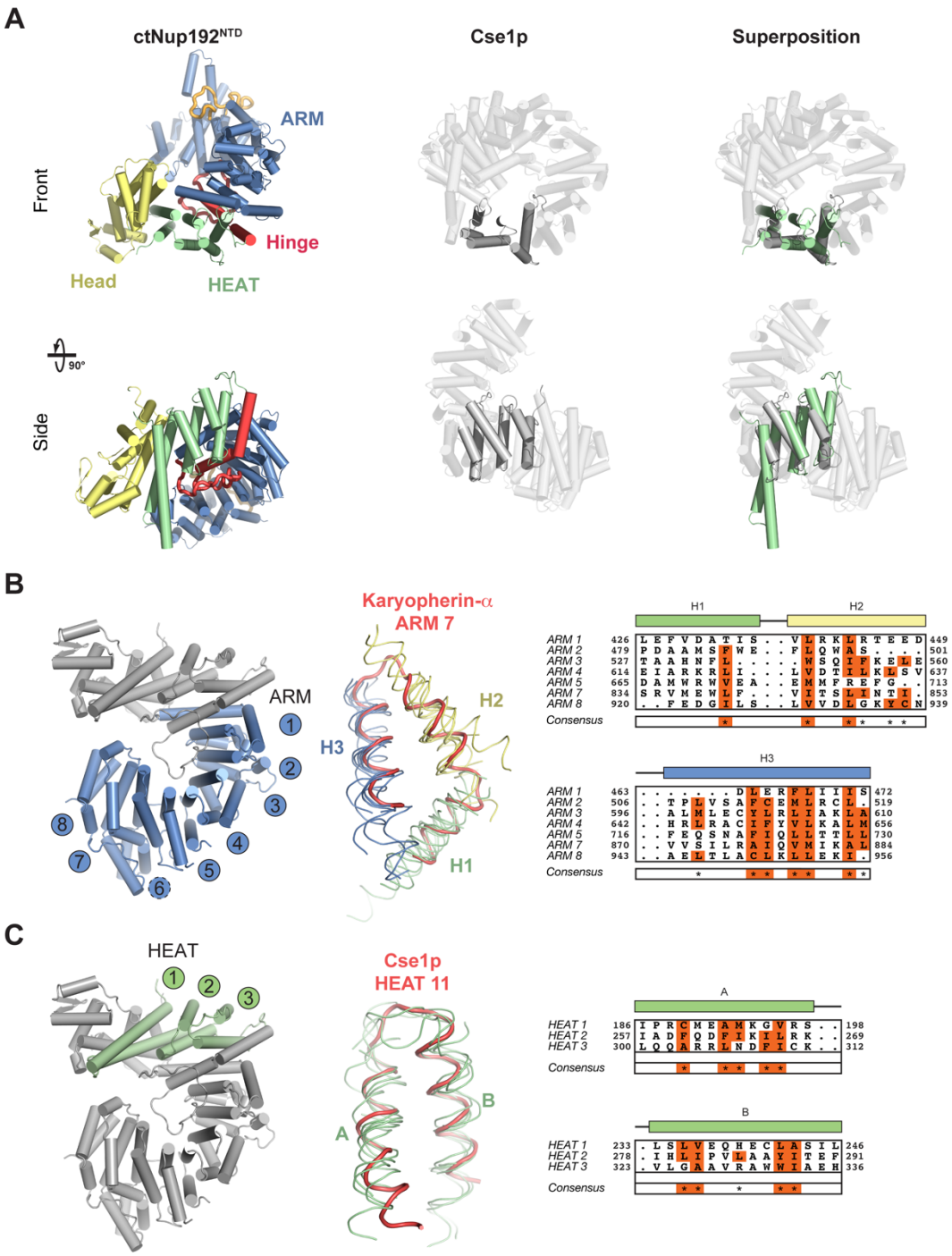


Fig. S3.

Analyses of the ARM and HEAT repeats in ctNup192^{NTD}. (*A*) Overview of the structures of ctNup192^{NTD}, Cse1p (grey, PDB 1WA5) (31), and their superposition on the right. Bottom panels depict 90º rotated views of the above structures. (*B*) An overview of the ctNup192^{NTD} structure is shown on the left with the ARM repeats highlighted in blue. The structural superposition of the ctNup192^{NTD} ARM repeats and ARM repeat 7 of Kap-α (PDB 1BK5) (11) is shown in the middle and the structure-guided sequence alignment is shown on the

right. Consensus hydrophobic positions are indicated below with asterisks and those residues that match these positions are highlighted in red. (*C*) The ctNup192^{NTD} HEAT repeats are highlighted in green on the left and their structural superposition with HEAT repeat 11 from Cse1p (PDB 1WA5) (31) is shown in the middle. A structure-guided sequence alignment is shown on the right, with consensus hydrophobic positions highlighted as in (*B*).

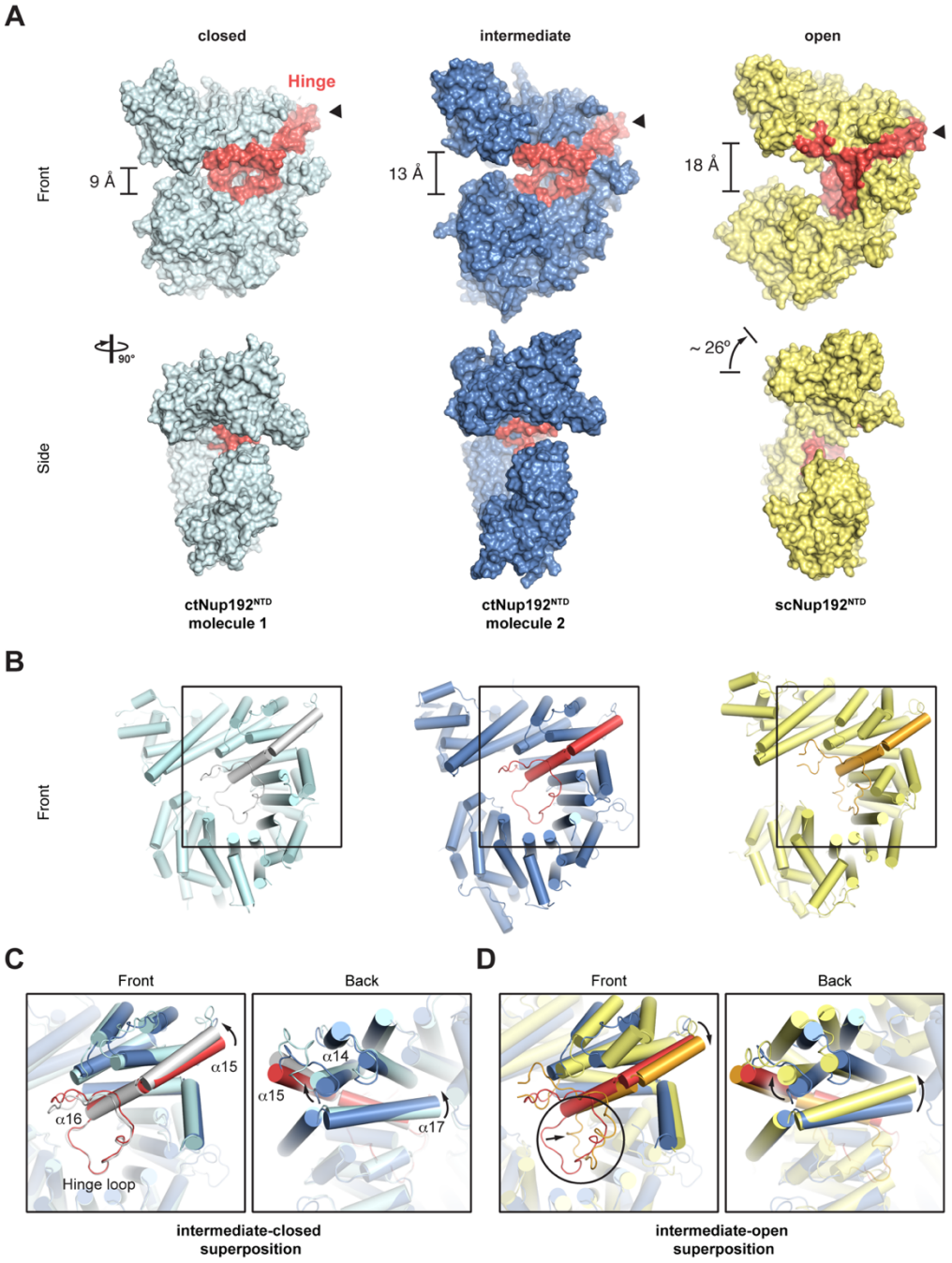


Fig. S4.

Conformational plasticity of Nup192^{NTD}. (*A*) Structures of ctNup192^{NTD} (left and middle) and scNup192^{NTD} (right, PDB 4IFQ) (6) depicted in surface representation. Their alignment on the C-terminal ARM module reveals conformational changes that are accompanied with an increased opening of the ring. The hinge module of each structure is colored in red and indicated by a black triangle. In the bottom panels, 90º rotated views show that the Head and HEAT modules also rotate along an axis out of the plane of the ring, reminiscent of the opening of a lock-washer. (*B*) Cartoon representations of ctNup192^{NTD} and scNup192^{NTD} are shown in the same orientation, coloring, and order as in (*A*). (*C*) Superposition of the closed

and intermediate states of Nup192^{NTD}. (*D*) Superposition of the intermediate and open states of Nup192^{NTD}. A large change in the conformation of the hinge loop of scNup192^{NTD}, highlighted by the black circle, is propagated to the hinge helices (α15 and α16), causing the ring to adopt a further open state. Arrows indicate the observed conformational changes of the hinge helices from the closed to the intermediate to the open states.

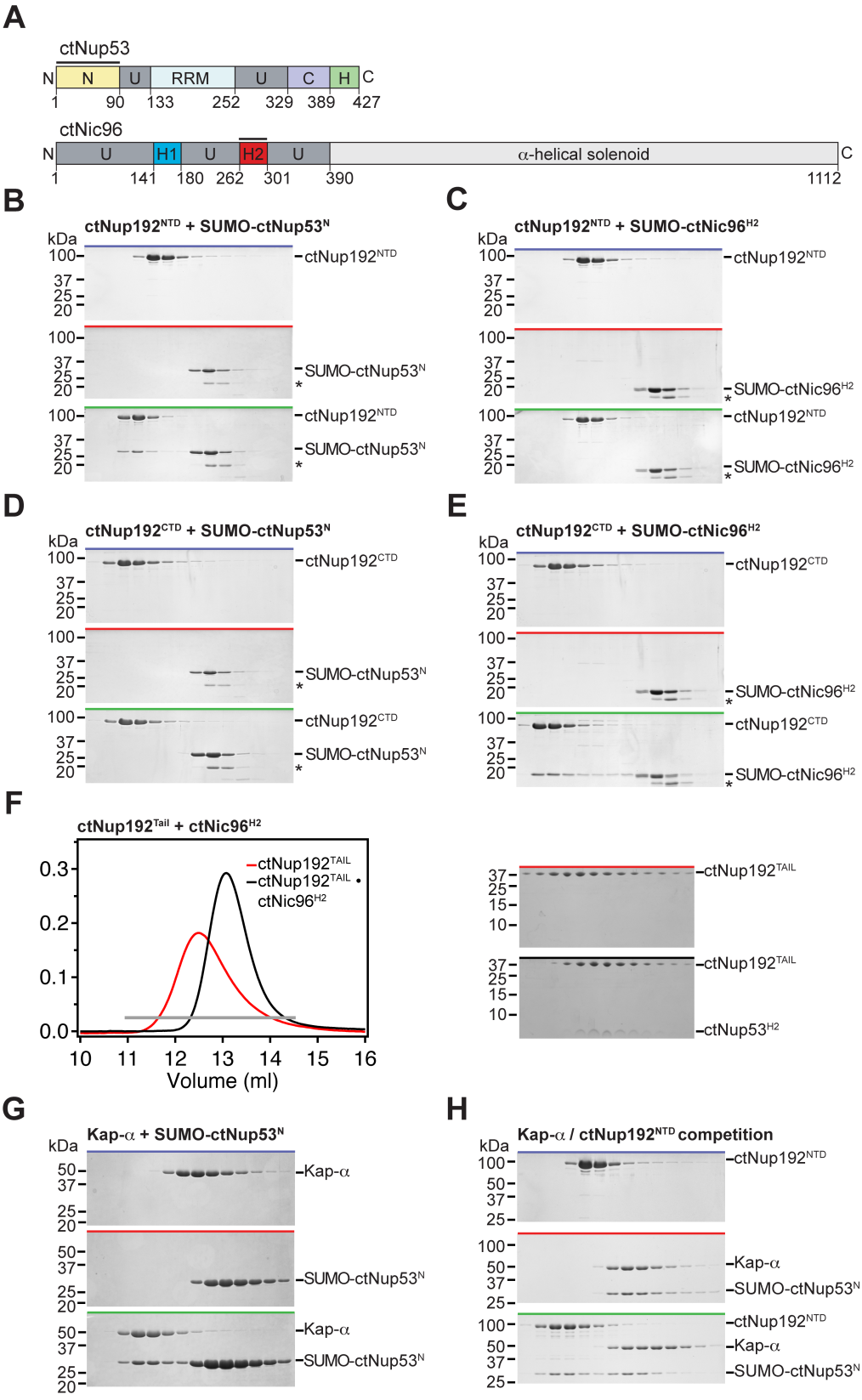

A
ctNup53
N
N
U
RRM
U
C
H
C
1
90
133
252
329
389
427
ctNic96
N
U
H1
U
H2
U
α-helical solenoid
C
1
141
180
262
301
390
1112
B
ctNup192NTD + SUMO-ctNup53N
kDa
100
37
25
20
ctNup192NTD
SUMO-ctNup53N
*
C
ctNup192NTD + SUMO-ctNic96H2
SUMO-ctNic96H2
D
ctNup192CTD + SUMO-ctNup53N
ctNup192CTD
E
ctNup192CTD + SUMO-ctNic96H2
F
ctNup192Tail + ctNic96H2
ctNup192TAIL
ctNup192TAIL •
ctNic96H2
0.3
0.2
0.1
0.0
10
11
12
13
14
15
16
Volume (ml)
37
25
15
10
ctNup53H2
G
Kap-α + SUMO-ctNup53N
50
Kap-α
H
Kap-α / ctNup192NTD competition

Fig. S5.

Biochemical analysis of ctNup192 interactions. (*A*) Schematic overview of the domain organization of *C. thermophilum* Nup53 and Nic96. (*B-E*) SDS-PAGE gels corresponding to the fractions indicated by gray bars in the gel filtration profiles of Fig. 4*B-E*. (*F*) Interaction between ctNup192^{TAIL} and ctNic96^{H2}. SEC profiles of purified ctNup192^{TAIL} alone or the purified ctNup192^{TAIL}•ctNic96^{H2} complex are shown on the left, with the corresponding SDS-PAGE gels shown on the right. The gray bar indicates the fractions analyzed. (*G, H*) SDS-PAGE gels corresponding to the fractions indicated by gray bars in the gel filtration profiles of Fig. 4*F, G*. Molecular mass standards and the positions of the proteins are indicated. Asterisks indicate degradation products of SUMO-ctNic96^{H2}. SDS-PAGE gels were stained with Coomassie brilliant blue.

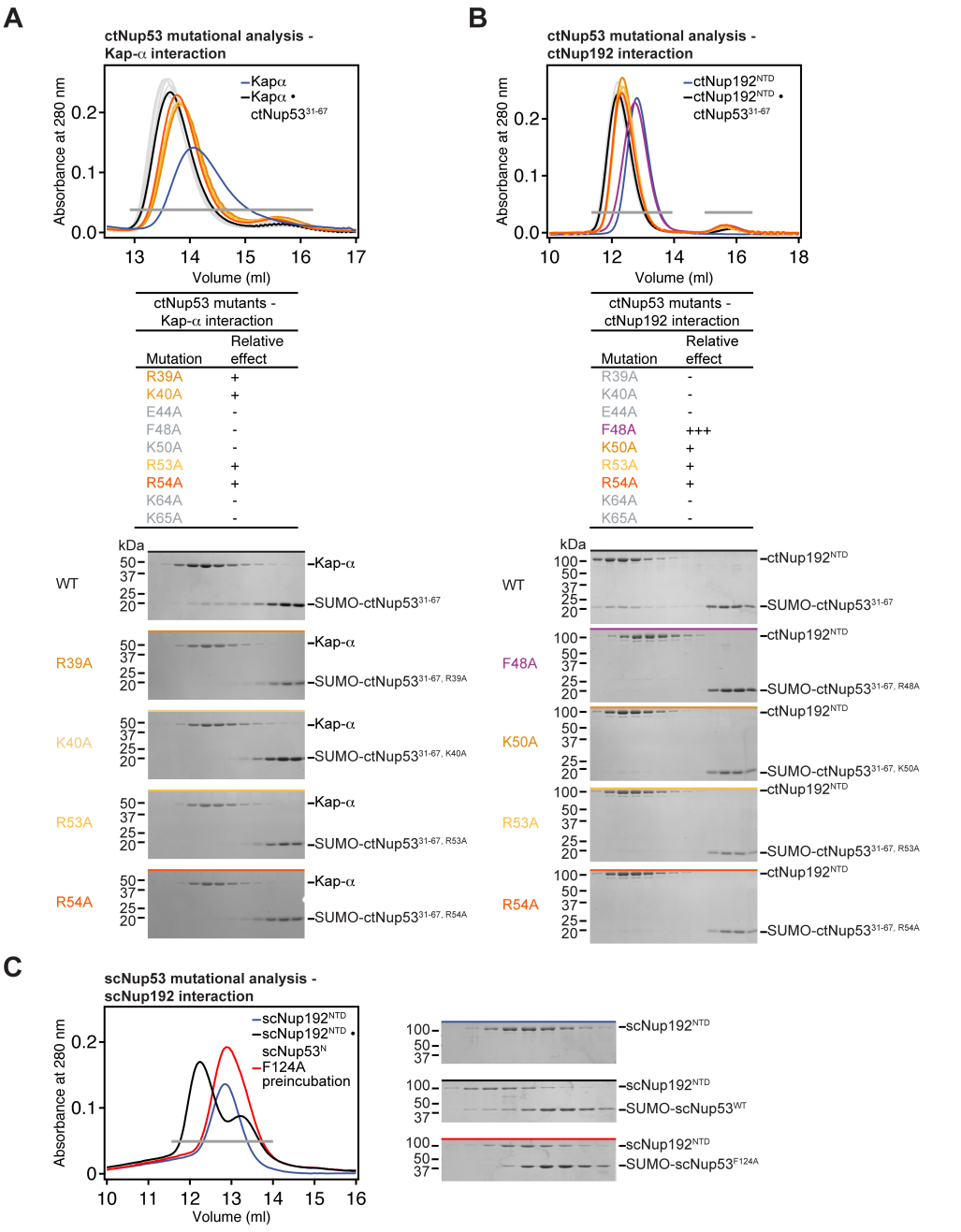
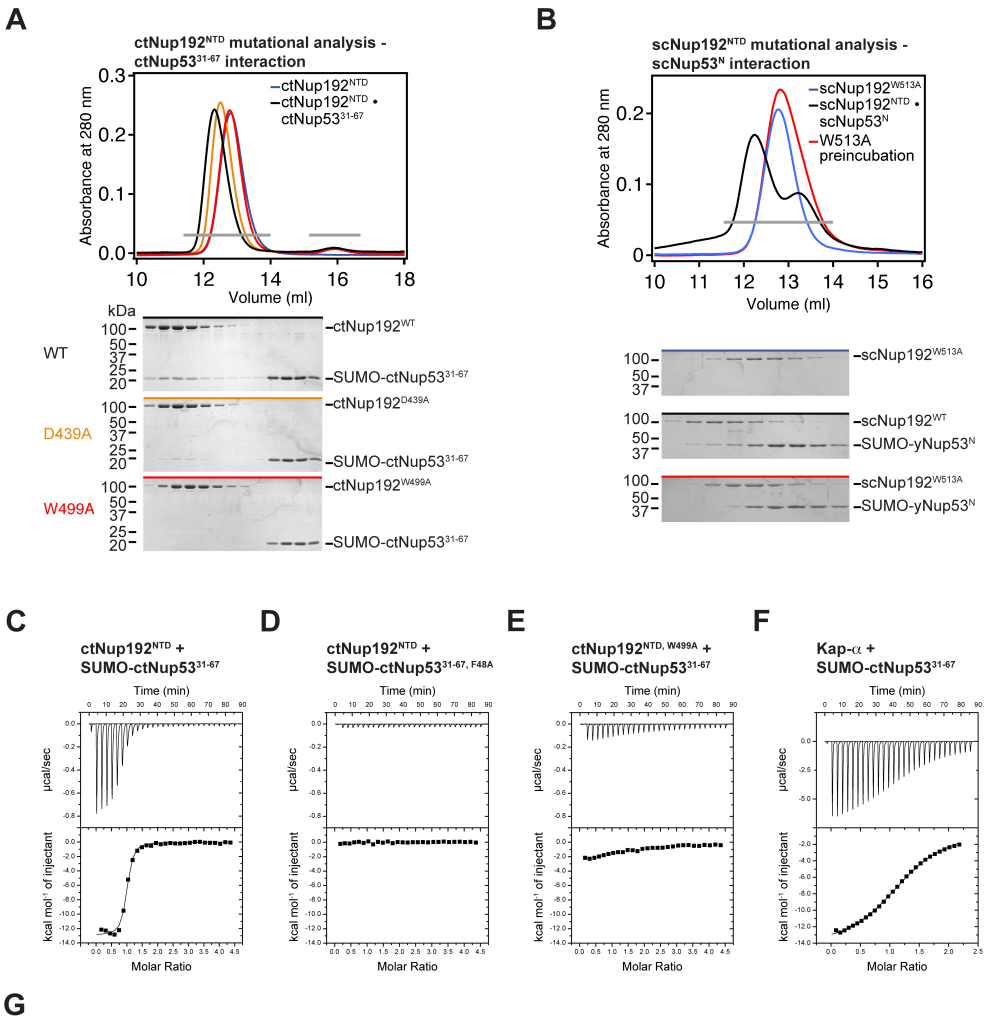


Fig. S6.

Mutational analysis of Nup53 interactions. (*A-B*) SEC interaction profiles of ctNup53$^{31\text{-}67}$ mutants and their effect on (*A*) Kap-α and (*B*) ctNup192^{NTD} binding. The results are summarized in Fig. 5*A*. Relative effects are categorized as no effect (-), reduced binding (+), and complete disruption (+++). Gray bars indicate fractions analyzed by Coomassie brilliant blue-stained gels. The corresponding gel filtration profiles are indicated by the colored bar above each gel. Molecular mass standards and the positions of the proteins are indicated. (*C*) Mutational analysis of *S. cerevisiae* Nup192^{NTD} and Nup53. The corresponding mutation F124A in scNup53^{N} also disrupts binding to scNup192^{NTD}, as shown by SEC analysis.



G

Isothermal titration calorimetry measurements				
Cell	ctNup192^{NTD}	ctNup192^{NTD}	ctNup192$^{NTD, W499A}$	Kap-α
Syringe	SUMO-ctNup53$^{31-67}$	SUMO-ctNup53$^{31-67, F48A}$	SUMO-ctNup53$^{31-67}$	SUMO-ctNup53$^{31-67}$
K_d (µM)	0.10 ± 0.02	ND[a]	ND[a]	25.3 ± 1.4
ΔH (kcal/mol)	-1.30 ± 0.16	-	-	-1.41 ± 0.15
ΔS (cal/mol/deg)	-12.0	-	-	-26.9
N (sites)	0.95 ± 0.01	-	-	1.34 ± 0.01

[a] ND, not determined

Fig. S7.

Mutational analysis of Nup192^{NTD} interactions. (*A*) Representative SEC interaction profiles of ctNup192^{NTD} mutants and their effect on ctNup53$^{31-67}$ binding. Results are summarized in Fig. 5*B*. Representative gel filtration profiles illustrating reduced binding and complete disruption of the interaction are shown in orange and red, respectively. Gray bars indicate fractions analyzed by Coomassie brilliant blue-stained gels. The corresponding gel filtration profiles are indicated by the colored bar above each gel. Molecular mass standards and the positions of the proteins are indicated. (*B*) Mutational analysis of *S. cerevisiae* Nup192^{NTD} and Nup53. The corresponding mutation W513A in scNup192^{NTD} also disrupts binding to scNup53^{N}, as shown by SEC analysis. (*C-F*) Isothermal titration calorimetry analysis of ctNup53$^{31-67}$ interactions. (*G*) Summary of the thermodynamic parameters determined using a single-site model.

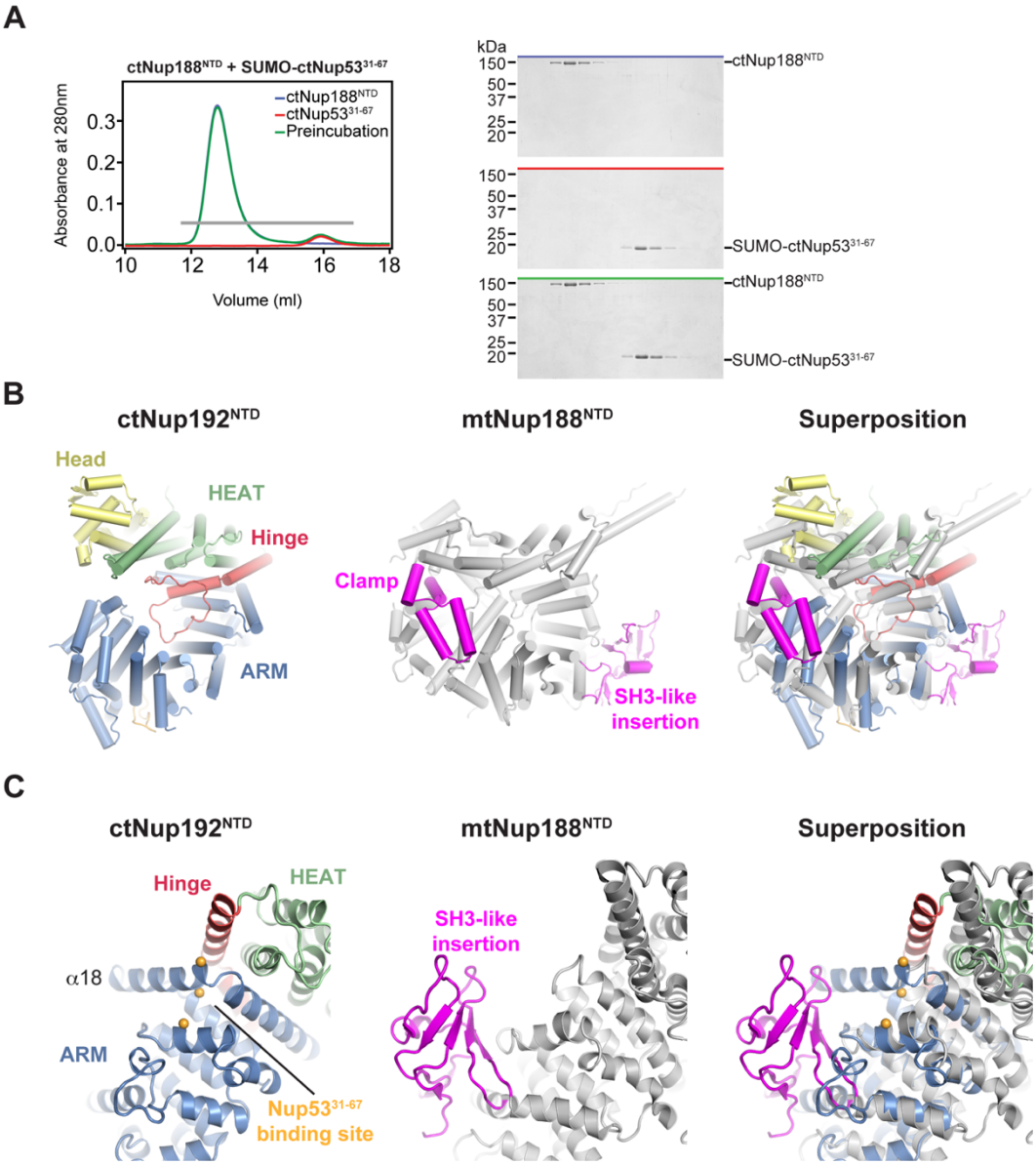


Fig. S8.

Structural basis for distinct large adaptor nucleoporin binding specificity. (*A*) SEC analysis of the interaction between ctNup188^{NTD} and ctNup53$^{31-67}$. Gray bars and colored lines designate the analyzed fractions in the respective Coomassie brilliant blue-stained SDS-PAGE gels on the right. Molecular mass standards and the positions of the proteins are indicated. (*B*) Cartoon representations of ctNup192^{NTD}, mtNup188^{NTD} (PDB 4KF7) (32), and their superposition on the right. mtNup188^{NTD}-specific insertions are colored in magenta. (*C*) Comparison of the ctNup53 binding site in ctNup192^{NTD} with the corresponding location in mtNup188^{NTD}, colored as in (*B*).

Table S1.
Data collection and refinement statistics

	Native	Selenomethionine peak	$[Ta_6Br_{12}]^{2+}$ peak
Data collection			
Protein	*ct*Nup192 NTD	*ct*Nup192 NTD	*ct*Nup192 NTD
Synchrotron	SSRL	SSRL	SSRL
Beamline	BL12-2	BL12-2	BL12-2
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions			
a, *b*, *c* (Å)	102.9, 102.9, 443.1	102.7, 102.7, 443.1	103.0, 103.0, 445.3
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength	1.0000	0.9795	1.2547
Resolution (Å)	50.0 – 2.70	50.0 – 3.40	50.0 – 3.60
R_{sym} (%)[a]	9.5 (100.0)	12.7 (93.4)	11.9 (83.2)
$< I > / <\sigma I >$[a]	21.6 (3.1)	12.4 (2.0)	19.2 (4.0)
Completeness (%)[a]	100.0 (100.0)	99.9 (100.0)	100.0 (100.0)
No. of observations	859,560	235,738	474,247
No. of unique reflections [a]	66,696 (6,506)	33,944 (3,334)	28,901 (2,801)
Redundancy[a]	12.9 (12.9)	6.9 (7.0)	16.4 (16.0)
Refinement			
Resolution (Å)	50.0 – 2.70		
No. of reflections	66,555		
No. of reflections test set	3,390		
Rwork / Rfree	19.2 / 23.1		
No. atoms (non-hydrogen)	14,574		
Protein	14,202		
Water	297		
Ligand/Ions	75		
B-factors			
Protein	70.2		
Water	49.1		
RMSD			
Bond lengths (Å)	0.002		
Bond angles (°)	0.569		
Ramachandran plot[d]			
Favored (%)	96.1		
Additionally allowed (%)	3.9		
Outliers (%)	0.0		

[a]Highest-resolution shell is shown in parentheses
[b]As determined by MolProbity

Table S2.

Expression constructs

Bacterial expression constructs				
Protein	Residues (mutations if applicable)	Expression vector	Restriction sites 5', 3'	N-terminal overhang
ctNup192^{NTD}	1-958	pET28a-PreS[a]	NdeI, NotI	GPH
ctNup192^{CTD}	976-1756	pET28a-PreS	NdeI, NotI	GPHM
ctNup192^{TAIL}	1358-1756	pET28a-PreS	NdeI, NotI	GPHM
ctNup192$^{\Delta HEAD}$	153-958	pET28a-PreS	NdeI, NotI	GPHM
ctNup188^{NTD}	1-1134	pET28a-PreS	AseI, BamHI	GPHN
scNup192^{NTD}	1-960	pET28a-PreS	NdeI, NotI	GPH
ctNup53^{N}	1-90	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67}$	31-67	pET28a-SUMO	BamHI, NotI	S
scNup53^{N}	1-181	pET28a-SUMO	BamHI, NotI	S
ctNic96^{H2}	814-960	pET28a-SUMO	BamHI, NotI	S
ctNic96^{H2}	814-960	pETMCN-SUMO	BamHI, NotI	S
ctNup53$^{31-67, R39A}$	31-67 (R39A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, K40A}$	31-67 (K40A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, E44A}$	31-67 (E44A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, F48A}$	31-67 (F48A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, K50A}$	31-67 (K50A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, R53A}$	31-67 (R53A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, R54A}$	31-67 (R54A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, K64A}$	31-67 (K64A)	pET28a-SUMO	BamHI, NotI	S
ctNup53$^{31-67, R65A}$	31-67 (R65A)	pET28a-SUMO	BamHI, NotI	S
scNup53$^{N, F124A}$	1-181 (F124A)	pET28a-SUMO	BamHI, NotI	S
ctNup192$^{NTD, E295A}$	1-958 (E295A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, E335A}$	1-958 (E335A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, E427A}$	1-958 (E427A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, D431A}$	1-958 (D431A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, S435A}$	1-958 (S435A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, N436A}$	1-958 (N436A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, D480A}$	1-958 (D480A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, E487A}$	1-958 (E487A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, D488A}$	1-958 (D488A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, N492A}$	1-958 (N492A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, D439A}$	1-958 (D439A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, L441A}$	1-958 (L441A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, K443A}$	1-958 (K443A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, R445A}$	1-958 (R445A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, R452A}$	1-958 (R452A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, Y475A}$	1-958 (Y475A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, W486A}$	1-958 (W486A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, L497A}$	1-958 (L497A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, W499A}$	1-958 (W499A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, R502A}$	1-958 (R502A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, R503A}$	1-958 (R503A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, F532A}$	1-958 (F532A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, F562A}$	1-958 (F562A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, E592A}$	1-958 (E592A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, M598A}$	1-958 (M598A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, Y062A}$	1-958 (Y602A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, E747A}$	1-958 (E747A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, R754A}$	1-958 (R754A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, L925A}$	1-958 (L925A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, L928A}$	1-958 (L928A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, V931A}$	1-958 (V931A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, V932A}$	1-958 (V932A)	pET28a-PreS	NdeI, NotI	GPH
ctNup192$^{NTD, L953A}$	1-958 (L953A)	pET28a-PreS	NdeI, NotI	GPH
scNup192$^{NTD, W513A}$	1-960 (W513A)	pET28a-PreS	NdeI, NotI	GPH
scKap-α	88-530	pProEX-HTb[b]	BamHI, XhoI	GAMGS

Table S3.

Yeast constructs

Yeast expression constructs				
Protein	Residues (mutations if applicable)	Shuffle Vector	Restriction Sites 5', 3'	Selection
scNup192^{FL}	1-1683	pRS416-mCherry	NotI, SacII	Ura
scNup192^{FL}	1-1683	pRS415-GFP	NotI, SacII	Leu
scNup192^{NTD}	1-954	pRS415-GFP	NotI, SacII	Leu
scNup192^{CTD}	955-1683	pRS415-GFP	NotI, SacII	Leu
scNup192$^{\Delta TAIL}$	1-1316	pRS415-GFP	NotI, SacII	Leu
scNup192$^{\Delta TAIL,\ W513A}$	1-1316 (W513A)	pRS415-GFP	NotI, SacII	Leu
scNup192$^{FL,\ W513A}$	1-1683 (W513A)	pRS415-GFP	NotI, SacII	Leu
scNup53^{FL}	1-475	pRS416-mCherry	BamHI, NotI	Ura
scNup53^{FL}	1-475	pRS415-GFP	BamHI, NotI	Leu
scNup53$^{FL,\ F124A}$	1-475	pRS415-GFP	BamHI, NotI	Leu

SUPPLEMENTAL REFERENCES

1. Kippert F & Gerloff DL (2009) Highly sensitive detection of individual HEAT and ARM repeats with HHpred and COACH. *PLoS One* 4:e7148.
2. Chook YM & Blobel G (1999) Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. *Nature* 399:230-237.
3. Cingolani G, Petosa C, Weis K, & Muller CW (1999) Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* 399:221-229.
4. Monecke T*, et al.* (2009) Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* 324:1087-1091.
5. Monecke T*, et al.* (2013) Structural basis for cooperativity of CRM1 export complex formation. *Proc Natl Acad Sci U S A* 110:960-965.
6. Sampathkumar P*, et al.* (2013) Structure, dynamics, evolution, and function of a major scaffold component in the nuclear pore complex. *Structure* 21:560-571.
7. Dong X*, et al.* (2009) Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* 458:1136-1141.
8. Hoelz A, Nairn AC, & Kuriyan J (2003) Crystal structure of a tetradecameric assembly of the association domain of Ca2+/calmodulin-dependent kinase II. *Mol Cell* 11:1241-1251.
9. Mossessova E & Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 5:865-876.
10. Romier C*, et al.* (2006) Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr D Biol Crystallogr* 62:1232-1242.

11. Conti E, Uy M, Leighton L, Blobel G, & Kuriyan J (1998) Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* 94:193-204.

12. Doublie S (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol Biol* 363:91-108.

13. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125-132.

14. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326

15. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.

16. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-221.

17. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* 56:965-972.

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

19. Painter J & Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62:439-450.

20. Laskowski RA, Macarthur MW, Moss DS, & Thornton JM (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Crystallogr* 26:283-291.

21. Davis IW*, et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375-W383.

22. Wyatt PJ (1997) Multiangle light scattering: The basic tool for macromolecular characterization. *Instrum Sci Technol* 25:1-18.

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

24. Gwizdek C*, et al.* (2006) Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. *Proc Natl Acad Sci U S A* 103:16376-16381.

25. Seo HS*, et al.* (2009) Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. *Proc Natl Acad Sci U S A* 106:14281-14286.

26. Amlacher S*, et al.* (2011) Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell* 146:277-289.

27. Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66:22-25.

28. Jeanmougin F, Thompson JD, Gouy M, Higgins DG, & Gibson TJ (1998) Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 23:403-405.

29. Barton GJ (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng* 6:37-40.

30. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037-10041.

31. Matsuura Y & Stewart M (2004) Structural basis for the assembly of a nuclear export complex. *Nature* 432:872-877.

32. Andersen KR*, et al.* (2013) Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* 2:e00745.

CHAPTER 3

ARCHITECTURE OF THE SYMMETRIC CORE OF THE NUCLEAR PORE

This chapter was adapted from:

Daniel H. Lin[#], Tobias Stuwe[#], Sandra Schilbach, Emily J. Rundlet, Thibaud Perriches, George Mobbs, Yanbin Fan, Karsten Thierbach, Ferdinand M. Huber, Leslie N. Collins, Andrew M. Davenport, Young E. Jeon, André Hoelz (2016). Architecture of the symmetric core of the nuclear pore, ***Science***, 352(6283):aaf1015

ABSTRACT

The nuclear pore complex (NPC) controls transport of macromolecules between the nucleus and cytoplasm, but its molecular architecture remains poorly defined. We biochemically reconstituted NPC core protomers and elucidated the underlying protein-protein interaction network. Flexible linker sequences, rather than interactions between the structured core scaffold nucleoporins, mediated the assembly of the inner ring complex and its attachment to the NPC coat. X-ray crystallographic analysis of these scaffold nucleoporins revealed the molecular details of their interactions with the flexible linker sequences and enabled construction of full-length atomic structures. By docking these structures into the cryoelectron tomographic reconstruction of the intact human NPC and validating their placement with our nucleoporin interactome, we built a composite structure of the NPC symmetric core containing ~320,000 residues and accounting for ~54 MDa of the NPC's structured mass. Our approach provides a paradigm for the structure determination of similarly complex macromolecular assemblies.

INTRODUCTION

The nuclear pore complex (NPC) is a massive molecular transport channel embedded in the nuclear envelope (*1*). In addition to its role as the sole mediator of bidirectional nucleocytoplasmic transport, the NPC is also involved in diverse cellular processes including transcription, mRNA maturation, and genome organization (*1, 2*). Despite their tremendous size (~120 MDa), NPCs are only composed of 34 different proteins (nucleoporins or nups), which assemble into eightfold-symmetric, ~1000-Å diameter pores that fuse the inner and outer nuclear membranes (Fig. 1, A and B) (*1*). Given their central role in cell biology, nucleoporins have been linked to a wide range of human diseases including viral infection, cancer, and neurodegenerative disease (*1, 3-9*). However, the structure of the NPC and the mechanisms by which it influences cellular processes remain enigmatic.

Most nucleoporins are symmetrically distributed in the NPC, forming a symmetric core that is decorated by asymmetric nucleoporins on the nuclear and cytoplasmic faces (Fig. 1, A and B). Many nucleoporins contain disordered repetitive sequences enriched in phenylalanine and glycine residues called FG repeats, which collectively form a central diffusion barrier that prevents passive diffusion of macromolecules with masses greater than ~40 kDa (Fig. 1, A and B). Larger macromolecules can only traffic through the NPC with the assistance of specialized karyopherin transport factors (*10*).

Despite extensive efforts, the protein-protein interaction network within the NPC remains incompletely characterized, presenting a fundamental limitation to our understanding of NPC architecture. Co-purification and mass spectrometry approaches have revealed some of the strongest interactions, but provide only general spatial restraints due to their limited resolution. The most well-characterized nucleoporin interactions are those within the coat nucleoporin complex (CNC), which comprises approximately half the mass of the symmetric core and contains Nup120, Nup85, Sec13, Nup145C, Nup84, Nup133, and,

in some species, Seh1, Nup37, Nup43, and ELYS (*1*). Structural and biochemical analyses of CNCs from multiple species reveal a highly conserved architecture in which α-helical domains form extensive interaction surfaces to assemble a large, Y-shaped complex (*11-14*). Recent advances have dramatically increased the resolution of cryoelectron tomographic (cryoET) reconstructions of intact NPCs, facilitating the unbiased placement of 32 copies of a ~400 kDa crystal structure of the yeast CNC into the intact human NPC (*13, 14*). The CNCs are arranged in pairs of concentric, eight-membered rings on both the nuclear and cytoplasmic faces of the NPC, accounting for the majority of the observed protein density in the outer rings of the NPC (*13, 14*).

In contrast to our understanding of the organization of the CNC in the intact NPC, relatively little is known about the molecular architecture of the inner ring that lines the central channel. The disordered N-terminal FG repeats of the channel nucleoporins Nup49, Nup57, and Nsp1 project into the central channel while their structured coiled-coil domains form a stable complex, termed the channel nucleoporin hetero-trimer (CNT) (*15, 16*). The CNT, Nic96, Nup192, and Nup145N collectively form a stable subcomplex called the inner ring complex (IRC) (*15, 17*). The remaining components of the symmetric core, Nup170 and Nup53, are thought to mediate interactions with the nuclear envelope, but the details of the interaction network that assembles these proteins and links them to the CNCs remain poorly defined (*18, 19*).

Crystal structures of many nucleoporin fragments from the symmetric NPC core have been determined, including the N-terminal domains (NTDs) of Nup192, Nup188, and Nup157; the C-terminal domain (CTD) of Nup170; the C-terminal tail domains (TAIL) of Nup192 and Nup188; the α-helical solenoid of Nic96; and the CNT bound to Nic96, CNT•Nic96^{R1} (*15, 20-26*). However, structures of full-length Nup192, Nup188, and Nup170

have remained elusive. Accurate placement of the existing structures into the intact NPC is limited by both their relatively small size and the resolution of available cryoET reconstructions. We have used an integrated approach to obtain complete atomic structures and accurate, high-resolution biochemical restraints for determining the near-atomic architecture of the NPC symmetric core.

A significant barrier preventing complete biochemical characterization of the NPC has been the difficulty of purifying significant quantities of full-length nucleoporins from a single species. To overcome this hurdle, we developed expression and purification protocols for all symmetric core nucleoporins from the thermophilic fungus *Chaetomium thermophilum*, which exhibit superior biochemical stability (*27*). By reconstituting NPC symmetric core protomers from purified proteins, we found that the interactions between flexible linker sequences and large scaffold nucleoporins drove NPC assembly. We generated a high-resolution biochemical map for these interactions and determined a series of crystal structures that revealed the structural basis for flexible linker sequence recognition by the large scaffold nucleoporins. These crystal structures enabled the construction of complete atomic structures for the ordered scaffolds of Nup170, Nup192, and Nic96. Using our biochemical restraints for validation, we performed unbiased searches to dock the atomic structures into a cryoET reconstruction of the intact human NPC with high confidence (*28*), thus determining a composite structure of the NPC symmetric core.

RESULTS

Reconstitution of NPC symmetric core protomers

We first directed our efforts towards reconstituting a soluble protomer that could recapitulate the interaction network within the assembled NPC using nucleoporins from the thermophilic fungus *C. thermophilum*. We used full-length proteins when possible, but FG repeats, disordered N- or C-terminal regions, or other sequences that prevented soluble protein expression were omitted (Fig. 1B; fig. S1; table S1 and S2). We first reconstituted a hetero-hexameric core CNC containing Nup120, Nup37, ELYS, Nup85, Sec13, and Nup145C, a complex analogous to the yeast CNC we previously crystallized (fig. S2, A and B) (*14*). This CNC hetero-hexamer was assembled with Nup84 and Nup133 to form a hetero-octameric CNC (fig. S2C). Due to the poor solubility of the intact hetero-octameric CNC, we focused our analysis on its hetero-hexameric core. Similarly, we extended our previous reconstitution of an IRC containing Nup192, Nic96, Nup145N, and the CNT by also incorporating Nup53 (fig. S3A and table S3) (*15*). We found that Nup188, which is evolutionarily related to Nup192, failed to incorporate into the IRC (fig. S3B). Rather, an analogous Nup188 complex formed in the absence of Nup192 (Fig. 1F). By preincubating the core CNC-hexamer with either the IRC or the Nup188 complex, we reconstituted two distinct 13-protein complexes representative of NPC symmetric core protomers (Fig. 1, C, D, and F and fig. S4, A and C).

With the ability to reconstitute NPC protomers, we sought to identify the interactions that linked the IRC or the analogous Nup188 complex to the CNC. Nup145N and Nup145C are components of the IRC/Nup188 complex and CNC, respectively, and originate from the same polypeptide chain after post-translational cleavage mediated by the Nup145N autoproteolytic domain (APD) located in the middle of the Nup145 pre-cursor polypeptide (Fig. 1B) (*29*). Nup145C is composed of a U-bend α-helical solenoid that is a core

component of the CNC and a disordered, ~300-residue N-terminal extension (NTE) (*14*). Previous studies indicate that Nup145N^{APD} can still bind the first six N-terminal residues of Nup145C after cleavage, thus offering a possible mechanism for linking the IRC to the CNC (*30*). Indeed, the complexes did not interact in the absence of Nup145N (Fig. 1, E and G and fig. S4, B and D). Moreover, Nup145N^{APD} alone could be incorporated into the CNC (fig. S5). These results indicated that the IRC and CNC were flexibly attached via the long, intrinsically disordered sequences of Nup145N and Nup145C. Nup145N^{APD} also binds to the β-propeller domain of the cytoplasmic filament nucleoporin Nup82 (*15*). However, since this interaction was outcompeted by CNC binding (fig. S6), another interaction must play a role in retaining the cytoplasmic filaments at the cytoplasmic face of the NPC.

Symmetric core assembly is driven by flexible linkers

We next performed a systematic analysis of the molecular interaction network within the IRC and Nup188 complex, extending our previous work and additionally including Nup170 in our analysis (*15*). These complexes contain large scaffold domains in Nic96, Nup170, Nup188, and Nup192 as well as long, intrinsically disordered sequences in the linker nucleoporins Nup53 and Nup145N (Fig. 1B). While the CNC is primarily held together by extensive interfaces between large α-helical solenoid domains (*1, 14*), recent studies hint that the architectural principles of the remaining symmetric core nucleoporins are different (*15, 17, 27*). For example, the scaffold nucleoporin Nic96 possesses a largely unstructured NTE containing two short helical regions, Nic96^{R1} and Nic96^{R2}, that are essential for IRC assembly: Nic96^{R1} recruits the CNT to the NPC while Nic96^{R2} binds Nup192 or Nup188 (*15*). To determine whether the structured domains of the scaffold nucleoporins interacted with each other to drive symmetric core assembly, we tested whether Nic96^{SOL}, Nup170,

Nup192, and Nup188 could form a complex, but observed no interaction (Fig. 2A and fig. S7A). Instead, complex formation was achieved only in the presence of the linker nucleoporins Nup53 and Nup145N (Fig. 2, B and C and fig. S7, B and C). Thus together with Nic96^{NTE}, the flexible linker nucleoporins Nup53 and Nup145N are the primary driving force of IRC/Nup188 complex assembly (*15*).

To further analyze the interaction network between the scaffold and linker nucleoporins, we tested which scaffolds interacted with Nup53 or Nup145N to form hetero-dimers, focusing first on the interactions within the IRC and with Nup170. Nup53 formed robust complexes with Nup192, Nic96^{SOL}, and Nup170 (Fig. 2D and fig. S8). We previously reported that Nup145N interacts weakly with Nic96^{SOL} (*15*), and here we found that Nup145N also binds to Nup192 and Nup170 (Fig. 2E and fig. S9). All of these scaffold-linker interactions are compatible, as demonstrated by the formation of hetero-trimeric complexes (Fig. 2, D and E and figs. S10 and S11). Nup192 and Nup170 can also bind to both linker nucleoporins simultaneously, indicating that the binding sites on the scaffolds are distinct (fig. S12, A and B). Indeed, we were able to reconstitute a stoichiometric hetero-tetramer composed of Nup192, Nup170, Nup53, and Nup145N (fig. S12C).

To improve the biochemical resolution of our interaction map, we next identified minimal sequence fragments of Nup53 and Nup145N sufficient for scaffold recognition. We previously mapped an interaction between a Nup53 fragment, encompassing residues 31-67 (Nup53^{R1}), with Nup192^{NTD} (Fig. 2F and fig. S13) (*20*). Here, we found an adjacent fragment containing residues 69-90 (Nup53^{R2}) that was recognized by Nic96^{SOL} (Fig. 2F and fig. S14, A and B). A Nup53 fragment including both of these binding sites (residues 1-90) was sufficient to link the two scaffolds into a hetero-trimeric complex (fig. S14C). We also identified a C-terminal Nup53 fragment containing residues 329-361 (Nup53^{R3}) that interacted specifically with Nup170^{NTD} (Fig. 2F and fig. S15). Conversely,

the association between Nup170 and Nup145N mapped to Nup145N residues 729-750 ($Nup145N^{R3}$) and $Nup170^{CTD}$ (Fig. 2F and fig. S16). Nup192 recognized a fragment of Nup145N encompassing residues 606-683 (Fig. 2F and fig. S17, A and B). $Nup192^{NTD}$ was sufficient for Nup145N binding (fig. S17C), but we also detected a weak interaction with $Nup192^{CTD}$ (fig. S17, D and E), suggesting that binding sites for Nup145N were distributed throughout Nup192. These minimal sequence fragments were specific for their binding partners (fig. S18).

While preincubation of the two linker nucleoporins with Nup192, Nup170, and $Nic96^{SOL}$ produced a robust pentameric complex, the analogous preincubation with Nup188 in place of Nup192 produced a mixture of species (Fig. 2, B and C and fig. S7, B and C). To understand this difference in behavior, we repeated the above analysis with Nup188 and identified a robust interaction with Nup145N whereas Nup53 binding was barely detectible (Fig. 2, D and E and figs. S8C and S9B). However, in contrast to our above results, Nup188 did not strongly bind Nup145N in the presence of Nup192 or Nup170 (Fig. 2, D and E and fig. S11). Similar to $Nup192^{NTD}$, $Nup188^{NTD}$ was sufficient for Nup145N binding (fig. S19, A and B). However, the minimal Nup192-binding fragment of Nup145N was not sufficient for Nup188 binding (fig. S19, D and E). Instead, we only detected robust complex formation with a much longer fragment encompassing both the Nup192 and Nup170 binding sites (residues 606-750), explaining the exclusivity of their interactions (Fig. 2F and fig. S19C). We found a similar architecture for the Nup192 and Nup188 binding sites in $Nic96^{R2}$. However, the Nup192 minimal binding fragment (residues 286-301) again was insufficient for Nup188 binding (fig. S20, A to C), which instead required a larger fragment (residues 274-301) (Fig. 2F and fig. S20D). Consistent with these findings, several mutations in the N-terminal region of $Nic96^{R2}$ ablated Nup188 binding but had no effect on Nup192 binding (fig. S20, E to G) (*15*). Thus, Nup192 and Nup188 bound competitively to

directly overlapping sequences in Nic96 and Nup145N, establishing the existence of a distinct Nup188 complex with an architecture analogous to the IRC.

In summary, we found that interactions between the large, ordered scaffold nucleoporins and flexible interaction motifs in Nup53, Nup145N, and Nic96^{NTE} were the dominant driving force for assembly of the NPC symmetric core outside of the CNCs. We built a biochemical map of these interactions by identifying minimal interaction motifs, revealing that the binding sites were spatially distributed throughout the scaffold nucleoporins, but that many of the binding sites on the linker nucleoporins were adjacent or overlapping in sequence (Fig. 2F). In doing so, we identified the exclusive interactions that provide a molecular basis for the formation of two distinct complexes, the Nup192-harboring IRC and an analogous Nup188 complex. As existing crystal structures have not captured interactions between scaffold and linker nucleoporins, we used these results to identify important structural targets for determining the structural basis for this mode of interaction.

Atomic architecture of the Nup170 interaction network

We determined a crystal structure of the Nup170^{NTD}•Nup53^{R3} complex at 2.1 Å resolution (Fig. 3, A and B and tables S4 and S5). In order to obtain high-resolution diffraction, we deleted residues 293-305 from the 3D4A loop of Nup170^{NTD} (fig. S21). Nup170^{NTD} was composed of a seven-bladed β-propeller and a C-terminal α-helical domain (Fig. 3B). An N-terminal α-helix packed against the C-terminal α-helical domain and was followed by three β-strands that formed a triple Velcro-closure against the β-propeller (Fig. 3B). Nup53 adopted an extended conformation and bound atypically to the side of the β-propeller, rather than the top, at blades 1 and 2 (Fig. 3B). The crystallized Nup53 fragment contained residues 329-361, but clear density was only observed for residues 342-355. Blade 2 of the Nup170

β-propeller deviated substantially from a canonical β-propeller blade to generate two hydrophobic pockets that accommodated Nup53 residues L346, L347, L353, and L354 (Fig. 3C). We identified several mutations in Nup170 and Nup53 that could disrupt their interaction (Fig. 3, D and E and fig. S22, A and B). Notably, we observed a complete loss of binding with mutations to Nup170 residues that are evolutionary conserved, F199, I203, and Y235, suggesting that the binding interface is evolutionarily conserved (Fig. 3, D and E; fig. S21; fig. S22, A and B).

Nup53 is anchored to the nuclear envelope by its C-terminal amphipathic helix, either directly or through an interaction with NDC1 (*18, 19*). Nup170 bound to Nup53^{R3}, which is directly adjacent to this C-terminal helix, prompting us to look for features in Nup170 that could also contribute to nuclear envelope binding. We identified two motifs next to the C-terminus of Nup53^{R3} that would be juxtaposed with the nuclear envelope. The first was a WF motif composed of solvent exposed, evolutionarily conserved tryptophan and phenylalanine residues in the 3CD loop (Fig. 3B and fig. S22C). As tryptophan residues are enriched at membrane interfaces, the WF motif may reinforce membrane binding (*31*). The second motif, residing in the 3D4A loop we deleted for crystallization, is predicted to form an amphipathic helix with a striking, evolutionarily conserved absence of charged residues, a feature characteristic of amphipathic lipid packing sensing (ALPS) motifs, which are also present in Nup120 and Nup133 (fig. S22, C and D) (*32, 33*). The Nup170 ALPS motif contained a universally conserved proline residue on the polar face of the helix, a feature reminiscent of antimicrobial membrane destabilizing peptides (fig. S22D) (*34*). We propose that these additional features on Nup170 act synergistically with Nup53 binding to the nuclear envelope to help maintain the extreme membrane curvature in nuclear pores.

We next determined a crystal structure of the Nup170^{CTD}•Nup145N^{R3} complex at 3.5

Å resolution, using a 2.1 Å-resolution structure of *apo* Nup170^{CTD} as a search model (Fig. 3, F and G and tables S4 and S6). Nup170^{CTD} formed an elongated α-helical solenoid containing two stacks of irregular helical pairs, arranged in a zig-zag fashion (Fig. 3G). The two stacks shared a long helix (α31) that capped the first stack and initiated the second stack. Nup145N^{R3} bound to a pair of deep hydrophobic pockets formed on either side of the first two helices of the second helical stack, inserting residues L733 and I735 into the first pocket and L743 and F744 into the second pocket (Fig. 3, G and H). Mutation of any of these residues completely abolished binding to Nup170 (Fig. 3, I and J and fig. S23B). Similarly, mutation of the residues that formed the hydrophobic pockets (F1171, F1154, I1131, and Y1157) strongly affected binding (Fig. 3, I and J and fig. S23A). The hydrophobic nature of both binding pockets is retained throughout eukaryotes (fig. S21), suggesting the evolutionary conservation of this interaction. Only minimal rearrangements of the binding pocket occur upon Nup145N binding (fig. S23C).

The Nup145N sequence that binds to Nup170 is also highly conserved throughout eukaryotes, and the homologous residues critical for Nup170 binding are conserved in humans (fig. S24A). During mitosis, extensive phosphorylation of *hs*Nup98, the human homologue of Nup145N, leads to NPC and nuclear envelope disassembly (*35*). The most abundant mitotic phosphorylation sites in *hs*Nup98 are at residues S608 and S612 (S741 and S745 in *C. thermophilum*), which flank L610 and F611 (L743 and F744 in *C. thermophilum*), residues we found to be critical for Nup170 binding (fig. S24, A and B) (*35*). To test the possibility that the interaction between Nup170 and Nup145N could be regulated by phosphorylation, we reconstituted a *hs*Nup155•*hs*Nup98 hetero-dimer homologous to our crystallized complex. We observed a robust interaction between *hs*Nup155^{CTD} and the corresponding minimal *hs*Nup98 fragment, which was partially disrupted by a

phosphomimetic mutation (S608E/S612E) (fig. S24C). As revealed by the Nup170^{CTD}•Nup145N^{R3} structure, S612 (S745 in *C. thermophilum*) formed a hydrogen bond with N609 (D743 in *C. thermophilum*). Phosphorylation would therefore destabilize the conformation required to insert the critical hydrophobic residues into the binding pocket in Nup170 (fig. S24B). Disruption of this hydrogen bond by mutagenesis also completely abolished binding (Fig. 3J and fig. S23B). Thus, the Nup170-Nup145N interaction is not only highly conserved, but its disruption is also a key step in mitotic NPC disassembly in humans.

Our structures of Nup170^{NTD} and Nup170^{CTD} did not overlap in sequence, preventing accurate modeling of the full-length protein. Therefore, we crystallized a larger fragment of Nup170 containing the α-helical solenoids of both domains (Nup170^{SOL}, residues 575-1402), and determined the crystal structure at 4.0 Å resolution (fig. S25A; tables S4 and S5;). While there was no conformational variability observed for the helices present in Nup170^{NTD}, the Nup170^{CTD} solenoid exhibited an ~20° movement resulting from a minor rearrangement of helix α27 (fig. S25A). We observed a similar conformational variability in the Nup170^{CTD}•Nup145N^{R3} complex structure, where all four molecules in the asymmetric unit adopted different conformations (fig. S25B). With the structure of Nup170^{SOL} as a template, we superposed the structures of Nup170^{NTD} and Nup170^{CTD}, obtaining a total of eight different conformations for full-length Nup170 (figs. S25B and S26).

Molecular basis for recognition of Nup53 by Nic96

We determined crystal structures of *apo* Nic96^{SOL} and a Nic96^{SOL}•Nup53^{R2} complex at 3.3 and 2.65 Å resolution, respectively (Fig. 4, A and B; fig. S27; tables S4 and S7). Nic96^{SOL} formed a rod-shaped molecule consisting of a U-bend α-helical solenoid with the N-terminus situated in the middle of the rod. Although residues 31-84 of Nup53 were included in the

crystallization construct, only residues 67-84 were visible in the electron density. Residues 67-84 of Nup53 formed an amphipathic helix that buried its hydrophobic face into a hydrophobic groove formed by helices α14, α15, and α16 near the U-bend end of the Nic96 solenoid (Fig. 4, B and C). Consistent with our crystal structure, we found that mutations to the hydrophobic residues in this pocket disrupted binding (figs. S28 and S29). Similar to the interaction between Nup170 and Nup145N, we observed minimal conformational rearrangements upon Nup53 binding (Fig. 4D).

Structure of Nup192

Nup192, the largest symmetric core nucleoporin, forms a question-mark shape at low resolution (*26*). Crystal structures exist for $Nup192^{NTD}$ (residues 1-958) and $Nup192^{TAIL}$ (residues 1397-1756), but an atomic structure of the entire molecule has remained elusive (*15, 20, 26*). To determine the complete atomic structure of Nup192, we obtained crystals of an engineered Nup192 truncation mutant, $Nup192^{\Delta HEAD}$, from which we deleted the N-terminal HEAD domain (residues 1-152) and replaced a loop encompassing residues 167-184 with a short glycine-serine linker. We determined the crystal structure of $Nup192^{\Delta HEAD}$ at 3.2 Å resolution (Fig. 4,E and F; fig. S30; tables S4 and S7).

The N-terminal portion of the previously unresolved middle domain of Nup192 ($Nup192^{MID}$) contained three additional ARM repeats (α46-α53, residues 959-1154) that continued the superhelical solenoid we previously observed in $Nup192^{NTD}$ (Fig. 4F) (*20*). Similarly, the C-terminal portion of $Nup192^{MID}$ contained a HEAT repeat (α60-α61, residues 1330-1376) that extended the $Nup192^{TAIL}$ solenoid such that the entire protein formed a continuous HEAT/ARM repeat solenoid (Fig. 4F) (*15*). However, we observed an unusual insertion (residues 1155-1329) between the ARM repeats and HEAT repeat in $Nup192^{MID}$

containing a ~50-residue helix, α58, that reached ~75 Å from the beginning of Nup192^{TAIL} to the C-terminus of Nup192^{NTD} (Fig. 4F). This insertion, which we termed the Tower helix, buried several hydrophobic residues against the bottom of Nup192^{NTD}, inducing minor rearrangements that facilitate packing of the Tower helix. While the Tower helix has only a moderate signature in secondary structure predictions, the evolutionarily-related Nup188 is also predicted to contain a similarly long, ~40 residue helix at the same location. However, previous models of either full-length Nup188 or Nup192 never anticipated the existence of the Tower helix, highlighting the importance of experimentally determining atomic resolution structures (*22, 36*).

Taking advantage of the extensive overlap between our crystal structure of Nup192$^{\Delta HEAD}$ with the existing structures of Nup192^{NTD} and Nup192^{TAIL}, we generated the structure of full-length Nup192 by superposition (Fig. 4G and fig S31). Inspection of the full-length protein revealed that the first loop in the HEAD domain between α1 and α2 was close enough to contact loops in the MID domain, predominantly with polar and charged residues (Fig. 4G and fig. S31). The binding sites on Nup192 for Nup53 and Nic96, which we previously identified via mutagenesis, were located at the top and bottom of the molecule, respectively, ~140 Å apart from each other (fig. S32) (*15, 20*).

Architecture of the NPC symmetric core

Recent advances in cryoET have produced rapidly improving reconstructions of intact NPCs, with the most recent reconstructions reporting average resolutions up to ~20 Å in the best resolved regions (*13, 28, 37*). We previously docked 32 copies of the yeast CNC into a ~34 Å reconstruction of the intact human NPC, taking advantage of the distinctive shape and large size of the complex (*14*). With the addition of the Nic96^{SOL}•Nup53^{R2} structure and the

full-length, superposition-generated structures of Nup192 and Nup170•Nup53^{R3}•Nup145N^{R3} reported here, as well as our recently reported crystal structure of the CNT•Nic96^{R1} complex, accurate structures were available for essentially the entire ordered mass of the NPC symmetric core (*15*). The domain architectures of the symmetric core nucleoporins are highly conserved from fungi to humans (fig. S33). We successfully located these structures in the recently reported ~23 Å reconstruction of the human NPC, with the arrangement of nucleoporins validated by our biochemical restraints (*28*). We utilized an incremental approach to confidently place the crystal structures, starting with the largest structures possessing the most distinctive shapes, and iteratively removing the occupied density to search for subsequent structures (fig. S34). As the cryoET map possesses eightfold rotational symmetry, each unique solution defined the location and orientation of eight copies of each molecule. We first tested our approach with the yeast CNC crystal structure and found four unique placements with exceptional scores compared to 50,000 other refined placements, in excellent agreement with our previous results (fig. S35A) (*14*). We also readily identified the location and orientation of human Nup84^{CTD}•Nup133^{CTD} in unbiased searches (fig. S35B) (*38*). Based on previously reported biochemical data, we manually docked the Nup37, Nup43, and Nup133 β-propellers and locally optimized their fit (fig. S35C) (*13, 39-41*).

We next performed unbiased searches for Nup170 and Nup192 using a map from which density corresponding to the CNCs had been removed (fig. S36). As our crystallographic data indicated significant flexibility in the Nup170 solenoid, we performed searches with the eight different conformations of Nup170. These searches identified two conformations that each yielded two distinct top scoring solutions (fig. S36, A and B). Searches with full-length Nup192 revealed six unique solutions, but because Nup192 and Nup188 could not be distinguished at this resolution, we assigned two of these as Nup188

using our biochemical results and previously reported cross-linking data, as detailed in the methods (fig. S36C) (*13, 42*). After removing the density assigned to Nup170, Nup192, and Nup188, we successfully located four unique copies of the Nic96^{SOL}•Nup53^{R2} and CNT•Nic96^{R1} complexes (fig. S37, A and B). Lastly, we inspected the remaining density in the inner ring in an attempt to locate the ordered domains of Nup53 and Nup145N. We determined crystal structures of Nup53^{RRM} and Nup145N^{APD}•Nup145C^{N} at 0.8 Å and 1.3 Å resolution, respectively, but could not unambiguously place them due to their small size and globular shape (fig. S38 and tables S4, S8, and S9). We attempted to generate biochemical restraints to dock Nup53^{RRM} confidently, but were unable to find any binding partners (fig. S39). However, we did find a pair of continuous densities that readily accommodated two additional Nup170 molecules in a third distinct conformation (fig. S40A). These placements were buried in our original global search, but the conformation of this Nup170 structure still differed slightly from the remaining map density, suggesting that our crystal structures did not capture the full conformational range of Nup170 (fig. S40, B and C).

With the CNC-hexamer, Nup84^{CTD}•Nup133^{CTD}, Nup133^{NTD}, Nup37^{NTD}, Nup43, Nup188^{NTD}, Nup188^{TAIL}, Nup192, Nup170•Nup53^{R3}•Nup145N^{R3}, Nic96^{SOL}•Nup53^{R2}, and CNT•Nic96^{R1} structures placed, we acquired a composite structure accounting for nearly all of the density in the NPC symmetric core, corresponding to ~54 MDa of protein mass or ~320,000 ordered residues (Fig. 5; fig. S41; table S10). In total, the composite structure contained a stoichiometry of 16 copies of Nup188, 32 copies of the CNC, Nup192, Nic96, and CNT, and 48 copies of Nup170, Nup53, and Nup145N. Overall this stoichiometry is in good agreement with a previous study that used mass spectrometry to measure the relative abundances of nucleoporins in the human NPC (*43*).

Spoke architecture

The symmetric core of the NPC consisted of eight spokes related by an eight-fold rotational axis of symmetry perpendicular to the nuclear envelope (Fig. 5A). Outer rings resided above the nuclear and cytoplasmic faces of the nuclear envelope, while an inner ring was embedded in the pore and spanned the nuclear envelope (Fig. 5B). When viewed from the cytoplasm, the nucleoporins formed distinct cylinders, with the CNTs lining the transport channel, surrounded by successive cylinders formed by Nup192, Nic96, Nup170, and the CNCs (Fig. 5A). Nic96^{NTE} and the linker nucleoporins Nup53 and Nup145N spanned these cylinders. Only C_8 rotational symmetry was applied to generate the cryoET reconstruction, yet we observed an additional two-fold axis of symmetry in our composite structure relating the nuclear and cytoplasmic sides within each spoke (Fig. 6, A to D) (*28*).

Each inner ring spoke contained four copies of the IRC (Fig. 6B). Our results provided spatial restraints for Nic96^{R1} and Nic96^{R2}, allowing us to trace the path of Nic96^{NTE}, which emerges from the middle of Nic96^{SOL}, to its binding sites on Nup192 and the CNT. We refer to these four distinct IRCs as nuclear peripheral, nuclear equatorial, cytoplasmic peripheral, and cytoplasmic equatorial IRCs in the following text (Fig. 6B). The nuclear and cytoplasmic equatorial IRCs were related to each other directly by the two-fold rotational axis of symmetry, as were the nuclear and cytoplasmic peripheral IRCs. Unexpectedly, the subunits in the equatorial and peripheral IRCs were in approximately the same relative orientation, which was readily apparent upon superposition (Fig. 6E). Because the subunits were placed independently, this surprising symmetry was an emergent property of the composite structure. The docking reveals that the CNT and Nup192 are in close proximity, suggesting that additional weaker interactions orient the CNTs in the fully assembled NPC. We observed additional knobs of density adjacent to the Nup57 α/β domains of each CNT,

which were unexplained by our composite structure (fig. S37D). When we superposed structures of CNT fragments from *Xenopus laevis* onto our docked CNT molecules (*16*), we found that the metazoan-specific ferredoxin-like domain of Nup57 accounted for these extra knobs of density (fig. S37, C and D), further validating our composite structure.

The mechanism by which FG repeats in the central transport channel form a diffusion barrier and facilitate transport remains controversial, partly because the stoichiometry and orientation of the FG repeats remain unknown. Our composite structure revealed a total of 32 CNTs in the inner ring, which would project 96 distinct polypeptide chains in clusters of three into the central channel (Fig. 6F). The orientation of the CNTs suggested that the FG repeats would emanate circumferentially towards the adjacent spoke rather than pointing radially towards the center of the channel. Unexpectedly, the N-termini of the peripheral and equatorial CNTs were evenly spaced and roughly planar such that they approximately formed two 16-membered rings (Fig. 6F).

In the outer rings, each spoke contained two CNCs on either face, which we refer to as proximal and distal CNCs based on their distance from the inner ring. The orientations of Nup133 relative to the CNC core differed slightly between the distal and proximal CNCs, but yielded the same overall architecture, as each pair of distal and proximal CNCs formed an arch over the nuclear envelope (fig. S42). The outer rings also contained a Nup188 molecule on either face (fig. S43). The majority of the CNC components were ~100 Å above the membrane, and the only contacts with the membrane were made by the β-propeller domains of Nup120 and Nup133 through their ALPS motifs (Fig. 6C). Similarly, the IRCs did not make direct contacts with the membrane, but instead were surrounded by a network of Nup170 molecules that formed the outermost layer of the inner ring (Fig. 6C). Each spoke contained three distinct pairs of Nup170 molecules, which we refer to as equatorial,

peripheral, and bridging Nup170 molecules. The equatorial pair occupied alternating orientations equatorially along the surface of the nuclear envelope (Fig. 6C and fig. S36A). The resolution of the cryoET reconstruction was high enough for these molecules that the central holes of the β-propellers were readily visible at higher contour levels (fig. S36A). The bridging pair of Nup170 molecules bridged the inner ring to the outer CNC rings, via a contact with Nup120 (Fig. 6C and figs. S36B and S43). The peripheral pair of Nup170 molecules, which had a weaker quality of fit and was identified only after placing all other symmetric core components, contacted both the equatorial and bridging Nup170 molecules (Fig. 6C and fig. S40). The equatorial Nic96 molecules contacted multiple Nup170 molecules, effectively bridging the nuclear and cytoplasmic networks. Notably, the ALPS and WF motifs we identified in $Nup170^{NTD}$ and the C-terminal amphipathic helix of Nup53 were oriented directly adjacent to the nuclear envelope (Fig. 6C). Thus, Nup170 and Nic96 constitute a membrane coat for the inner ring analogous to the CNCs in the outer ring.

Inter-spoke interactions

We next searched for interfaces mediating interactions between spokes. Several potential interactions have been identified in previous studies, including one between $Nup133^{NTE}$ and $Nup120^{NTD}$ (*40*). Our composite structure revealed four additional interactions between CNC components that could link adjacent spokes: (1) between the neighboring proximal Nup84 and distal Nup85, (2) between the proximal Nup133 α-helical solenoid and distal Nup120 α-helical solenoid, (3) between the proximal Nup133 β-propeller and proximal Nup120 β-propeller, and (4) between the distal Nup133 β-propeller and distal Nup120 β-propeller (fig. S43, A to C). The space between the proximal and distal CNCs contained density that readily accommodated a Nup188 molecule. Nup188 recognized a special niche generated in

the CNC inter-spoke interface, bridging four CNC molecules by potentially making contacts with: (5) the distal Sec13, (6) the distal Nup85, (7) the proximal Nup43, and (8) the proximal Nup133 of a neighboring spoke (fig. S43, B and C). The Nic96^{R2} binding site in Nup188^{TAIL} was not occluded by any additional density (fig. S43B). Due to the absence of strong density in this region, we cannot determine whether only Nup188 or entire an Nup188 complex would be anchored to the outer rings at this site. However, any flexibly tethered components of the NPC, including the remainder of the Nup188 complex, may not be clearly visible in cryoET reconstructions.

Inspection of the interaction surfaces also provided a molecular explanation for how two CNC rings are assembled. Nup133 and Nup170 bind to distinct Nup120 molecules via overlapping interfaces on the proximal and distal Nup120 molecules, respectively, effectively capping the CNCs on either side (fig. 43, D and E). These results are in agreement with a common evolutionary origin for Nup120, Nup133, and Nup170, which all possess similar domain architectures, contact the nuclear envelope via ALPS motifs, and interact with each other at inter-spoke interfaces. The U-bend solenoid nucleoporins, Nup85, Nup145C, Nup84, and Nic96, bridge these inter-spoke interfaces to form a continuous membrane-bending coat. These results also support the protocoatomer hypothesis of a common evolutionary origin for vesicle coats and the NPC coat, wherein the extant nucleoporins derived from an ancient membrane coat containing these protein folds (*44*).

We could only identify a single interaction that would analogously link the inner ring spokes, which would be mediated by an interaction between Nup53^{R1} and Nup192 (Fig. 7). In our composite structure, peripheral Nic96 molecules oriented Nup53^{R2} directly adjacent to the Nup53-binding site on equatorial Nup192 molecules from a neighboring spoke (Fig. 7B). Our biochemical mapping experiments identified adjacent binding sites for Nic96 and Nup192 on Nup53 (Fig. 2F) and that a fragment containing both these binding sites could

bridge the two nucleoporins (fig. S14C). Thus, binding of Nup53^{R1} in trans to a Nup192 molecule from a neighboring spoke would link the inner ring spokes.

An open question regarding nucleocytoplasmic transport has been the mechanism of inner nuclear membrane (INM) protein transport through the NPC, particularly for INM proteins with large globular nuclear domains. Peripheral channels on the order of ~100 Å have been proposed as routes for INM transport (*45*), but we observed no such channels through the inner ring in either our composite structure or the cryoET reconstruction (Fig. 5). Given the dense packing within each spoke, traffic of INM proteins through a spoke would require significant disruption of NPC structure (Fig. 7A). Rather, the most likely path through the inner ring would be at the inter-spoke interfaces where Nup53^{R1} and Nup192 interact (Fig. 7). The CNCs form an ~100-Å arch above this interface, providing an uninterrupted path to the inner ring and possibly explaining the previously observed upper limit for the size of nuclear domains (Fig. 6A) (*46*). However, the channel at the inner ring was much smaller, suggesting that rearrangements at the inter-spoke interface may be necessary to traffic large nuclear domains. Thus, our composite structure of the NPC symmetric core enables rational design of experiments to further understand the mechanism of INM protein import.

A flexible linker mediates CNC oligomerization

While reconstituting the CNCs, we noticed that assembly of the CNC-octamer spontaneously generated a separate solution phase (fig. S44A). Similar phase transitions were previously seen in other systems with multiple binding valencies, suggesting that the oil droplet formation we observed resulted from oligomerization of the CNC (*47*). We previously identified an interaction between Nup133^{NTE} and Nup120 that would mediate head-to-tail CNC ring formation consistent with the composite structure of the NPC (*40*). Removal of the unstructured Nup133^{NTE} completely ablated both oil droplet and complex formation,

suggesting that this flexible interaction serves as a driving force for CNC ring formation (fig. S44). We also found that Nup170 could be incorporated into this separate solution phase and that this incorporation was ablated by C-terminal truncation, which was consistent with the interaction between the proximal Nup120 and bridging Nup170 molecules we observed in our composite structure (fig. S44A).

Conservation of NPC architecture

Our results establish the principles that drive the assembly of nucleoporins in the NPC. While the outer rings assemble largely via structurally rigid interaction surfaces, inner ring assembly is primarily driven by flexible linker sequences within Nup53, Nup145N, and Nic96^{NTE}. This dichotomy may reflect the different roles of the respective complexes. The outer rings provide a structural scaffold for the NPC and, given their location above the plane of the nuclear envelope, their assembly would not be affected dramatically by the dynamic generation of membrane curvature during fusion of the inner and outer nuclear membranes. In contrast, the proteins in the inner ring occupy an environment that only exists after membrane fusion, likely necessitating conformational flexibility over the course of NPC assembly.

The importance of these flexible interactions in the NPC is highlighted by their evolutionary conservation despite poor overall sequence conservation in the linker nucleoporins Nup53 and Nup145N. The overall folds of the scaffold proteins are well-conserved in *S. cerevisiae* (fig. S45) and furthermore, point mutations in the binding pockets of *S. cerevisiae* Nic96, Nup170, and Nup157 also disrupted their interaction with linker sequences (figs. S46 to S48). A complete understanding of the interaction network in *S. cerevisiae* has been partially intractable because of the genetic redundancy that arises from several gene duplications (Nup170/Nup157, Nup53/Nup59, and

Nup145N/Nup100/Nup116). We found that the paralogs mostly retained the ability to form these interactions, but did not detect an interaction between *sc*Nup188 and any of the Nup145N paralogs (figs. S46 to S50).

Our structural data also highlighted the evolutionary conservation of nucleoporin structure and their interactions. While crystal structures of the human scaffold nucleoporins have not been determined, previous comparisons of the fungal CNC with low-resolution reconstructions of the human CNC suggest a conserved architecture (*13, 14*). Superposition of the structures of $Nup53^{RRM}$ and $Nup145N^{APD}$ with their human homologues also revealed that their folds were identical (fig. S38) (*30*). In addition, the mechanism of interaction between Nup170 and Nup145N is conserved in humans, and we found that phosphomimetic mutations weakened the interaction between *hs*Nup98 and *hs*Nup155. Several other phosphorylation sites have been identified in Nup98, many of which potentially overlap with scaffold binding sites (*35*). Therefore, phosphorylation could also regulate other key interactions, including those that occur between Nup145N and Nup188, Nup192, and the CNC. We note that Nup53 is similarly phosphorylated in a cell-cycle dependent manner, and thus phosphorylation of both linker nucleoporins would be an effective means to disassemble the entire inner ring of the NPC (*48*).

CONCLUSIONS

Determining the molecular details of nucleocytoplasmic transport has been a longstanding challenge, at least in part due to an incomplete understanding of the architecture and biochemistry of the NPC itself. We have used purified, recombinant proteins to systematically characterize the nucleoporin interaction network and determine atomic resolution structures of nucleoporin complexes. This approach was crucially complemented by recent advances in cryoET reconstructions (*28*). Using the results of our divide-and-conquer approach, we were able to dock the available crystal structures into a cryoET reconstruction of the human NPC, yielding a composite structure for the entire NPC symmetric core. This union of bottom-up and top-down approaches offers a paradigm for determining the architectures of similarly complex macromolecular assemblies.

Our composite structure differs dramatically from the previously reported computational models not only in relative and absolute stoichiometry, but also in overall architecture (*49*). These discrepancies highlight the complexities that must be accommodated when attempting a holistic, computational approach. We observed a remarkable degree of symmetry in the structure of the NPC, which explains how such a limited vocabulary of proteins can generate such a large macromolecular structure. Most nucleoporins also occupy multiple, distinct biochemical environments. Nup170 offers a dramatic example of this property, as biochemically distinct versions of Nup170 are either buried in the inner ring or are exposed in the bridge between the inner and outer rings. Similarly, Nup120 utilizes overlapping, exclusive interfaces to contact Nup170 and Nup133. Due to this diversification of nucleoporin function, the NPC can be encoded by a relatively small number of genes. The gene duplications of nucleoporins in *S. cerevisiae* may reflect the gradual separation of these distinct functions into several genes. Nup170 appears to also adopt different conformations in each of its distinct biochemical environments, which is consistent with the wide

conformational range we observed in crystal structures. It is possible that the different conformations are the result of different mechanical forces acting on Nup170 at each position.

Biochemical diversification of proteins within the same protein complex also generates enormous challenges for computationally modeling the structure of the NPC and similar complexes, as distance restraints such as crosslinks that are valid in one biochemical environment may be violated in another. This challenge is exacerbated by the possibility of flexibly tethered domains or nucleoporins, such as those in the Nup188 complex. Our results also highlight the confounding effect generated by the flexible linker nucleoporins Nup53 and Nup145N, which occupy binding sites that span the entirety of the inner ring, rather than a single globular volume. The structure of the NPC could be used as a template for the development of methods that can accommodate these additional complexities.

Our composite structure of the NPC provides a rich platform for contextualizing previous results, not all of which can be commented upon here. The structure also permits the rational design of new experiments to not only further validate the structure, but also begin structure-function interrogation of the NPC. Our biochemical results built a map of the strongest and most conserved interactions, but our composite structure clearly indicates that many additional interactions can occur in the context of the assembled NPC. However, a structural understanding of the entire NPC at single residue resolution still requires several advances. Successful placement of crystal structures of fungal nucleoporins into a cryoET reconstruction of the intact human NPC highlights not only the evolutionary conservation of NPC structure, but also the need for structural characterization of human nucleoporins and further improvement of the resolution of cryoET reconstructions to improve the accuracy of the composite structure of the NPC. Improved resolution in cryoET reconstructions where secondary structure elements can be visualized would enable flexible fitting of high-

resolution crystal structures. We were not able to dock any of the asymmetric components of NPC into the composite structure, due to the small size of extant structures and the absence of high-quality restraints (fig. S41). A similar approach to the one used here will be necessary to extend our analysis to the many proteins in the cytoplasmic filaments and nuclear basket that interact directly with transport factors and are implicated in human disease. Lastly, high-resolution structural characterization of the remaining interactions, especially those involving flexible linker sequences, remains critical to building a truly complete structure of the NPC, as these interactions may never be resolved in cryoET reconstructions.

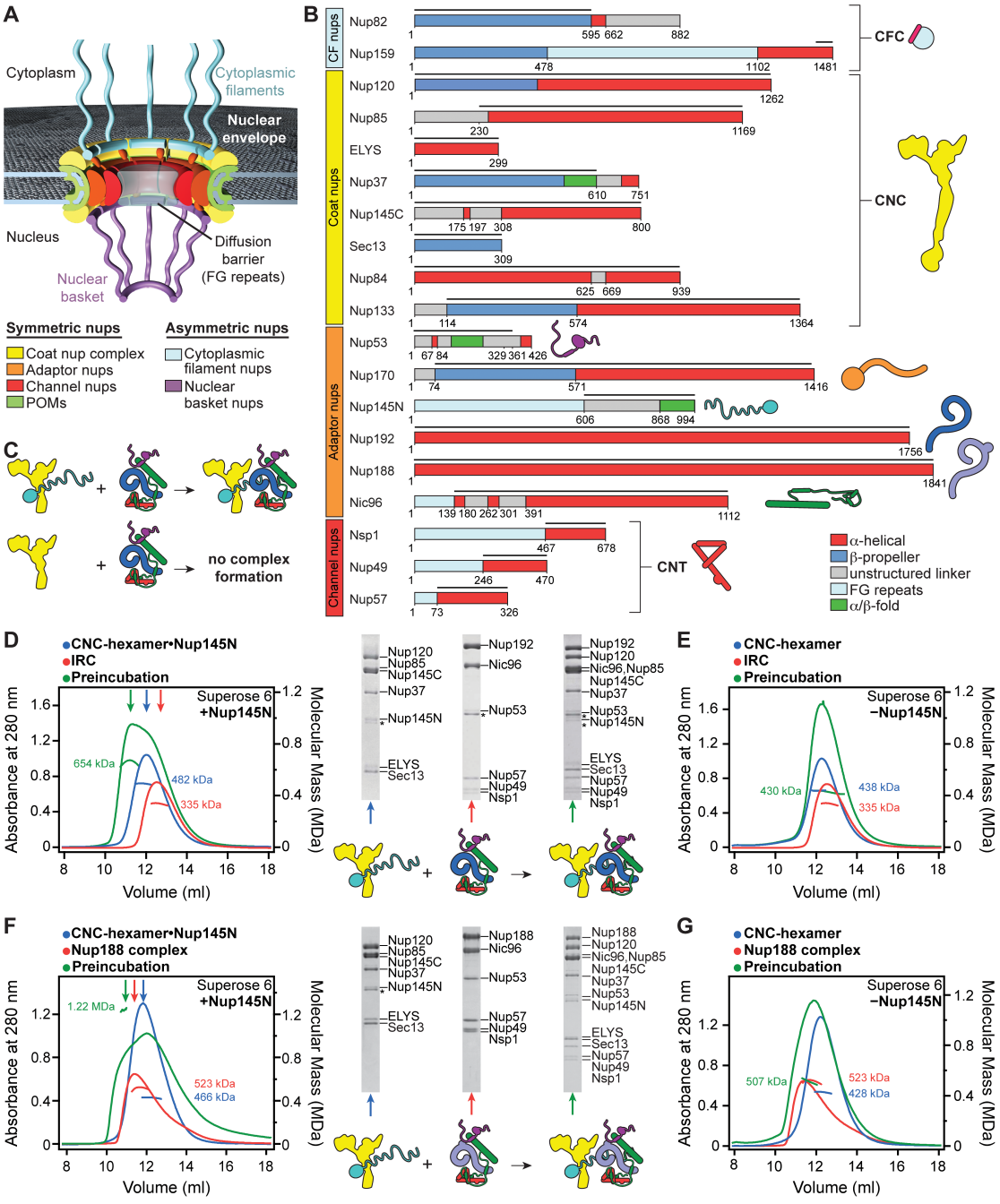


Fig. 1.

Reconstitution of NPC symmetric core protomers. (**A**) Cross-sectional conceptual schematic of the NPC. CNCs, colored yellow, form outer rings above the membrane on the nuclear and cytoplasmic faces. Adaptor and channel nucleoporins, colored orange and red, respectively, form concentric cylinders in the inner ring. Asymmetric nucleoporins decorate the cytoplasmic and nuclear faces of the NPC. (**B**) Domain organization of nucleoporins used in biochemical reconstitution experiments. Black lines indicate the construct boundaries used. Domains are drawn as boxes colored according to the legend at the bottom right for observed or predicted folds. Cartoons to the right are used throughout the text to represent

the nucleoporins. Some nucleoporins form stable complexes (CNC, coat nucleoporin complex; CFC cytoplasmic filament nucleoporin complex; CNT, channel nucleoporin hetero-trimer) and are drawn as a unit. (**C**) Cartoon of experimental setup in panels (D-G). The CNC was preincubated with IRC/Nup188 complex in the presence or absence of Nup145N. Complex formation was only observed in the presence of Nup145N. (**D**, **E**) Reconstitution of NPC symmetric core protomers containing the CNC-hexamer and IRC is dependent on Nup145N. Identical experiments were performed in the (D) presence or (E) absence of Nup145N. Size-exclusion chromatography coupled to multiangle light scattering (SEC-MALS) profiles of the complexes in isolation (red and blue) and after preincubation (green) are shown. Representative Coomassie-stained SDS-PAGE gel slices for the peak fractions are shown with a colored arrow above the chromatogram indicating the resolved peak fraction. Measured molecular masses are indicated for each peak. Cartoons below the gel slices illustrate the respective complexes. (**F**, **G**) Reconstitution of NPC symmetric core protomers containing the CNC-hexamer and Nup188-complex is dependent on Nup145N. Identical experiments were performed in the (F) presence or (G) absence of Nup145N. Complete SDS-PAGE gels for all panels are shown in fig. S4.

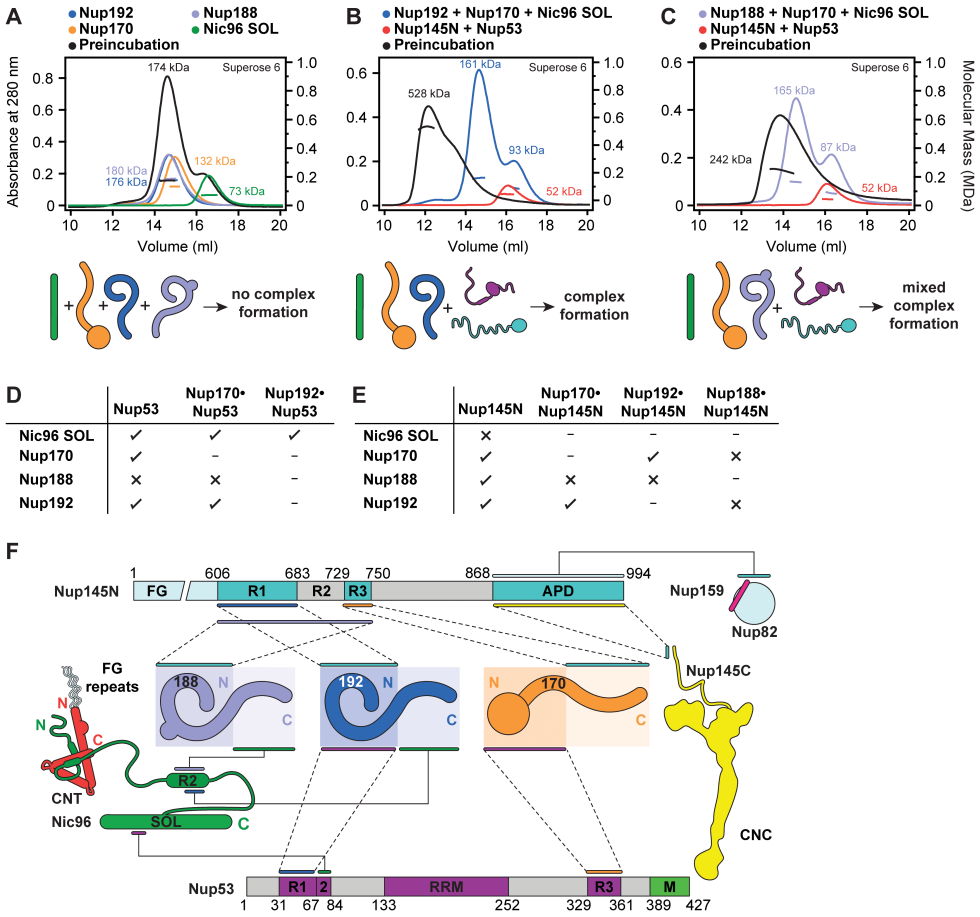


Fig. 2.

Biochemical analysis of the interactions mediating NPC symmetric core assembly. (**A**) Scaffold domains of Nup192, Nup188, Nup170, and Nic96^{SOL} do not interact with each other. SEC-MALS profiles of the individual scaffolds alone (blue, purple, orange, and green) and after their preincubation with each other (black) are shown. Measured molecular masses are indicated for each peak. (**B**, **C**) Linker nucleoporins Nup53 and Nup145N mediate scaffold nucleoporin assembly. SEC-MALS profiles of a scaffold nucleoporin mixture (blue or purple), a linker nucleoporin mixture (red), and after their preincubation with each other (black) are shown. Mixtures of scaffold nucleoporins contained either (B) Nup192 or (C) Nup188. Complete SDS-PAGE gels for all chromatograms are shown in fig. S7. (**D**, **E**) Interaction network between scaffold nucleoporins and (D) Nup53 or (E) Nup145N. Scaffold nucleoporins were tested in SEC-MALS interaction experiments for their ability to form hetero-dimeric complexes with Nup53 or Nup145N and their compatibility to also form hetero-trimeric complexes. Check marks indicate complexes that can form in SEC-MALS experiments, crosses indicate complexes that do not form, and dashes indicate complexes that were not tested. Complete SEC-MALS chromatograms and SDS-PAGE gels are shown in figs. S8 to S11. (**F**) Biochemical interaction map revealed by SEC-MALS interaction experiments. Minimal regions of Nup145N and Nup53 sufficient for binding to components of the NPC are depicted using colored bars and dashed lines between interacting regions. Interactions that map to the same regions on Nup145N and Nup53 do not occur simultaneously. Complete SEC-MALS chromatograms and SDS-PAGE gels are shown in figs S13 to S20. The nucleoporin schematics are according to Fig. 1B.

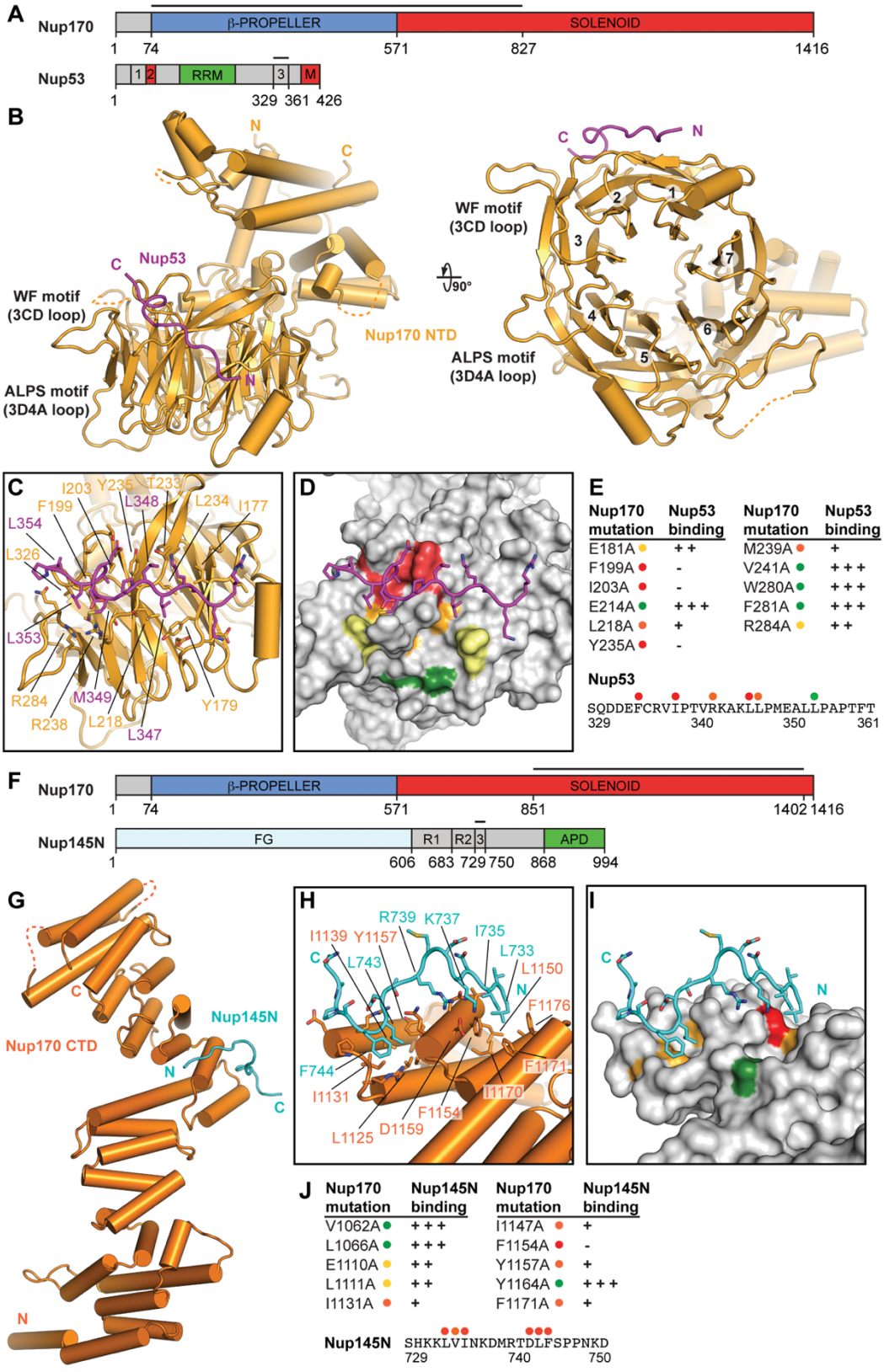

A
Nup170
β-PROPELLER
SOLENOID
1
74
571
827
1416
Nup53
1
2
RRM
3
M
1
329
361
426
B
N
C
C
Nup53
WF motif
(3CD loop)
ALPS motif
(3D4A loop)
N
Nup170 NTD
C
N
WF motif
(3CD loop)
90°
ALPS motif
(3D4A loop)
1
2
3
4
5
6
7
C
I203
Y235
T233
F199
L348
L234
I177
L354
L326
L353
R284
M349
Y179
R238
L218
L347
D
E
Nup170 mutation
Nup53 binding
E181A + +
F199A -
I203A -
E214A + + +
L218A +
Y235A -
M239A +
V241A + + +
W280A + + +
F281A + + +
R284A + +
Nup53
SQDDEFCRVIPTVRKAKLLPMEALLPAPTFT
329
340
350
361
F
Nup170
β-PROPELLER
SOLENOID
1
74
571
851
1402
1416
Nup145N
FG
R1
R2
3
APD
1
606
683
729
750
868
994
G
C
Nup145N
Nup170 CTD
N
C
N
H
R739
K737
I735
I1139
Y1157
L733
C
L743
L1150
N
F1176
F744
F1171
I1131
I1170
F1154
L1125
D1159
I
C
N
J
Nup170 mutation
Nup145N binding
V1062A + + +
L1066A + + +
E1110A + +
L1111A + +
I1131A +
I1147A +
F1154A -
Y1157A +
Y1164A + + +
F1171A +
Nup145N
SHKKLVINKDMRTDLFSPPNKD
729
740
750

Fig. 3.

Structural and biochemical analyses of the Nup170 interaction network. (**A**) Domain structures of Nup170 and Nup53. Black lines indicate fragments used for crystallization. (**B**) Crystal structure of the $Nup170^{NTD}$•$Nup53^{R3}$ complex shown in cartoon representation. A 90º rotation is shown on the right. The potential WF and ALPS membrane interaction motifs are indicated. (**C**) Close-up view of the interaction between Nup170 and Nup53. Nup170 and Nup53 residues involved in the interaction are labeled in orange and purple, respectively. (**D**) Graphic summary of mutational analysis of the $Nup170^{NTD}$-$Nup53^{R3}$ interaction. Nup170 is shown in surface representation from the same view as in panel (C). Residues are colored in red, orange, yellow, and green to indicate mutations that had a strong, moderate, weak, or no effect on binding, respectively. (**E**) Tabular summary of tested Nup170 mutants and their effect on Nup53 binding; (+++) wild-type binding, (++) moderately weakened binding, (+) weak binding, and (-) no binding. Mutations in Nup53 that were tested for binding are indicated by dots above the sequence using the same color code as in panel (D). Chromatograms and representative SDS-PAGE gels are shown in fig. S22. (**F**) Domain structures of Nup170 and Nup145N. Black lines indicate fragments used for crystallization. (**G**) Crystal structure of the $Nup170^{CTD}$•$Nup145N^{R3}$ complex shown in cartoon representation. (**H**) Close-up view of the interaction between Nup170 and Nup145N. Nup170 and Nup145N residues involved in the interaction are labeled in orange and cyan, respectively. (**I**) Graphic summary of mutational analysis of the $Nup170^{CTD}$-$Nup145N^{R3}$ interaction. Nup170 is shown in surface representation from the same view as in panel (H). Coloring is according to panel (D). (**J**) Summary of mutational analysis of the $Nup170^{CTD}$-$Nup145N^{R3}$ interaction, coloring and key are same as in panel (E). Chromatograms and representative SDS-PAGE gels are shown in fig. S23.

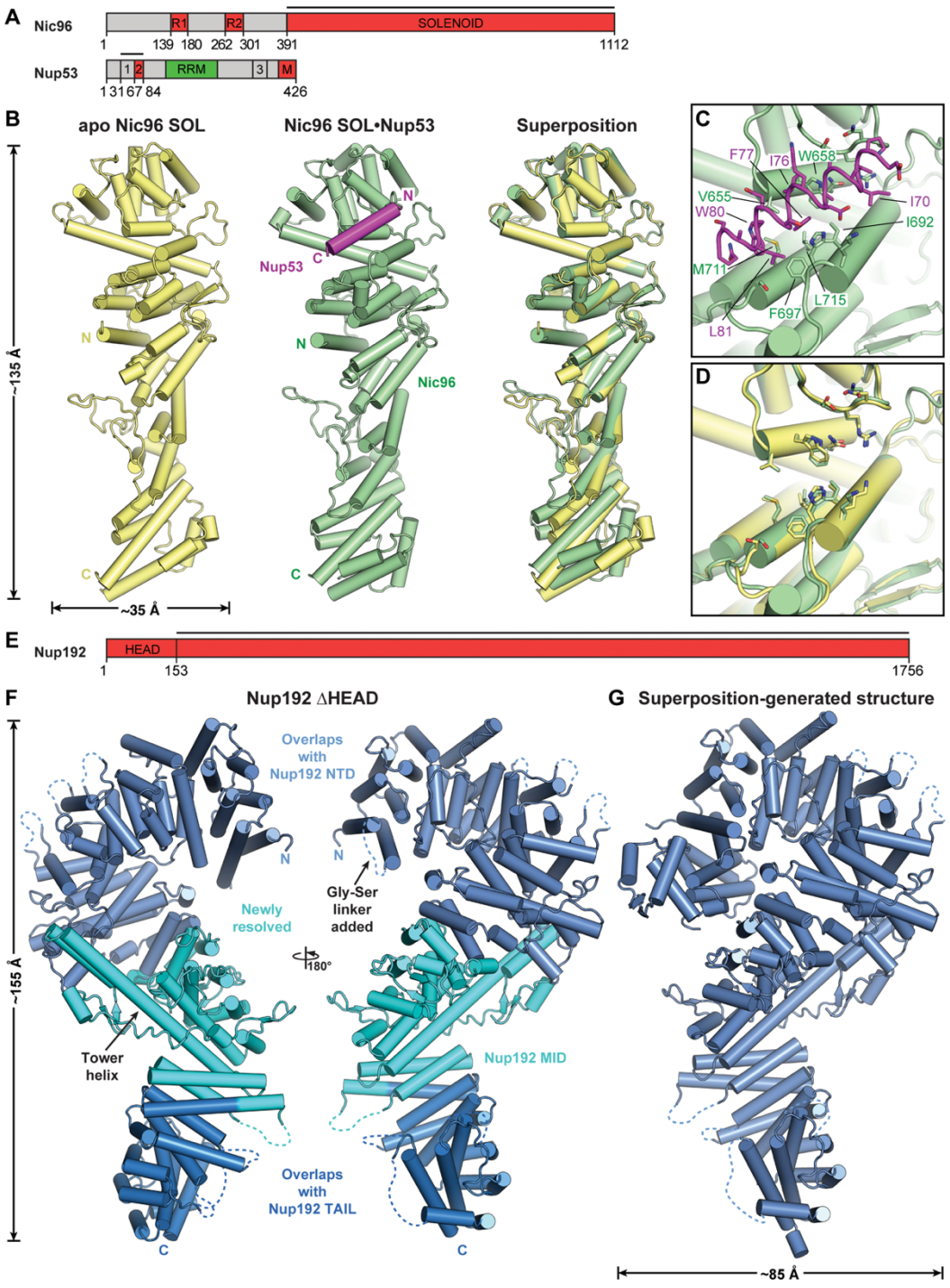


Fig. 4.

Structural analysis of the inner ring complex nucleoporins Nic96 and Nup192. (**A**) Domain structures of Nic96 and Nup53. Black lines indicate fragments used for crystallization. (**B**) Crystal structures of *apo* Nic96^{SOL} (yellow) and the Nic96^{SOL}•Nup53^{R2} complex (green and purple) and their superposition are shown in cartoon representation. (**C**) Close-up view of the Nup53^{R2}-binding site in Nic96^{SOL}. For clarity, Nup53 is shown in ribbon

representation. Nic96 and Nup53 residues involved in the interaction are labeled in green and purple, respectively. (**D**) Close-up view of the superposition of *apo* and Nup53-bound structures of Nic96^{SOL} reveals minimal conformational changes. (**E**) Domain structure of Nup192. A black line indicates the fragment used for crystallization. (**F**) Crystal structure of Nup192$^{\Delta HEAD}$ shown in cartoon representation. A 180º rotated view is shown on the right. Regions of Nup192 that were resolved in previous crystal structures are colored in shades of blue, while the region of the protein that was not included in previous crystallographic analyses is shaded cyan. (**G**) Structure of full-length Nup192 generated by superposing fragment crystal structures. See also fig. S31.

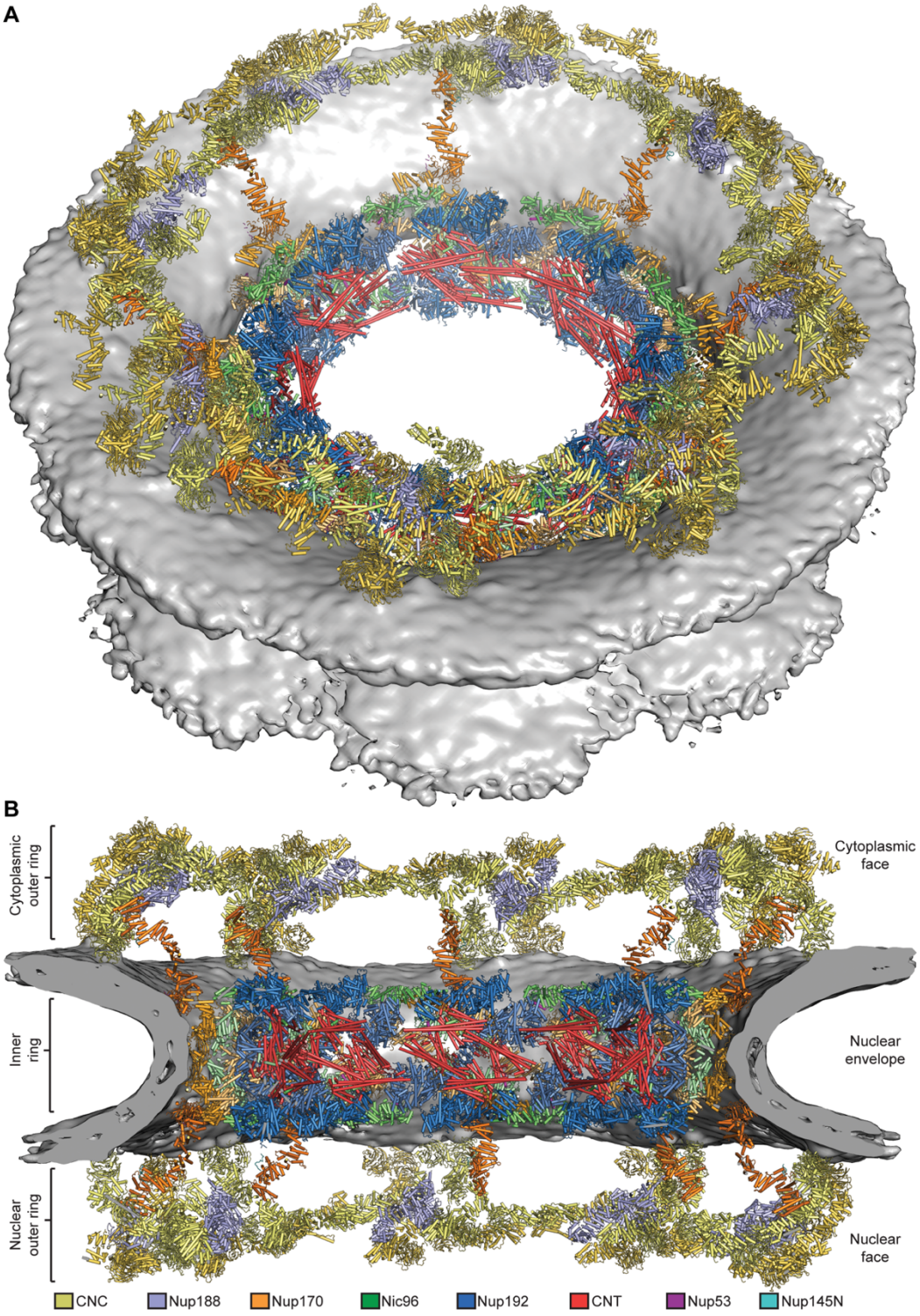


Fig. 5.

Architecture of the NPC symmetric core. Composite structure of the NPC symmetric core generated by docking nucleoporin and nucleoporin complex crystal structures into the cryoET reconstruction of the intact human NPC (EMD-3103). The density corresponding to

the nuclear envelope is shown as a gray surface. Proteins are color-coded according to the legend at the bottom. (**A**) View from above the cytoplasmic face and (**B**) a cross-sectional view from within the transport channel. Details are discussed in the text. See also figs. S34 to S37; S40;.

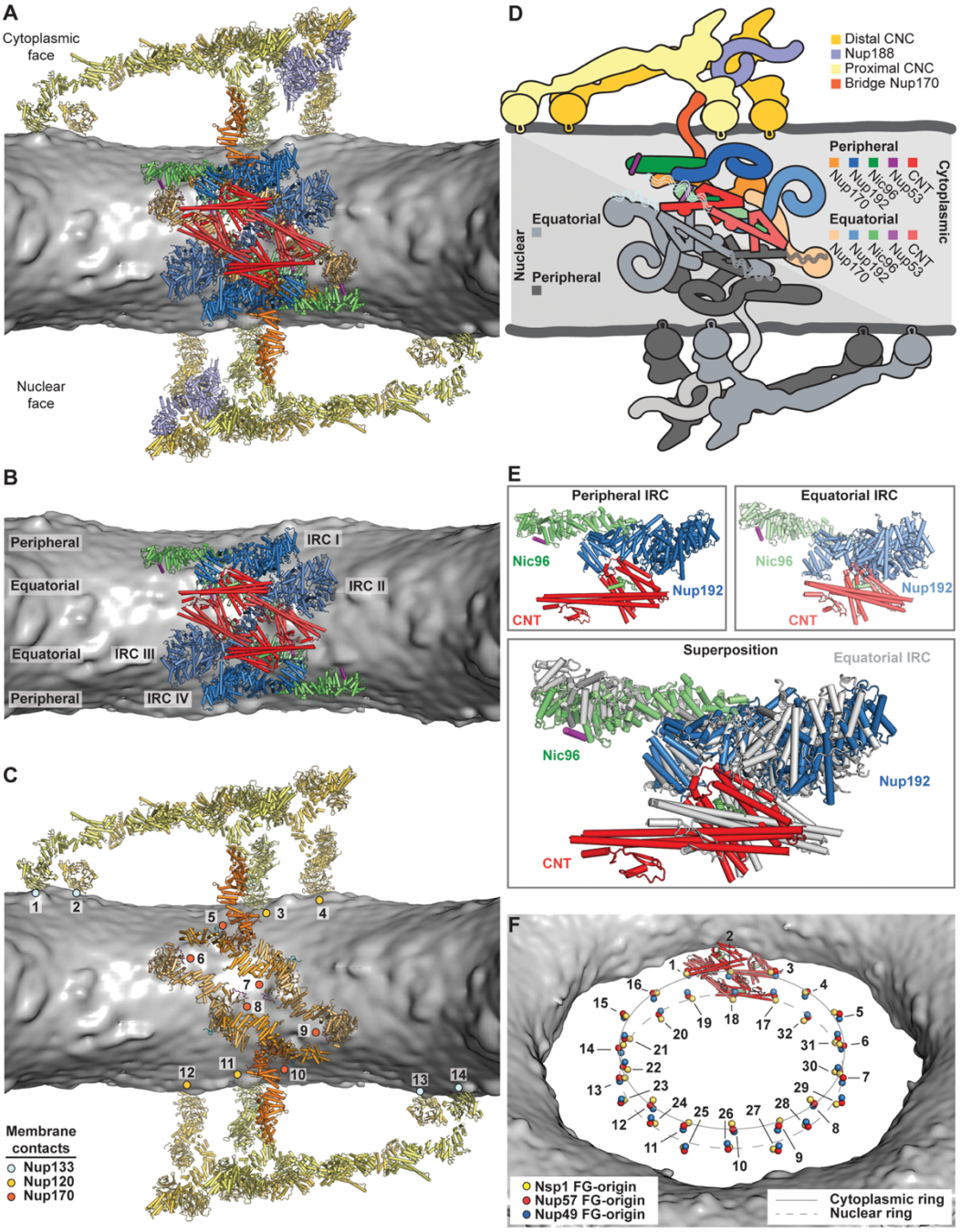


Fig. 6.

Architecture of the NPC spoke. (**A**) A single NPC symmetric core spoke is shown in cartoon representation from the same cross-sectional view as in Fig. 5B. Different shades of colors are used to indicate biochemically distinct dockings of the same protein. (**B**) Same view as in panel (A), but with Nup170 and the nucleoporins from the outer ring removed to highlight the organization of the four inner ring complexes (IRCs). (**C**) The membrane coat of the NPC. Nup192 and Nic96 molecules and CNTs have been removed for clarity. Contacts between the nuclear envelope and the ALPS motifs in Nup120, Nup133, and Nup170 are indicated by dots. (**D**) Schematic of the NPC spoke. The proteins corresponding to the nuclear

side of the spoke are colored in gray, demonstrating the two-fold rotational symmetry relating the cytoplasmic and nuclear halves of each spoke. (**E**) Equatorial and peripheral IRCs adopt similar conformations. Identical views of the cytoplasmic peripheral IRC, the cytoplasmic equatorial IRCs, and their superposition are shown. For clarity, the equatorial IRC is colored in gray in the superposition. (**F**) Organization of FG repeats in the inner ring. Colored spheres indicate the positions of the N-termini of the three channel nucleoporins for all 32 CNT copies in the inner ring. Despite four distinct CNT positions in each spoke, the N-termini for the cytoplasmic and nuclear complexes are arranged in approximately the same plane. Thus, while not possessing true 16-fold symmetry, the CNT N-termini approximately form two 16-membered rings, which are indicated by solid and dashed lines. Notably, the FG repeats of the 16-membered CNT rings project circumferentially into the central transport channel with opposite directionality; cytoplasmic CNT ring (counterclockwise), nuclear CNT ring (clockwise). The CNT molecules for a single spoke are shown in cartoon representation to indicate the directionality of the FG repeats.

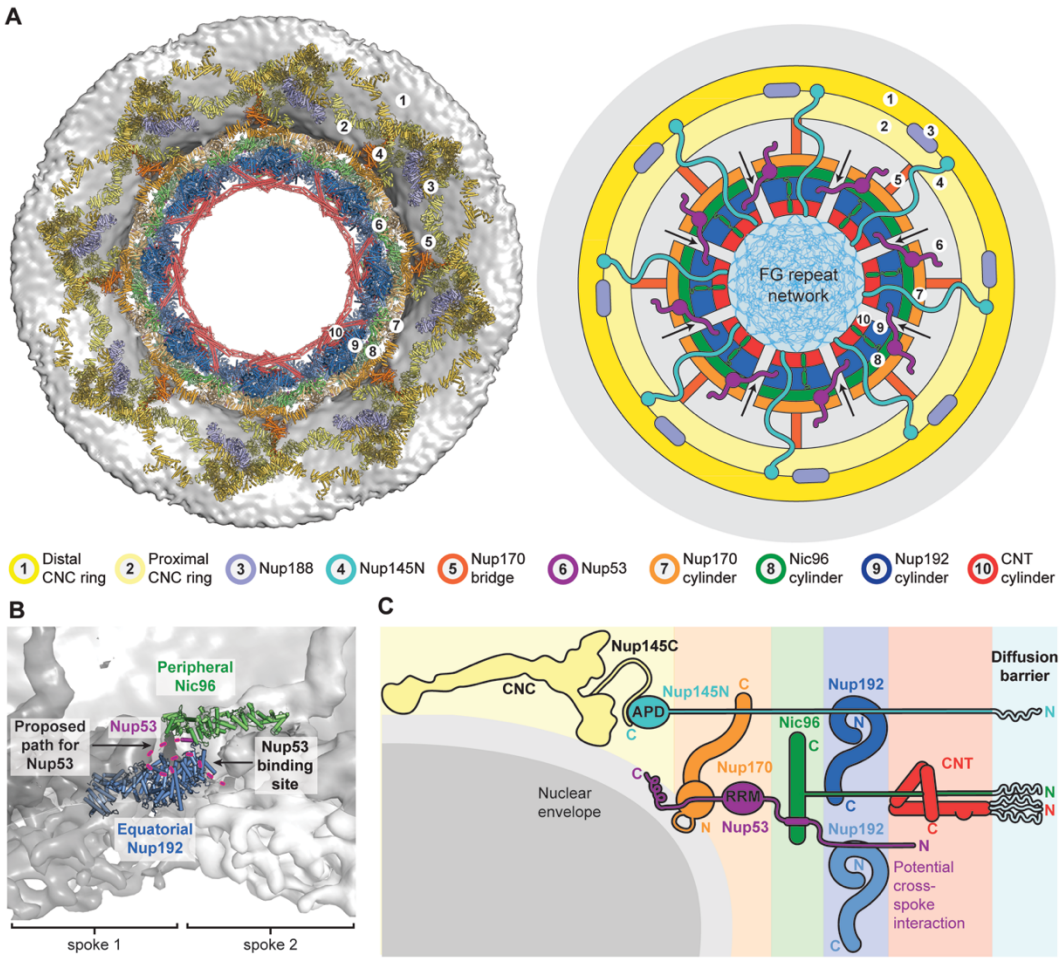


Fig. 7.

Cylindrical organization of the NPC. (**A**) A top view of the composite structure of the NPC symmetric core viewed from the cytoplasm is shown (left). Coloring is according to Fig. 5. Schematic representation on the right illustrates the distinct concentric cylinders observed in the symmetric core. Cartoons of the linker nucleoporins and the unstructured NTE of Nic96 are shown to indicate how they span across multiple cylinders. Arrows indicate gaps between the inner ring spokes, which represent proposed paths for inner nuclear membrane protein transport. (**B**) Only a single contact is observed between the two adjacent inner ring spokes. A close-up view of the inter-spoke interface in the inner ring is shown in an orientation similar to Fig. 5A. Two adjacent spokes are depicted as gray and white surfaces, on the left and right, respectively. An equatorial Nup192 molecule (spoke 1) and a peripheral Nic96 molecule bound to Nup53 (spoke 2) are shown in cartoon representation. The Nup53 binding site on Nup192 is labeled and the proposed path of the peptide is drawn as a purple dashed line. (**C**) A schematic representation illustrating the concentric cylinder organization of the symmetric core shown from the side. The order and arrangement of the binding sites for the linker nucleoporins observed in the composite structure are depicted. The proposed inter-spoke interaction between Nup53 and Nup192 is drawn at the bottom.

MATERIALS AND METHODS

Bacterial expression constructs. DNA fragments were amplified by polymerase chain reaction using C. thermophilum and H. sapiens cDNA. SUMO (small ubiquitin-like modifier)–tagged proteins were cloned into a modified pET28a or pET-MCN vector containing an N-terminal hexahistidine tag followed by a SUMO tag by using BamHI and NotI restriction sites (50, 51). Hexahistidine-tagged proteins were cloned into a modified pET28a vector containing an N-terminal hexahistidine tag followed by a protease cleavage site (PreScission) by using NdeI and NotI restriction sites (52). Glutathione S-transferase (GST)–tagged proteins were cloned into a pGEX-6P1 vector by using the BamHI and NotI restriction sites. Mutants were generated by QuikChange mutagenesis and confirmed by DNA sequencing. Details of bacterial expression constructs are shown in table S1.

Protein expression and purification. Proteins were expressed in Escherichia coli BL21-CodonPlus(DE3)-RIL cells (Stratagene) in Luria-Bertani media and induced at an OD 600 (optical density at 600 nm) of 0.8 with 0.5 mM isopropyl-β-D-thiogalactopyranoside. Details regarding expression times and temperatures are given in table S1. Seleno-L-methionine (SeMet)–labeled proteins were expressed with a methionine pathway inhibition protocol, as previously described (53). Cells were harvested by centrifugation and resuspended in a buffer containing 20 mM TRIS [tris(hydroxymethyl)aminomethane, pH 8.0], 500 mM sodium chloride, and 5 mM 2-mercaptoethanol (β-ME) for hexahistidine-tagged proteins, or a buffer containing 20 mM TRIS (pH 8.0), 200 mM sodium chloride, and 4 mM DTT for GST-tagged proteins; both were supplemented with complete EDTA-free protease inhibitor cocktail (Roche), unless otherwise noted. For purification, the cell suspensions of all proteins and protein complexes were supplemented with 1 mg deoxyribonuclease I (Roche) and lysed by using a cell disruptor (Avestin). Cell lysates were cleared by means of centrifugation at

30,000g for 1 hour. The supernatants were filtered through a 0.45-µm filter (Millipore) and purified using standard chromatography methods (details are given in table S2). Proteins were concentrated to ~10 to 20 mg/ml for biochemical interaction experiments, complex reconstitution, and crystallization. Representative chromatograms and SDS–polyacrylamide gel electrophoresis (SDS-PAGE) analyses are shown for the purifications for Nup120•Nup37•ELYS•Nup85, Sec13•Nup145C, Nup84•Nup133, Nic96•CNT, Nup192, Nup188, Nup170, Nup53, and Nup145N (figs. S51 to S59).

Reconstitution of Nup120•Nup37•ELYS•Nup85. Purified Nup120•Nup37•ELYS heterotrimer was mixed with a twofold molar excess of purified Nup85, incubated for 30 min on ice, and injected onto a HiLoad Superdex 200 16/60 prep grade (PG) gel filtration column equilibrated in 20 mM TRIS (pH 8.0), 150 mM sodium chloride, 5 mM dithiothreitol (DTT), and 5% (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~8 mg/ml for biochemical studies and further reconstitutions (see also fig. S51 for alternative copurification).

Reconstitution of Nup120•Nup37•ELYS•Nup85•Sec13•Nup145C (CNC hexamer). Purified Nup120•Nup37•ELYS•Nup85 heterotetramer was mixed with a twofold molar excess of purified Sec13•Nup145C heterodimer, incubated for 30 min on ice, and injected onto a HiLoad Superdex 200 16/60 PG gel filtration column equilibrated in 20 mM TRIS (pH 8.0), 150 mM sodium chloride, 5 mM DTT, and 5 % (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~8 mg/ml for biochemical studies and further reconstitutions.

Reconstitution of Nup120•Nup37•ELYS•Nup85•Sec13•Nup145C•Nup84•Nup133ΔNTE (CNC octamer) Purified Nup120•Nup37•ELYS•Nup85•Sec13•Nup145C (CNC hexamer) was mixed with a 1.1-fold molar excess of purified Nup84•Nup133ΔNTE heterodimer, incubated for 30 min on ice, and injected onto a HiLoad Superdex 200 16/60 PG gel filtration column equilibrated in 20 mM TRIS (pH 8.0), 150 mM sodium chloride, 5 mM DTT, and 5 % (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~8 mg/ml for biochemical studies and further reconstitutions.

Reconstitution of Nic96SOL•SUMO-Nup531–90. Purified Nic96SOL was mixed with a 1.2-fold molar excess of purified SUMO-Nup531–90, incubated on ice for 30 min, and injected over a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT. Complex-containing fractions were pooled and concentrated at room temperature to ~10 mg/ml for biochemical studies.

Purification of Nic96•CNT. Cells containing Nic96 and CNT were grown individually and resuspended in 50 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM β-ME, and 10 % (v/v) glycerol supplemented with complete EDTA-free protease inhibitor cocktail, 2 mM bovine lung aprotinin (Sigma), and 1 mM phenylmethanesulfonyl fluoride (Sigma). Before lysis, cell suspensions containing CNT and Nic96 were mixed in a 3:1 ratio. Clarified lysate was loaded onto a Ni–nitrilotriacetic acid column equilibrated in a buffer containing 25 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM β-ME, 20 mM imidazole, and 5 % (v/v) glycerol and eluted with an imidazole gradient. Peak fractions were pooled and desalted using a HiPrep 26/20 Desalting column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 150 mM sodium chloride, 5 mM DTT, and 5 % (v/v) glycerol in the presence

of ULP1 protease. After SUMO cleavage, the protein was loaded onto a MonoQ 5/50 GL column (GE Healthcare) equilibrated in 20 mM TRIS (pH 8.0), 150 mM sodium chloride, 5 mM DTT, and 5 % (v/v) glycerol and eluted with a sodium chloride gradient. Complex-containing fractions were pooled for further biochemical reconstitutions (see also fig. S54).

Reconstitution of Nup192•Nup145N, Nup192•Nup53, and Nup192•Nup145N•Nup53. Purified Nup192 was mixed with a 1.5-fold molar excess of purified Nup53 and/or Nup145N. The mixture was incubated on ice for 30 min before injection onto a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT. Complex-containing fractions were pooled and concentrated to ~10 mg/ml for biochemical studies and further reconstitutions.

Reconstitution of Nup192•Nic96•Nup53•CNT and Nup192•Nic96•Nup145N•Nup53•CNT (IRC•Nup53). Purified Nup192•Nup53 heterodimer or Nup192•Nup145N•Nup53 heterotrimer was mixed with a twofold molar excess of purified Nic96•CNT heterotetramer, incubated for 30 min on ice, and injected onto a HiLoad Superdex 200 16/60 PG gel filtration column equilibrated in 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT, and 2% (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~5 mg/ml for biochemical studies.

Reconstitution of Nup82NTD•Nup159T•Nup145N. Purified Nup82NTD•Nup159T heterodimer (T, TAIL; residues 1440 to 1481) was mixed with a 1.2-fold molar excess of purified Nup145N, incubated on ice for 30 min, and injected onto a HiLoad Superdex 200 16/60 PG column equilibrated in buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium

chloride, and 5 mM DTT. Complex-containing fractions were pooled and concentrated to ~30 mg/ml for biochemical studies.

Reconstitution of Nup170•Nup53, Nup170•Nup145N, and Nup170•Nup53•Nup145N. Purified Nup170 was mixed with a twofold molar excess of purified Nup53 and/or Nup145N. The mixture was incubated on ice for 30 minutes before injection onto a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 200 mM sodium chloride, 5 mM DTT, and 5% (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~10 mg/ml for biochemical studies and further reconstitutions.

Reconstitution of Nup170•Nup53•Nup192. Purified Nup170•Nup53 heterodimer was mixed with a 1.2-fold molar excess of purified Nup192, incubated on ice for 30 min, and injected onto a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 150 mM sodium chloride, and 5 mM DTT. Complex-containing fractions were pooled and concentrated to ~8 mg/ml for biochemical studies.

Reconstitution of Nup188•Nup145N. Purified Nup188 was mixed with a twofold molar excess of purified Nup145N, incubated on ice for 30 min, and injected onto a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT. Complex-containing fractions were pooled and concentrated to ~10 mg/ml for biochemical studies and further reconstitutions.

Reconstitution of Nup188•Nic96•Nup53•CNT. Purified Nup188 was mixed with twofold molar excess of purified Nup53 and Nic96•CNT heterotetramer, incubated for 30 min on ice, and injected onto a HiLoad Superdex 200 16/60 PG gel filtration column equilibrated in 20

mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT, and 2% (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~5 mg/ml for biochemical studies.

Reconstitution of scNup170CTD•scNup145N and scNup157CTD•scNup145N. Purified scNup170CTD or scNup157CTD was mixed with a 1.5-fold molar excess of purified scNup145N. The mixture was incubated on ice for 30 min prior to injection over a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 300 mM sodium chloride, 5 mM DTT, and 5% (v/v) glycerol. Complex-containing fractions were pooled and concentrated to ~10 mg/ml for biochemical studies.

Reconstitution of scNup170CTD•scNup100. Purified scNup170CTD was mixed with a 1.5-fold molar excess of purified scNup100, incubated on ice for 30 min, and injected onto a HiLoad Superdex 200 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5% (v/v) glycerol, and 5 mM DTT. Complex-containing fractions were pooled and concentrated to ~10 mg/ml for biochemical studies.

Multiangle light scattering coupled to analytical size-exclusion chromatography. Purified proteins and complex formations were characterized by inline multiangle light scattering (MALS) after separation on a Superdex 200 10/300 GL or a Superose 6 10/300 GL column. All experiments conducted using the Superdex 200 10/300 GL column were performed in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT, whereas experiments conducted using the Superose 6 10/300 GL column were performed in a buffer containing 20 mM TRIS (pH 8.0), 150 mM sodium chloride, and 5 mM DTT. The chromatography system was connected in series with an 18-angle light-scattering detector (DAWN HELEOS II, Wyatt Technology), a dynamic light-scattering

detector (DynaPro Nanostar, Wyatt Technology), and a refractive index detector (Optilab t-rEX, Wyatt Technology). Data were collected at 21°C every 1 s at a flow rate of 0.4 ml/min and analyzed using ASTRA 6 software, thereby generating molar mass and mass distribution (polydispersity) measurements of the samples (54). For interaction studies, proteins were mixed and preincubated on ice for 30 min before being applied to the gel filtration column. Protein-containing fractions were analyzed by SDS-PAGE followed by Coomassie brilliant blue staining. Experimental and theoretical masses of all proteins and complexes are listed in table S10.

Immunoblotting. Protein from peak fractions was resolved on a 12% SDS-PAGE gel, transferred to polyvinylidene difluoride membranes, and immunoblotted with a mouse antibody against the aviTag (Genscript, 1:1000 dilution), followed by an alkaline phosphatase-conjugated goat antibody against mouse immunoglobulin G (Promega, 1:4000 dilution). Blots were developed with SigmaFast bromochloroindolyl phosphate–nitro blue tetrazolium tablets (Sigma) for 2 to 3 min and imaged immediately.

Crystallization and structure determination. Crystallization trials for all proteins were performed at 21°C in hanging drops containing 1 μl of protein and 1 μl of reservoir solution. SeMet-labeled crystals were grown under similar conditions. Details of the crystallization and cryoprotection are given in table S4. X-ray diffraction data were collected at 100 K at beamline 8.2.2 at the Advanced Light Source (ALS) at Lawrence Berkley National Laboratory, beamline BL12-2 at the Stanford Synchrotron Radiation Source (SSRL), and beamline GM/CA-CAT 23ID-D at the Advanced Photon Source (APS). The x-ray diffraction data were processed using XDS (55). Structures were phased with Phaser, and maps were improved by density modification with RESOLVE (56, 57). Iterative rounds of model

building and refinement were performed using COOT and PHENIX (58, 59). All models had excellent stereochemistry and geometry, as determined by Molprobity (60). Details of refinement and data processing are given in tables S5 to S9.

Structure determination of Nup170NTD•Nup53329–361. Purified Nup170NTD was mixed with a threefold molar excess of synthesized Nup53329–361 peptide (Zhejiang Ontores Biotechnologies). To obtain high-resolution diffraction, residues 293 to 305 of Nup170NTD, which were predicted to reside in a large loop, were deleted. Crystals were grown in 1.0 M sodium potassium phosphate (pH 6.9) and cryoprotected by gradual addition of glycerol to the crystallization drop. The Nup53329–361 peptide was added to all cryoprotectant solutions. The structure was determined by single-wavelength anomalous dispersion (SAD) with Phaser, using anomalous x-ray diffraction data collected from a crystal containing SeMet-labeled Nup170NTD.

Structure determination of apo Nup170CTD. A quadruple methionine mutant (Y905M, L1007M, L1183M, and V1292M) was used for crystallization and collection of SeMet SAD data for apo Nup170CTD. Crystals were grown in 0.1 M MES (pH 6.3), 10% (w/v) PEG-20,000 (polyethylene glycol, molecular weight 20,000), 10% (v/v) ethylene glycol, and 0.2 M potassium thiocyanate and cryoprotected by gradual addition of ethylene glycol to the crystallization drop. The four additional methionine positions were introduced as sequence markers for high-confidence sequence assignment. An initial model was used for molecular replacement in Phaser with diffraction data collected from crystals of native apo Nup170CTD that diffracted to 2.1 Å resolution. Crystals of native apo Nup170CTD were grown in 0.1 M Hepes (pH 7.0) and 1.0 M sodium acetate and cryoprotected by gradual addition of ethylene glycol to the crystallization drop.

Structure determination of Nup170CTD•Nup145N729–750. Purified Nup170CTD was incubated with a threefold molar excess of synthesized Nup145N729–750 peptide (Zhejiang Ontores Biotechnologies). Attempts to obtain crystals with stoichiometric Nup145N incorporation revealed that two regions of Nup170CTD could bind to the Nup145N-binding pocket (1403 to 1416 and 1375 to 1377), blocking the Nup170-Nup145N interaction. To obtain crystals of the complex, both of these regions of Nup170 were deleted. Complex crystals were grown in 0.2 M lithium acetate and 12% (w/v) PEG-3350. Crystals were cryoprotected by gradual addition of ethylene glycol to the crystallization drop. Phasing was performed by means of molecular replacement in Phaser, using the structure of apo Nup170CTD as a search model.

Structure determination of Nup170SOL. Crystals of Nup170SOL were grown in 0.1 M Hepes (pH 7.1) and 0.5 M ammonium sulfate and cryoprotected by gradual addition of ethylene glycol to the crystallization drop. The Nup170SOL structure was determined with the MR-SAD routine in PHENIX, combining molecular replacement phases obtained by using the corresponding fragments of the Nup170NTD and Nup170CTD structures with SAD phases obtained from anomalous x-ray diffraction data that were collected from a crystal containing SeMet-labeled Nup170SOL. The crystals diffracted with a high degree of anisotropy, with diffraction past 3.9 Å for two axes, but diffraction was limited to 4.5 Å for the third axis. Minimal model building was performed after the high-resolution structures were refined as rigid bodies, with the exception of the loops connecting helices α16, α17, and α18, which were not observed in the other structures. The overall correctness of the observed conformation was confirmed by calculating anomalous difference Fourier maps to locate the methionine positions. Refinement was performed with anisotropically truncated data, which yielded more interpretable maps.

Structure determination of apo Nic96SOL and Nic96SOL•Nup5331–84. The structure of apo Nic96SOL was determined using SAD x-ray diffraction data collected from crystals grown with SeMet-labeled Nic96SOL in 4% (v/v) tacscimate (pH 7.4) and 14% (w/v) PEG-3350 and cryoprotected by exchanging the crystallization buffer with paratone-N. Purified Nic96SOL•Nup5331–84 was crystallized in 0.1 M TRIS (pH 7.0) and 12% (w/v) PEG-6000, and 1.2 M sodium chloride, and crystals were also cryoprotected by exchanging the crystallization buffer with paratone-N. An initial model of apo Nic96SOL was built using the experimental electron density map and used as a search model in molecular replacement to determine the structure of Nic96SOL•Nup5331–84 at 2.65 Å resolution. Residues 31 to 66 of Nup53 are presumed to be disordered, because only residues 67 to 84 were visible in the electron density map.

Structure determination of Nup192ΔHEAD. Despite extensive optimization and screening, crystals of Nup192ΔHEAD diffracted to ~10 Å resolution at best. Inspection of the crystal structure of Nup192NTD suggested helix α9, which protrudes from the surface of the protein, might be preventing the formation of stable crystal contacts. Thus, we replaced residues 167 to 184 with a GSGS linker. Crystals of this mutant grown in 0.1 M TRIS (pH 7.9), 6% (w/v) PEG-4000, and 5% (v/v) polypropylene glycol displayed improved diffraction to ~5 Å resolution. While screening for heavy metal derivatives, we noticed that the diffraction quality of these crystals was improved further by soaking with 0.2-μl saturated potassium hexachloroosmate (K2OsCl6) solution in a total drop volume of 5 μl for 1 min. Crystals were cryoprotected by gradually supplementing the drop with polypropylene glycol in 5% steps. The structure was determined by molecular replacement with Phaser, using the corresponding fragments of Nup192NTD [Protein Data Bank identifier (PDB ID) 4KNH] and Nup192TAIL (PDB ID 5CWV), followed by SAD phasing, using anomalous x-ray

diffraction data collected at the osmium L3 edge (15, 20). Although the crystals only grew in the presence of Nup5331–67 peptide, no strong electron density was observed for the peptide, probably because of partial occupancy of the peptide in the crystal.

Structure determination of Nup53RRM. Nup53RRM crystallized in two different crystal forms. Crystals grown in 0.1 M Hepes (pH 7.0), 24% (w/v) PEG-3350, and 0.2 M potassium iodide crystallized in the space group P212121. Crystals were cryoprotected by gradual addition of ethylene glycol to the crystallization drop. The crystal structure was determined by SAD phasing with Phaser, using anomalous x-ray diffraction data from a crystal grown with SeMet-labeled Nup53RRM. Iodide ions were placed in the native structure based on peaks in the anomalous difference Fourier map. These crystals diffracted exceptionally well, with diffraction clearly visible up to 0.7 Å resolution, but the geometry of the beamline did not permit positioning of the detector close enough to measure all spots with very high Bragg angles. As a result, the completeness of the data in the highest-resolution shells is low, but these data were included because they improved the quality of the electron density maps. The second crystal form of Nup53RRM crystallized in the space group P3121 in 4% (v/v) tacsimate (pH 4.1) and 15% (w/v) PEG-3350. The crystals were cryoprotected in 32% (v/v) tacsimate (pH 4.1) and 19% (w/v) PEG-3350. The structure was determined by SAD phasing with Phaser, using anomalous x-ray diffraction data collected from potassium hexachloroosmate–derivatized native crystals.

Structure determination of Nup145NAPD•Nup145C^{N}. Crystals of Nup145NAPD were grown in 0.1 M citric acid (pH 3.0) and 25% (w/v) PEG-3350 and cryoprotected with mother liquor supplemented with 25% (v/v) ethylene glycol. The structure was determined by molecular replacement with Phaser, using the crystal structure of Nup145NAPD (PDB ID

5CWW) as a search model (15). We also crystallized a catalytically dead mutant, Nup145NAPD T994A, that was fused to the N-terminal six residues of Nup145C. Crystals grew in 0.1 M TRIS (pH 8.8), 32% (w/v) PEG-4000, and 0.2 M lithium sulfate and were cryoprotected with mother liquor supplemented with 10% (v/v) glycerol. The structure was determined by molecular replacement with Phaser, using the crystal structure of Nup145NAPD as the search model.

Generation of full-length structures by superposition for docking into cryo-ET reconstructions. Superposition-generated structures of full-length Nup170 and Nup192 were generated by careful superposition of crystal structures of the respective fragments. For Nup170, the crystal structure Nup170NTD•Nup53329–361 was superposed onto the crystal structure of Nup170SOL. Minor differences in the conformations of helices were observed at the truncated ends of the structures, but the conformation included in the models always corresponded to the structure in which no truncation was made. The various structures of Nup170CTD were superposed onto the crystal structure of Nup170SOL by using only helices α19 to α24, which displayed only minimal conformation rearrangements. In the superposition-generated full-length Nup170 structure, residues 1 to 808 were derived from the Nup170NTD•Nup53329–361 crystal structure, residues 809 to 844 were derived from the Nup170SOL crystal structure, and residues 845 to 1416 were derived from the various Nup170CTD crystal structures (see also fig. S25). A similar approach was used to superpose the structure of Nup192NTD (PDB ID 4KNH) and Nup192TAIL (PDB ID 5CWV) onto the crystal structure of Nup192ΔHEAD, superposing only helices that had identical conformations in both crystal structures (15, 20). In the full-length Nup192 superposition-generated structure, residues 1 to 253 were derived from the Nup192NTD crystal structure,

residues 254 to 1750 from the Nup192ΔHEAD crystal structure, and residues 1750 to 1756 from the Nup192TAIL crystal structure (see also fig. S31).

Docking crystal structures into the cryo-ET reconstruction of the intact human NPC. An incremental approach was used to dock available crystal structures and superposition-generated structures into the previously reported ~23 Å cryo-ET reconstruction of the intact human NPC (28). For all structures other than Nup37, Nup43, and Nup133NTD, global searches in the cryo-ET reconstruction were performed using the UCSF Chimera Fit Map command (61). Searches consisted of 50,000 initial placements that were locally optimized and scored based on correlation between a simulated 20 Å map of the docked model and the cryo-ET reconstruction, taking into account the C8 rotational symmetry of the latter. The top-scoring solutions for the search were inspected for their agreement with previously published biochemical data and the biochemical data presented here. A flowchart of our incremental docking approach can be found in fig. S34. The average Fourier shell correlation between a single spoke and the map corresponding to the docked spoke was calculated using Refmac5 to a value of 0.79 at the reported resolution of the map (23 Å) (62).

Docking of the CNC. Searches for the CNC hexamer (PDB ID 4XMM) and the hsNup84•hsNup133 heterodimer (PDB ID 3CQC) were performed in maps from which the density corresponding to the nuclear envelope had been removed (14, 38). Because of its large size and distinctive shape, the four top-scoring solutions of the CNC hexamer have much higher scores than any other solution. Nup37 was docked by superposing the crystal structure of the S. pombe Nup120NTD•Nup37 heterodimer (PDB ID 4FHN) onto the CNC hexamer (39). hsNup43 was docked into an unoccupied density next to Nup85, based on a previously identified cross-link and the previously reported interaction (13, 41). hsNup133

(PDB ID 1XKS) was docked in an unoccupied density next to the Nup120 β-propeller, and the fit was optimized locally with the Fit Map command (63). Because the resolution of the map cannot distinguish between different orientations of the Nup43 and Nup133 β-propellers, their exact orientation is ambiguous. Details regarding CNC docking can be found in fig. S35.

Docking of Nup170, Nup192, and Nup188. Before performing global searches for Nup170 and Nup192, the density corresponding to the docked CNCs was removed to obtain placements that would not clash with the CNCs. Searches in maps that included the density assigned to the CNCs produce similar results. Searches with Nup170 were performed using multiple conformations. Most conformations produced similar results, referred to as conformation I, and occupied two symmetry-related positions in the inner ring. A different conformation (conformation II), which was derived from the structure of the apo Nup170CTD quadruple methionine mutant, occupied symmetry-related positions bridging the inner and outer rings. For Nup192, we accepted five placements from the global search results, two placements in the inner ring, two placements in the outer rings, and one placement on the nuclear peripheral side of the inner ring. A matching cytoplasmic peripheral placement was not found in global searches, but the manual placement on the cytoplasmic peripheral side of the inner ring generated a score that would have ranked as the 9th highest. The four molecules placed in the inner ring were assigned as Nup192, and the two molecules placed in the outer rings were assigned as Nup188, based on the following considerations: (i) Nup188 binding to Nup145N was incompatible with both Nup192 and Nup170, making it unlikely that Nup188 is a component of the tightly packed inner ring core; (ii) cross-linking and mass spectrometry data identified cross-links between Nup188 and Nup85 (13, 42); and (iii) Nup188 mislocalization behavior was found to be consistent with other outer ring

components upon genetic perturbation (64). Details regarding Nup170, Nup192, and Nup188 docking can be found in fig. S36.

Docking of Nic96SOL•Nup5331–84 and CNT•Nic96R1. Searches for the Nic96SOL•Nup5331–84 complex were performed using a map of the inner ring from which the density corresponding to the placed Nup170, Nup188, and Nup192 molecules had also been removed. Many high-scoring solutions of Nic96SOL were related to each other by rotations along the long axis of the rod-shaped molecule, reflecting the ambiguity of the exact rotational orientation of the molecule at the resolution of the cryo-ET reconstruction. Searches for the CNT•Nic96R1 complex (PDB ID 5CWS) were performed in a map from which the density corresponding to the placed Nic96SOL molecules also had been removed (15). Superposition of the X. laevis (xl) CNT crystal structures (PDB IDs 5C2U and 5C3L) was performed by superposing xlNup54 from the xlCNT structure directly onto Nup57 from the CNT structure, and then superposing the ferredoxin-like domain onto the xlCNT structure (16). Details regarding Nic96SOL•Nup5331–84 and CNT•Nic96R1docking can be found in fig. S37.

Manual identification of a third pair of Nup170 molecules. After successful placement of 32 copies each of Nup170, Nup192, Nic96SOL, and the CNT in the inner ring, we continued our analysis by removing the density corresponding to the docked molecules. There were four major distinct densities remaining in the inner ring, each of which appeared on both the nuclear and cytoplasmic sides (fig. S40). The first set of densities corresponded to the ferredoxin-like domain that is only present in metazoan Nup57. The second set corresponded to the density adjacent to the C-terminal TAIL domains of the equatorial Nup192 molecules, which we assume arises from our structure imperfectly capturing the conformation of

Nup192 in the assembled NPC. The third major density was adjacent to the equatorial Nup170 β-propellers. We were unable to assign this density confidently to any remaining subunit, such as Nup53RRM, Nup145NAPD, or any other nucleoporin of known structure. Lastly, a pair of extended densities was adjacent to the nuclear envelope (fig. S40). These densities were very similar in shape to the Nup170 molecules that we had already placed, but neither of the conformations that we previously placed yielded a high-quality fit. Instead, we found the best fit with a third conformation, although visible differences suggest that even this structure did not perfectly capture the conformation of this third molecule inside the NPC. Details regarding docking for this conformation of Nup170 can be found in fig. S40.

Protein labeling for oligomerization experiments. Proteins were labeled under conditions that were selective for the N-terminal amino group of proteins. Before the labeling reaction, purified CNC hexamer, Nup84•Nup133, and Nup84•Nup133ΔNTE complexes at concentrations of 10 to 20 μM were dialyzed in a buffer containing 100 mM sodium bicarbonate (pH 8.0), 100 mM sodium chloride, and 5 mM DTT. 20 μM Nup170 and Nup170ΔC were dialyzed against a buffer containing 100 mM sodium bicarbonate (pH 8.0), 200 mM sodium chloride, and 5 mM DTT. For the labeling reactions, BODIPY (boron-dipyrromethene) FL NHS Ester (succinimidyl ester) or Alexa Fluor 647 NHS Ester (succinimidyl ester) dyes (Molecular Probes) were added at ratios equimolar to the dialyzed protein samples and incubated for 50 min at room temperature. The reaction was subsequently quenched by dialysis against buffers containing 100 mM TRIS (pH 8.0) instead of sodium bicarbonate.

Oligomerization experiments. Oil droplet formation was analyzed by fluorescence microscopy with a Carl Zeiss AxioImagerZ.1 equipped with an AxioCamMRm camera.

Experiments performed without Nup170 were performed with BODIPY-labeled CNC hexamer and Alexa Fluor 647–labeled Nup84•Nup133 or Nup84•Nup133ΔNTE. Proteins were mixed at equimolar ratios directly on a cover slide and analyzed immediately. For experiments involving the incorporation of Nup170, BODIPY-labeled CNC hexamer and unlabeled Nup84•Nup133 were first mixed together in equimolar ratios and then subsequently incubated with a 1.5-fold molar excess of Alexa Fluor 647–labeled Nup170 or Nup170ΔC.

Figures and movies. Gel filtration profiles and MALS graphs were generated in IGOR (WaveMetrics) and assembled in Adobe Illustrator. Sequence alignments were generated using MUSCLE and colored with ALSCRIPT (65, 66). Electrostatic potentials were calculated with APBS (Adaptive Poisson-Boltzmann Solver) software (67). All structural figures and movies were generated using PyMol (www.pymol.org).

REFERENCES

1. A. Hoelz, E. W. Debler, G. Blobel, The structure of the nuclear pore complex. *Annu. Rev. Biochem.* **80**, 613-643 (2011).
2. A. Ibarra, M. W. Hetzer, Nuclear pore proteins and the control of genome functions. *Genes Dev.* **29**, 337-349 (2015).
3. V. Le Sage, A. J. Mouland, Viral subversion of the nuclear pore complex. *Viruses* **5**, 2019-2042 (2013).
4. A. Kohler, E. Hurt, Gene regulation by nucleoporins and links to cancer. *Mol. Cell* **38**, 6-15 (2010).
5. K. Zhang *et al.*, The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61 (2015).
6. A. C. Woerner *et al.*, Cytoplasmic protein aggregates interfere with nucleo-cytoplasmic transport of protein and RNA. *Science* **351**, 173-176 (2015).
7. H. M. Kaneb *et al.*, Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **24**, 1363-1373 (2015).
8. B. D. Freibaum *et al.*, GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129-133 (2015).
9. A. Jovicic *et al.*, Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* **18**, 1226-1229 (2015).
10. A. Cook, F. Bono, M. Jinek, E. Conti, Structural biology of nucleocytoplasmic transport. *Annu. Rev. Biochem.* **76**, 647-671 (2007).

11. K. Thierbach *et al.*, Protein interfaces of the conserved Nup84 complex from Chaetomium thermophilum shown by crosslinking mass spectrometry and electron microscopy. *Structure* **21**, 1672-1682 (2013).

12. K. Kelley, K. E. Knockenhauer, G. Kabachinski, T. U. Schwartz, Atomic structure of the Y complex of the nuclear pore. *Nat. Struct. Mol. Biol.* **22**, 425-431 (2015).

13. K. H. Bui, A. von Appen *et al.*, Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* **155**, 1233-1243 (2013).

14. T. Stuwe, A. R. Correia *et al.*, Architecture of the nuclear pore complex coat. *Science* **347**, 1148-1152 (2015).

15. T. Stuwe, C. J. Bley, K. Thierbach, S. Petrovic *et al.*, Architecture of the fungal nuclear pore inner ring complex. *Science* **350**, 56-64 (2015).

16. H. Chug, S. Trakhanov, B. B. Hulsmann, T. Pleiner, D. Gorlich, Crystal structure of the metazoan Nup62*Nup58*Nup54 nucleoporin complex. *Science* **350**, 106-110 (2015).

17. J. Fischer, R. Teimer, S. Amlacher, R. Kunze, E. Hurt, Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. *Nat. Struct. Mol. Biol.* **22**, 774-781 (2015).

18. N. Eisenhardt, J. Redolfi, W. Antonin, Interaction of Nup53 with Ndc1 and Nup155 is required for nuclear pore complex assembly. *J. Cell Sci.* **127**, 908-921 (2014).

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

20. T. Stuwe, D. H. Lin, L. N. Collins, E. Hurt, A. Hoelz, Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. *Proc. Natl. Acad. Sci. USA* **111**, 2530-2535 (2014).

21. H. S. Seo, B. J. Blus, N. Z. Jankovic, G. Blobel, Structure and nucleic acid binding activity of the nucleoporin Nup157. *Proc. Natl. Acad. Sci. USA* **110**, 16450-16455 (2013).

22. K. R. Andersen *et al.*, Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* **2**, e00745 (2013).

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

24. N. Schrader *et al.*, Structural basis of the nic96 subcomplex organization in the nuclear pore channel. *Mol. Cell* **29**, 46-55 (2008).

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

26. P. Sampathkumar *et al.*, Structure, dynamics, evolution, and function of a major scaffold component in the nuclear pore complex. *Structure* **21**, 560-571 (2013).

27. S. Amlacher *et al.*, Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell* **146**, 277-289 (2011).

28. A. von Appen *et al.*, In situ structural analysis of the human nuclear pore complex. *Nature* **526**, 140-143 (2015).

29. M. T. Teixeira *et al.*, Two functionally distinct domains generated by in vivo cleavage of Nup145p: a novel biogenesis pathway for nucleoporins. *EMBO J.* **16**, 5086-5097 (1997).

30. A. E. Hodel *et al.*, The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. *Mol. Cell* **10**, 347-358 (2002).

31. J. A. Killian, G. von Heijne, How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429-434 (2000).

32. G. Drin *et al.*, A general amphipathic alpha-helical motif for sensing membrane curvature. *Nat. Struct. Mol. Biol.* **14**, 138-146 (2007).

33. S. J. Kim *et al.*, Integrative structure-function mapping of the nucleoporin Nup133 suggests a conserved mechanism for membrane anchoring of the nuclear pore complex. *Mol. Cell. Proteomics* **13**, 2911-2926 (2014).

34. D. I. Fernandez, T. H. Lee, M. A. Sani, M. I. Aguilar, F. Separovic, Proline facilitates membrane insertion of the antimicrobial peptide maculatin 1.1 via surface indentation and subsequent lipid disordering. *Biophys J.* **104**, 1495-1507 (2013).

35. E. Laurell *et al.*, Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. *Cell* **144**, 539-550 (2011).

36. D. Flemming *et al.*, Analysis of the yeast nucleoporin Nup188 reveals a conserved S-like structure with similarity to karyopherins. *J. Struct. Biol.* **177**, 99-105 (2012).

37. M. Eibauer *et al.*, Structure and gating of the nuclear pore complex. *Nat. Commun.* **6**, 7532 (2015).

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

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

40. H. S. Seo *et al.*, Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. *Proc. Natl. Acad. Sci. USA* **106**, 14281-14286 (2009).

41. C. Xu *et al.*, Crystal structure of human nuclear pore complex component NUP43. *FEBS Lett.* **589**, 3247-3253 (2015).

42. D. I. Kim *et al.*, Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci. USA* **111**, E2453-2461 (2014).

43. A. Ori *et al.*, Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol. Syst. Biol.* **9**, 648 (2013).

44. D. Devos *et al.*, Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol.* **2**, e380 (2004).

45. S. S. Katta, C. J. Smoyer, S. L. Jaspersen, Destination: inner nuclear membrane. *Trends Cell Biol.* **24**, 221-229 (2014).

46. R. Ungricht, U. Kutay, Establishment of NE asymmetry-targeting of membrane proteins to the inner nuclear membrane. *Curr. Opin. Cell Biol.* **34**, 135-141 (2015).

47. P. Li *et al.*, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012).

48. C. P. Lusk *et al.*, Nup53p is a target of two mitotic kinases, Cdk1p and Hrr25p. *Traffic* **8**, 647-660 (2007).

49. F. Alber *et al.*, The molecular architecture of the nuclear pore complex. *Nature* **450**, 695-701 (2007).

ACKNOWLEDGEMENTS

We thank Jee-Young Mock, Stefan Petrovic, and Alina Patke for critical reading of the manuscript. We thank Martin Beck for sharing the cryoET reconstruction of the intact human NPC prior to publication, Ed Hurt for providing material and discussions, and Camille Bayes, Slava Butkovich, Christopher Frick, Antoine Koehl, Hanna Kuberczyk, Roland Tran, and Stephan Zimmerman for experimental support. We also acknowledge Jens Kaiser and the scientific staff of the Stanford Synchrotron Radiation Laboratory (SSRL) Beamline 12-2, the National Institute of General Medical Sciences and National Cancer Institute Structural Biology Facility (GM/CA) at the Advanced Photon Source (APS), and the Advanced Light Source (ALS) beamline 8.2.1 for their support with X-ray diffraction measurements. We acknowledge the Gordon and Betty Moore Foundation, the Beckman Institute, and the Sanofi-Aventis Bioengineering Research Program for their support of the Molecular Observatory at the California Institute of Technology (Caltech). The operations at the SSRL, ALS, and APS are supported by the U.S. Department of Energy and the National Institutes of Health (NIH). GM/CA has been funded in whole or in part with federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). D.H.L and A.M.D. were supported by a NIH Research Service Award (5 T32 GM07616). D.H.L. was also supported by an Amgen Graduate Fellowship through the Caltech-Amgen Research Collaboration. T.S. was supported by a Postdoctoral Fellowship of the Deutsche Forschungsgemeinschaft. F.M.H. was supported by a Ph.D. fellowship of the Boehringer Ingelheim Fonds. Y.F. was supported by a visiting Ph.D. student scholarship from the China Scholarship Council (CSC). A.H. was supported by Caltech startup funds, the Albert Wyrick V Scholar Award of the V Foundation for Cancer Research, the 54th Mallinckrodt Scholar Award of the Edward Mallinckrodt Jr. Foundation, a Kimmel Scholar Award of the Sidney Kimmel Foundation for Cancer Research, a Camille-

Dreyfus Teacher Scholar Award, NIH grant R01-GM111461, and is a Heritage Principal Investigator of the Heritage Medical Research Institute. The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 5HAX ($Nup170^{NTD}$•$Nup53^{R3}$), 5HAY ($Nup170^{CTD}$), 5HAZ ($Nup170^{CTD}$), 5HB0 ($Nup170^{CTD}$•$Nup145N^{R3}$), 5HB1 ($Nup170^{SOL}$), 5HB2 ($Nic96^{SOL}$), 5HB3 ($Nic96^{SOL}$•$Nup53^{R2}$), 5HB4 ($Nup192^{\Delta HEAD}$), 5HB5 ($Nup145N^{APD}$), 5HB6 ($Nup145N^{APD}$•$Nup145C^{N}$), 5HB7 ($Nup53^{RRM}$, space group $P2_12_12_1$), 5HB8 ($Nup53^{RRM}$, space group $P3_121$). A PyMol session file containing the composite structure of the NPC symmetric core can be obtained from our webpage (http://ahweb.caltech.edu). The authors declare no financial conflicts of interest. This work is dedicated to the memory of Mandy Hoelz.

Table S1.

Bacterial expression constructs and expression conditions

#	Protein	Residues	Expression vector	Restriction sites 5', 3'	N-terminal overhang	C-terminal overhang	Expression conditions
1	Nup82 NTD	1-595	pET28a-PreS	NdeI, NotI	GPH	none	18 °C / 18 hours
2	Nup159 T	1440-1481	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 18 hours
3	Nup120 Nup37 ELYS	1-1262 1-610 1-299	pET-MCN-SUMO	NdeI, BamHI NdeI, BamHI BamHI, NotI	GPHMHHHHHH Q S	none none none	37 °C / 2 hours
4	Nup85	230-1169	pET24a	NdeI, BamHI	MGHHHHHH	none	37 °C / 2 hours
5	Sec13	1-309	pETMCN	NdeI, BamHI	Q	none	23 °C / 18 hours; coexpressed with Nup145C
6	Nup145C	1-800	pET-MCN-SUMO	NcoI, NotI	S	none	23 °C / 18 hours; coexpressed with Sec13
7	Nup84	1-939	pET24a	NdeI, BamHI	MGHHHHHH	none	23 °C / 18 hours; coexpressed with Nup133
8	Nup133	1-1364	pET-MCN-SUMO	EcoRI, HindIII	SEF	none	23 °C / 18 hours; coexpressed with Nup84
9	Nup133 ΔNTE	105-1364	pET-MCN-SUMO	EcoRI, HindIII	SEF	none	23 °C / 18 hours; coexpressed with Nup84
10	Nup53	1-361	pET28a-SUMO	BamHI, XhoI	S	LEHHHHHH	23 °C / 18 hours
11♦	Nup53 RRM	133-253	pET28a-PreS	NdeI, XhoI	GPHM	none	23 °C / 18 hours
12	Nup53 R2	69-90	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
13♦	Nup53	31-84	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
14	Nup53 R1	31-67	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
15	Nup53	1-90	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
16	Nup53 R3	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
17	Nup53 R3 (F334A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
18	Nup53 R3 (I338A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
19	Nup53 R3 (R342A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
20	Nup53 R3 (K343A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
21	Nup53 R3 (L346A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
22	Nup53 R3 (L347A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
23	Nup53 R3 (E350A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
24	Nup53 R3 (E351A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
25	Nup53 R3 (L353A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
26	Nup53 R3 (L354A)	329-361	pGEX-6P-1	BamHI, NotI	not cleaved	none	37 °C / 2 hours
27	Nup170	74-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
28	Nup170 ΔC	74-1037	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
29♦	Nup170 SOL	575-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
30	Nup170 CTD	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
31♦	Nup170 CTD (ΔTLR)	851-1402 (Δ1375-1377)	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
32♦	Nup170 NTD	74-827 (Δ293-305)	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
33♦	Nup170 NTD	74-843 (Δ293-305)	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
34	Nup170 NTD (F199A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
35	Nup170 NTD (Y235A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
36	Nup170 NTD (V241A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
37	Nup170 NTD (W280A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
38♦	Nup170 CTD	851-1416	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
39	Nup170 NTD (F281A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
40	Nup170 NTD (R284A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
41	Nup170 NTD (I203A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
42	Nup170 NTD (M239A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
43	Nup170 NTD (L218A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
44	Nup170 NTD (E214A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
45	Nup170 NTD (E181A)	74-843	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
46	Nup170 CTD (Y1050A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
47	Nup170 CTD (V1062A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours

#	Protein	Residues	Expression vector	Restriction sites 5', 3'	N-terminal overhang	C-terminal overhang	Expression conditions
48	Nup170 CTD (L1066A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
49	Nup170 CTD (E1110A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
50	Nup170 CTD (L1111A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
51	Nup170 CTD (I1131A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
52	Nup170 CTD (I1147A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
53	Nup170 CTD (Y1164A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
54	Nup170 CTD (F1171A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
55	Nup170 CTD (F1154A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
56	Nup170 CTD (Y1157A)	851-1402	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
57♦	Nup170 CTD (Y905M/ L1007M/L1183M/V1292M)	832-1416	pET28a-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
58	Nup145N	606-993	pET28a-SUMO	BamHI, NotI	S	none	23 °C / 18 hours
59♦	Nup145N APD	858-993	pET28a-PreS	NdeI, XhoI	GPHM	none	23 °C / 18 hours
60♦	Nup145N APD (T994A) •Nup145C N	858-1000	pET28a-PreS	NdeI, XhoI	GPHM	none	23 °C / 18 hours
61	Nup145N R1	606-683	pET28a-SUMO	BamHI, NotI	S	AAALEHHH HHH	37 °C / 2 hours
62	Nup145N	606-750	pET28a-SUMO	BamHI, NotI	S	AAALEHHH HHH	37 °C / 2 hours
63	Nup145N	649-750	pET28a-SUMO	BamHI, NotI	S	AAALEHHH HHH	37 °C / 2 hours
64	Nup145N R3	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
65	Nup145N R3 (L733A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
66	Nup145N R3 (V734A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
67	Nup145N R3 (I735A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
68	Nup145N R3 (M739A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
69	Nup145N R3 (D742A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
70	Nup145N R3 (L743A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
71	Nup145N R3 (F744A)	729-750	pGEX-6P-1	BamHI, NotI	not cleaved	none	18 °C / 18 hours
72	Avi-Nup192	1-1756	pET28a-PreS	NdeI, NotI	GPLMSGLNDIF EAQKIEWHEGS AGGSGH	none	37 °C / 2 hours
73♦	Nup192 ΔHEAD	153-1756 (Δ167-184, replaced with GSGS)	pET28a-PreS	NdeI, NotI	GPHM	none	37 °C / 2 hours
74	Nup192 NTD	1-958	pET28a-PreS	NdeI, NotI	GPH	none	37 °C / 2 hours
75	Nup192 CTD	921-1756	pET28a-PreS	NdeI, NotI	GPHM	none	23 °C / 18 hours
76	Nup192 TAIL	1397-1756	pET28a-PreS	NdeI, NotI	GPHM	none	37 °C / 2 hours
77	Nic96	110-1112	pET28a-SUMO	BamHI, NotI	S	AAALEHHH HHH	18 °C / 18 hours
78♦	Nic96 SOL	391-1112	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 18 hours
79	Nic96 R2Δ2	274-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours coexpressed with Nup188 TAIL
80	Nic96 R2Δ4	286-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
81	Nic96 R2	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
82	Nic96 R2 (F275A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
83	Nic96 R2 (D276A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
84	Nic96 R2 (F278A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
85	Nic96 R2 (N282A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
86	Nic96 R2 (L285A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
87	Nic96 R2 (W287A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
88	Nic96 R2 (I294A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
89	Nic96 R2 (F298A)	262-301	pET-MCN-SUMO	BamHI, NotI	S	none	37 °C / 2 hours
90	Nic96 SOL (L715A)	391-1112	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 18 hours
91	Nic96 SOL (V655A/W658A)	391-1112	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 18 hours
92	Nic96 SOL (F697A)	391-1112	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 18 hours
93	Nup188	1-1858	pET28a-PreS	AseI, BamHI	GPHN	none	23 °C / 18 hours
94	Nup188 NTD	1-1134	pET28a-PreS	AseI, BamHI	GPHN	none	23 °C / 18 hours
95	Nup188 TAIL	1447-1858	pET28a-SUMO	BamHI, NotI	SM	none	18 °C / 18 hours

#	Protein	Residues	Expression vector	Restriction sites 5', 3'	N-terminal overhang	C-terminal overhang	Expression conditions
96	Nsp1 Nup57	467-674 74-319	pETDuet1	NcoI, NotI NdeI, XhoI	None M	none none	18 °C / 18 hours coexpressed with Nup49
97	Nup49	246-470	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 18 hours coexpressed with Nsp1 and Nup57
98	*hs*Nup155 CTD	870-1391	pET28a-PreS	NdeI, NotI	GPHM	none	37 °C / 2 hours
99	*hs*Nup98	596-617	pGEX-6P-1	BamHI, NotI	not cleaved	none	23 °C / 18 hours
100	*hs*Nup98 (S608E/S612E)	596-617	pGEX-6P-1	BamHI, NotI	not cleaved	none	23 °C / 18 hours
101	*sc*Nup159 T *sc*Nup82 NTD	1425-1460 1-452	pETDuet1	NcoI, NotI NdeI, XhoI	GPHM none	none none	18 °C / 16 hours
102	*sc*Nup120	1-1037	pET8c	NdeI, NotI	MGSSHHHHHH SD	none	18 °C / 16 hours
103	*sc*Nup85	44-744	pETDuet1	NdeI, NotI	MGSSHHHHHH SQDP	none	18 °C / 16 hours
	*sc*Seh1	1-349		NheI, XhoI	N/A	none	
104	*sc*Nup145C	1-712	pETDuet1	BamHI, NotI	MGSSHHHHHH SQDP	none	18 °C / 16 hours
	*sc*Sec13	1-297		NdeI, XhoI	N/A	none	
105	*sc*Nup145C ΔN *sc*Sec13	75-712 1-297	pETDuet1	BamHI, NotI NdeI, XhoI	MGSSHHHHHH SQDP N/A	none none	18 °C / 16 hours
106	*sc*Nup84 NTD	1-451	pET28a-PreS	NdeI, NotI	GPH	none	18 °C / 16 hours
107	*sc*Nup53	374-460	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
108	*sc*Nup170 CTD	999-1502	pET28a-SUMO	SacI, NotI	SEFEL	none	18 °C / 8 hours
109	*sc*Nup170 CTD (F1308A)	999-1502	pET28a-SUMO	SacI, NotI	SEFEL	none	18 °C / 8 hours
110	*sc*Nup157 CTD	894-1391	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
111	*sc*Nup157 NTD	70-893	pGEX-6P-1	BamHI, NotI	GPLGS	none	18 °C / 8 hours
112	*sc*Nup157 NTD (F214A)	70-893	pGEX-6P-1	BamHI, NotI	GPLGS	none	18 °C / 8 hours
113	*sc*Nup100	571-959	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
114	*sc*Nup116	752-1113	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
115	*sc*Nup145N	210-605	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
116	*sc*Nup192 NTD	1-960	pET28a-PreS	NdeI, NotI	GPH	none	21 °C / 8 hours
117	*sc*Nic96 SOL (I432A/W435A/Y469A)	204-839	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
118	*sc*Nup53	1-240	pET28a-SUMO	BamHI, NotI	S	none	18 °C / 16 hours
119	*sc*Nup188	1-1655	pET28a-SUMO	BamHI, NotI	S	none	21 °C / 16 hours
120	Nup145N PreS	606-993	pET28a-SUMO	BamHI, NotI	S	none	23 °C / 18 hours

Proteins are from *C. thermophilum* unless otherwise noted

♦ Constructs that were used for crystallization

Table S2.

Protein purification protocols

Protein(s)	Expression constructs	Purification step	Buffer A	Buffer B
Avi-Nup192*	Individual (#72)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. MonoQ 10/100 GL 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup192 ΔHEAD♦	Individual (#73)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. MonoQ 10/100 GL 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup192 NTD	Individual (#74)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup192 CTD	Individual (#75)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. MonoQ 10/100 GL 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup192 TAIL	Individual (#76)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A3, 5 mM β-ME / PreS 3. Ni-A3, 5 mM β-ME 4. SEC-C, 1 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. N/A
Nup188*	Individual (#93)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG 6. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT 6. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A 6. N/A
Nup188 NTD	Individual (#94)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup188 TAIL	Individual (#95)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM β-ME / ULP1 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nic96 SOL♦ wild-type and mutants	Individual (#78, 82-92)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. Ni-NTA 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 5 mM β-ME / ULP1 3. Ni-A2, 5 mM β-ME 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. N/A
Nup170*	Individual (#27)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME, 5 % Glycerol 2. Ni-A3, 5 mM β-ME, 5 % Glycerol / ULP1 3. Ni-A3, 5 mM β-ME, 5 % Glycerol 4. IEX-A4, pH 8.0, 5 mM DTT, 5 % Glycerol 5. SEC-C, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 5 mM β-ME, 5 % Glycerol 2. N/A 3. Ni-B1, 5 mM β-ME, 5 % Glycerol 4. IEX-B, pH 8.0, 5 mM DTT, 5 % Glycerol 5. N/A
Nup170 SOL♦	Individual (#29)	1. Ni-NTA 2. Dialysis/Cleavage (48 hr) 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME, 2. Ni-A3, 5 mM β-ME / ULP1 3. Ni-A3, 5 mM β-ME 4. IEX-A4, pH 8.0, 5 mM DTT 5. SEC-C, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup170 NTD♦ wild-type and mutants	Individual (#32-47, 39-45)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME, 5 % Glycerol 2. Ni-A3, 5 mM β-ME, 5 % Glycerol / ULP1 3. Ni-A3, 5 mM β-ME, 5 % Glycerol 4. IEX-A2, pH 8.0, 5 mM DTT, 5 % Glycerol 5. SEC-C, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 5 mM β-ME, 5 % Glycerol 2. N/A 3. Ni-B1, 5 mM β-ME, 5 % Glycerol 4. IEX-B, pH 8.0, 5 mM DTT, 5 % Glycerol 5. N/A
Nup170 CTD♦ wild-type and mutants	Individual (#30-31, 38, 46-57)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM β-ME / ULP1 3. Ni-A2, 5 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
***hs*Nup155 CTD**	Individual (#98)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME, 5 % Glycerol 2. Ni-A3, 5 mM β-ME, 5 % Glycerol / PreS 3. Ni-A3, 5 mM β-ME, 5 % Glycerol 4. IEX-A3, pH 8.0, 5 mM DTT, 5 % Glycerol 5. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 5 mM β-ME, 5 % Glycerol 2. N/A 3. Ni-B1, 5 mM β-ME, 5 % Glycerol 4. IEX-B, pH 8.0, 5 mM DTT, 5 % Glycerol 5. N/A

Protein(s)	Expression constructs	Purification step	Buffer A	Buffer B
Nup120•Nup37•ELYS*	Co-expression (#3)	1. Ni-NTA 2. Dialysis/Cleavage 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME, 5 % Glycerol 2. Ni-A3, 5 mM DTT, 5 % Glycerol / ULP1-PreS 3. IEX-A4, pH 8.0, 5 mM DTT, 5 % Glycerol 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME, 5 % Glycerol 2. N/A 3. IEX-B, pH 8.0, 5 mM DTT, 5 % Glycerol 4. N/A
Nup85*	Individual (#4)	1. Ni-NTA 2. Dialysis/Cleavage 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM DTT / PreS 3. IEX-A2, pH 8.0, 5 mM DTT 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME, 5 % Glycerol 2. N/A 3. IEX-B, pH 8.0, 5 mM DTT 4. N/A
Sec13•Nup145C*	Co-expression (#5, 6)	1. Ni-NTA 2. HiPrep 26/20 Desalting /Cleavage 3. Ni-NTA 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM β-ME / ULP1 3. Ni-A2, 5 mM β-ME 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B2, 5 mM β-ME 4. N/A
Nup84•Nup133*	Co-expression (#7, 8)	1. Ni-NTA 2. Dialysis 3. HiLoad Superdex 200 16/60 PG 4. MonoQ 10/100 GL	1. Ni-A1, 5 mM β-ME 2. SEC-A, 5 mM DTT 3. SEC-A, 5 mM DTT 4. IEX-A2, pH 8.0, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. N/A 4. IEX-B, pH 8.0, 5 mM DTT
Nup84•Nup133ΔNTE	Co-expression (#7, 9)	1. Ni-NTA 2. Dialysis 3. HiLoad Superdex 200 16/60 PG 4. MonoQ 10/100 GL	1. Ni-A1, 5 mM β-ME 2. SEC-A, 5 mM DTT 3. SEC-A, 5 mM DTT 4. IEX-A2, pH 8.0, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. N/A 4. IEX-B, pH 8.0, 5 mM DTT
Nup145N* wild-type and PreS	Individual (#58, 120)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. MonoS 5/50 GL 4. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 7.0, 5 mM DTT / ULP1 3. IEX-A2, pH 7.0, 5 mM DTT 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. IEX-B, pH 7.0, 5 mM DTT 4. N/A
Nup145N APD♦	Individual (#59)	1. Ni-NTA 2. Dialysis/cleavage 3. Ni-NTA 4. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, pH 8.0, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. N/A
GST-Nup145N wild-type and mutants	Individual (#64, 65-71)	1. GST Affinity 2. HiTrap Q HP 3. HiLoad Superdex 200 16/60 PG	1. GST-A, 5 mM DTT 2. IEX-A2, pH 8.0, 5 mM DTT 3. SEC-A, 5 mM DTT	1. GST-B, 5mM DTT 2. IEX-B, pH 8.0, 5 mM DTT 3. N/A
GST-*hs*Nup98 wild-type and mutant	Individual (#99, 100)	1. GST Affinity 2. HiTrap Q HP 3. HiLoad Superdex 200 16/60 PG	1. GST-A, 5 mM DTT 2. IEX-A2, pH 8.0, 5 mM DTT 3. SEC-A, 5 mM DTT	1. GST-B, 5mM DTT 2. IEX-B, pH 8.0, 5 mM DTT 3. N/A
SUMO-Nup145N various fragments	Individual (#61-63)	1. Ni-NTA 2. HiPrep 26/20 Desalting /Cleavage 3. Ni-NTA 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, 5 mM β-ME / ULP1 3. Ni-A2, 5 mM β-ME 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM **β**-ME 2. N/A 3. Ni-B2, 5 mM β-ME 4. N/A
Nup53*	Individual (#10)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. Ni-NTA 4. HiPrep 26/20 Desalting 5. MonoS 5/50 GL 6. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A1, 5 mM β-ME / ULP1 3. Ni-A1, 5 mM β-ME 4. IEX-A2, pH 7.0, 5 mM DTT 5. IEX-A2, pH 7.0, 5 mM DTT 6. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. N/A 5. EX-B, pH 7.0, 5 mM DTT 6. N/A
Nup53 RRM♦	Individual (#11)	1. Ni-NTA 2. Desalting/Cleavage 3. Ni-NTA 4. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A2, pH 8.0, 5 mM β-ME / PreS 3. Ni-A2, 5 mM β-ME 4. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. Ni-B1, 5 mM β-ME 4. N/A
SUMO-Nup53 various fragments	Individual (#12-15)	1. Ni-NTA 2. HiPrep 26/20 Desalting 3. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. SEC-A, 5 mM DTT 3. SEC-A, 5 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. N/A
GST-Nup53 R3 wild-type and mutants	Individual (#16, 17-26)	1. GST Affinity 2. HiTrap Q HP 3. HiLoad Superdex 200 16/60 PG	1. GST-A, 5 mM DTT 2. IEX-A2, pH 8.0, 5 mM DTT 3. SEC-A, 5 mM DTT	1. GST-B, 5mM DTT 2. IEX-B, pH 8.0, 5 mM DTT 3. N/A
CNT	Co-expression (#96, 97)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 4 mM β-ME 2. Ni-A2, 4 mM β-ME / ULP1 3. Ni-A2, 4 mM β-ME 4. IEX-A2, pH 8.0, 5 mM DTT 5. SEC-A, 5 mM DTT	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 5 mM DTT 5. N/A
Nup82 NTD•SUMO-Nup159T	Individual (#1, 2) co-lysis ratio 1:1	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. Ni-A1, 5 mM β-ME/ PreS / ULP1 3. Ni-A1, 5 mM β-ME 4. SEC-A, 5 mM DTT	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. N/A
His_6-Nup192 TAIL	Individual (#76)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A1, pH 9.0, 1 mM DTT 3. IEX-A1, pH 9.0, 1 mM DTT 4. SEC-C, 1 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. IEX-B, pH 9.0, 1 mM DTT 4. N/A
SUMO-Nup188 TAIL	Individual (#95)	1. Ni-NTA 2. Dialysis 3. MonoQ 10/100 GL 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 1 mM DTT 3. IEX-A2, pH 8.0, 1 mM DTT 4. SEC-C, 1 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 1 mM DTT 4. N/A
SUMO-Nup188 TAIL•SUMO-Nic96-R2Δ2	Co-expression (#95, 79)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 1 mM DTT 3. IEX-A2, pH 8.0, 1 mM DTT 4. SEC-C, 1 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 1 mM DTT 4. N/A

Protein(s)	Expression constructs	Purification step	Buffer A	Buffer B
SUMO-Nic96 R2	Individual (#81)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 1 mM DTT 3. IEX-A2, pH 8.0, 1 mM DTT 4. SEC-C, 1 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 1 mM DTT 4. N/A
SUMO-Nic96 R2Δ4	Individual (#80)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 5 mM β-ME 2. IEX-A2, pH 8.0, 5 mM β-ME 3. IEX-A2, pH 8.0, 1 mM DTT 4. SEC-C, 1 mM DTT	1. Ni-B1, 5 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 1 mM DTT 4. N/A
SUMO-*sc*Nup53	Individual (#107)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol 3. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol 4. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 4 mM β-ME 4. N/A
***sc*Nup100**	Individual (#113)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A3, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A3, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
SUMO-*sc*Nup100	Individual (#113)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol 3. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol 4. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 4 mM β-ME 4. N/A
***sc*Nup116**	Individual (#114)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A3, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A3, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
***sc*Nup145N**	Individual (#115)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A3, pH 8.0, 4 mM **β**-ME 5 % Glycerol 4. IEX-A3, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
SUMO-*sc*Nup145N	Individual (#115)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol 3. IEX-A3, pH 8.0, 4 mM β-ME, 5 % Glycerol 4. SEC-B, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 4 mM β-ME 4. N/A
***sc*Nup188**	Individual (#119)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A2, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-A, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
***sc*Nup192 NTD**	Individual (#116)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A2, pH 8.0, 4 mM β-ME, 5 % Glycerol / PreS 3. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-A, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
***sc*Nup170 CTD** wild-type and mutant	Individual (#108, 109)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A4, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A4, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A4, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-C, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
SUMO-*sc*Nup170 CTD	Individual (#108)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A4, pH 8.0, 4 mM β-ME, 5 % Glycerol 3. IEX-A4, pH 8.0, 4 mM β-ME 5 % Glycerol 4. SEC-C, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 4 mM β-ME 4. N/A
***sc*Nup157 CTD**	Individual (#110)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A4, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A4, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A4, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-C, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM **β**-ME 5. N/A
SUMO-*sc*Nup157 CTD	Individual (#110)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A4, pH 8.0, 4 mM β-ME, 5 % Glycerol 3. IEX-A4, pH 8.0, 4 mM β-ME 5 % Glycerol 4. SEC-C, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 4 mM β-ME 4. N/A
***sc*Nup82 NTD•*sc*Nup159 T**	Co-expression (#101)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A2, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-A, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
***sc*Sec13•*sc*Nup145C**	Co-expression (#105)	1. Ni-NTA 2. Dialysis/Cleavage 3. Ni-NTA 4. HiTrap Q HP 5. HiLoad Superdex 200 16/60 PG	1. Ni-A4, 4 mM β-ME, 5 % Glycerol 2. IEX-A2, pH 8.0, 4 mM β-ME, 5 % Glycerol / ULP1 3. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 4. IEX-A2, pH 8.0, 4 mM β-ME 5 % Glycerol 5. SEC-A, 5 mM DTT, 5 % Glycerol	1. Ni-B1, 4 mM β-ME 2. N/A 3. Ni-B1, 4 mM β-ME 4. IEX-B, pH 8.0, 4 mM β-ME 5. N/A
***sc*Nic96 SOL** triple mutant	Individual (#117)	1. Ni-NTA 2. HiPrep 26/20 Desalting/Cleavage 3. HiLoad Superdex 200 16/60 PG	1. Ni-A1, 4 mM β-ME 2. SEC-A, 5 mM DTT / ULP1 3. SEC-A, 5 mM DTT	1. Ni-B1, 4 mM β-ME 2. N/A 3. N/A

Protein(s)	Expression constructs	Purification step	Buffer A	Buffer B
SUMO-*sc*Nup53	Individual (#118)	1. Ni-NTA 2. Dialysis 3. HiTrap Q HP 4. HiLoad Superdex 75 16/60 PG	1. Ni-A1, 4 mM β-ME 2. IEX-A1, pH 8.0, 5 mM DTT 3. IEX-A1, pH 8.0, 5 mM DTT 4. SEC-A, 5 mM DTT	1. Ni-B1, 4 mM β-ME 2. N/A 3. IEX-B, pH 8.0, 5 mM DTT 4. N/A
scNup157 NTD wild-type and mutant	Individual (#111-112)	Detailed description in SI methods		
Nic96•CNT	Individual (#77, 96, 97) co-lysis	Detailed description in SI methods		

♦ Constructs that were used for crystallization

***** Detailed purification provided in supplementary figures

Ni-A1: 20 mM TRIS (pH 8.0), 500 mM NaCl, 20 mM imidazole
Ni-A2: 20 mM TRIS (pH 8.0), 100 mM NaCl, 20 mM imidazole
Ni-A3: 20 mM TRIS (pH 8.0), 200 mM NaCl, 20 mM imidazole
Ni-A4: 75 mM TRIS (pH 8.0), 500 mM NaCl, 20 mM imidazole
Ni-B1: 20 mM TRIS (pH 8.0), 500 mM NaCl, 500 mM imidazole
Ni-B2: 20 mM TRIS (pH 8.0), 100 mM NaCl, 500 mM imidazole

GST-A: 20 mM TRIS (pH 8.0), 200 mM NaCl
GST-B: 20 mM TRIS (pH 8.0), 200 mM NaCl, 20 mM glutathione

IEX-A1: 20 mM TRIS, 50 mM NaCl
IEX-A2: 20 mM TRIS, 100 mM NaCl
IEX-A3: 20 mM TRIS, 150 mM NaCl
IEX-A4: 20 mM TRIS, 200 mM NaCl
IEX-B: 20 mM TRIS, 2.0 M NaCl

SEC-A: 20 mM TRIS (pH 8.0), 100 mM NaCl
SEC-B: 20 mM TRIS (pH 8.0), 150 mM NaCl
SEC-C: 20 mM TRIS (pH 8.0), 200 mM NaCl

Table S3.

SEC-MALS analysis

Figure	Nucleoporin or nucleoporin complex	Experimental mass (kDa)	Theoretical mass (kDa)	Stoichiometry
fig. S3A	Nup192 Nup145N Nup53	250	278	
	Nic96 CNT (Nic96 Nup57 Nup49 Nsp1)	160	187	
	IRC Nup53 (Nup192 Nic96 Nup145N Nup53 CNT)	380	465	conc. dependent
fig. S3B	IRC Nup53 (Nup192 Nic96 Nup145N Nup53 CNT)	335	423	
	Nup188	180	204	
	IRC Nup53 + Nup188	341 / 220	627	no interaction
Fig. 1D	CNC-hexamer Nup145N			
fig. S4A	(Nup120 Nup37 ELYS Nup85 Sec13 Nup145C Nup145N)	482	507	
	Nup192 Nic96 Nup53 CNT	335	423	
	CNC-hexamer IRC Nup53	654	930	conc. dependent
Fig. 1E	CNC-hexamer			
fig. S4B	(Nup120 Nup37 ELYS Nup85 Sec13 Nup145C)	438	465	
	Nup192 Nic96 Nup53 CNT	335	423	
	CNC-hexamer + IRC Nup53	430	888	no interaction
Fig. 1F	CNC-hexamer Nup145N	466	507	
fig. S4C	Nup188 Nic96 Nup53 CNT	523	431	
	CNC-hexamer Nup188 Nic96 Nup145N Nup53 CNT	1220	938	superstoichiometric
Fig. 1G	CNC-hexamer	428	465	
fig. S4D	Nup188 Nic96 Nup53 CNT	523	431	
	CNC-hexamer + Nup188 Nic96 Nup53 CNT	507	896	no interaction
fig. S5A	CNC-hexamer	428	465	
	Nup145N	40	42	
	CNC-hexamer Nup145N	508	507	stoichiometric
fig. S5B	CNC-hexamer	428	465	
	Nup145N APD	17	15	
	CNC-hexamer Nup145N APD	447	480	stoichiometric
fig. S5C	CNC-hexamer	428	465	
	Nup145N PreS	36 / 15	27 / 15	
	CNC-hexamer Nup145N APD + Nup145N MID	690	480	superstoichiometric
fig. S6A	CNC-hexamer	428	465	
	Nup82 NTD Nup159 T Nup145N	112	122	
	CNC-hexamer Nup145N + Nup82 NTD Nup159 T	579	587	stoichiometric
fig. S6B	Nup188 Nic96 Nup53 CNT	523	431	
	Nup82 NTD Nup159 T Nup145N	126	122	
	Nup188 Nic96 Nup145N Nup53 CNT			
	+ Nup82 NTD Nup159 T	670	553	superstoichiometric
Fig. 2A	Nup192	176	196	
fig. S7A	Nup188	180	204	
	Nup170	132	147	
	Nic96 SOL	73	81	
	Nup192 + Nup188 + Nup170 + Nic96 SOL	174	628	no interaction
Fig. 2B	Nup192 + Nup170 + Nic96 SOL	161 / 93	172[a] / 81	
fig. S7B	Nup53 + Nup145N	52	41[a]	
	Nup192 Nup170 Nic96 SOL Nup53 Nup145N	528	506	stoichiometric
Fig. 2C	Nup188 + Nup170 + Nic96 SOL	165 / 87	176[a] / 81	
fig. S7C	Nup53 + Nup145N	52	41[a]	
	Nup188 + Nup170 + Nic96 SOL + Nup53 + Nup145N	242	512	weak interaction
Fig. 2D	Nic96 SOL	72	81	
fig. S8A	Nup53	40	40	
	Nic96 SOL Nup53	146	121	superstoichiometric
Fig. 2D	Nup170	145	147	
fig. S8B	Nup53	35	40	
	Nup170 Nup53	197	187	stoichiometric
Fig. 2D	Nup188	188	204	
fig. S8C	Nup53	40	40	
	Nup188 + Nup53	188	244	weak interaction
Fig. 2D	Nup192	178	196	
fig. S8D	Nup53	40	40	
	Nup192 Nup53	223	236	stoichiometric
Fig. 2D	CNT (Nup57, Nup49, Nsp1)	70	77	
fig. S8E	Nup53	38	40	
	CNT + Nup53	73	117	weak interaction
Fig. 2E	Nup170	145	147	
fig. S9A	Nup145N	42	42	
	Nup170 Nup145N	196	189	stoichiometric

Figure	Nucleoporin or nucleoporin complex	Experimental mass (kDa)	Theoretical mass (kDa)	Stoichiometry
Fig. 2E	Nup188	189	204	
fig. S9B	Nup145N	40	42	
	Nup188 Nup145N	216	246	stoichiometric
Fig. 2E	Nup192	178	196	
fig. S9C	Nup145N	40	42	
	Nup192 Nup145N	265	238	superstoichiometric
Fig. 2D	Nup170 Nup53	154	187	
fig. S10A	Nic96 SOL	70	81	
	Nup170 Nup53 Nic96 SOL	208	268	stoichiometric
Fig. 2D	Nup170 Nup53	154	187	
fig. S10B	Nup188	194	204	
	Nup170 Nup53 + Nup188	188	391	no interaction
Fig. 2D	Nup170 Nup53	154	187	
fig. S10C	Nup192	196	196	
	Nup170 Nup53 Nup192	259	383	conc. dependent
Fig. 2D	Nup192 Nup53	215	236	
fig. S10D	Nic96 SOL	70	81	
	Nup192 Nup53 Nic96 SOL	252	317	conc. dependent
fig. S10E	Nup170 Nup53 Nup192	255	383	
	Nic96 SOL	70	81	
	Nup170 Nup53 Nup192 Nic96 SOL	286	464	conc. dependent
Fig. 2E	Nup170 Nup145N	178	189	
fig. S11A	Nup188	184	204	
	Nup170 Nup145N + Nup188 Nup145N	227	393	partial exchange
Fig. 2E	Nup170 Nup145N	178	189	
fig. S11B	Nup192	178	196	
	Nup170 Nup145N Nup192	305	385	conc. dependent
Fig. 2E	Nup188 Nup145N	203	246	
fig. S11C	Nup170	137	147	
	Nup188 Nup145N + Nup170 Nup145N	193	393	partial exchange
Fig. 2E	Nup192 Nup145N	227	238	
fig. S11D	Nup170	137	147	
	Nup192 Nup145N Nup170	276	385	conc. dependent
Fig. 2E	Nup192 Nup145N	233	238	
fig. S11E	Nup188	190	204	
	Nup192 Nup145N + Nup188 Nup145N	223	442	partial exchange
Fig. 2E	Nup188 Nup145N	215	246	
fig. S11F	Nup192	180	196	
	Nup188 Nup145N + Nup192 Nup145N	224	442	partial exchange
Fig. 2E	Avi-Nup192	178	194	
fig. S11G	Nup188	178	204	
	Nup145N	52	42	
	Avi-Nup192 Nup145N + Nup188 Nup145N	223	440	partial exchange
fig. S12A	Nup170 Nup145N	178	189	
	Nup53	38	40	
	Nup170 Nup145N Nup53	*inconclusive	229	stoichiometric
fig. S12B	Nup192	179	196	
	Nup53	38	40	
	Nup145N	40	42	
	Nup192 Nup53 Nup145N	280	278	stoichiometric
fig. S12C	Nup170 Nup53 Nup145N	183	229	
	Nup192	176	196	
	Nup170 Nup53 Nup145N Nup192	416	425	stoichiometric
Fig. 2F	Nup192	183	196	
fig. S13A	SUMO-Nup53 31-67	16	14	
	Nup192 SUMO-Nup53 31-67	197	210	stoichiometric
Fig. 2F	Nic96 SOL	70	81	
fig. S14A	SUMO-Nup53 1-90	21	21	
	Nic96 SOL SUMO-Nup53 1-90	88	102	stoichiometric
Fig. 2F	Nic96 SOL	81	81	
fig. S14B	SUMO-Nup53 69-90	16	14	
	Nic96 SOL SUMO-Nup53 69-90	88	95	stoichiometric
Fig. 2F	Nup192	176	196	
fig. S14C	Nic96 SOL SUMO-Nup53 1-90	81	102	
	Nup192 Nic96 SOL SUMO-Nup53 1-90	220	298	conc. dependent
Fig. 2F	Nup192	178	196	
fig. S17A	SUMO-Nup145N 606-683	20	19	
	Nup192 SUMO-Nup145N 606-683	189	215	stoichiometric

Figure	Nucleoporin or nucleoporin complex	Experimental mass (kDa)	Theoretical mass (kDa)	Stoichiometry
Fig. 2F	Nup192	176	196	
fig. S17B	Nup145N PreS	28 / 14	27 / 15	
	Nup192 Nup145N MID + Nup145N APD	223	223	stoichiometric
Fig. 2F	Nup192 NTD	100	109	
fig. S17C	Nup145N	53	42	
	Nup192 NTD Nup145N	151	151	stoichiometric
Fig. 2F	Nup192 CTD	127	91	
fig. S17D	Nup145N	40	42	
	Nup192 CTD + Nup145N	137	133	stoichiometric
Fig. 2F	Nup192 TAIL	38	41	
fig. S17E	Nup145N	46	42	
	Nup192 TAIL + Nup145N	45 / 40	83	stoichiometric
fig. S18A	Nup170	125	147	
	SUMO-Nup53 31-67	16	14	
	Nup170 + SUMO-Nup53 31-67	129	161	no interaction
fig. S18B	Nup170	125	147	
	SUMO-Nup53 69-90	14	14	
	Nup170 + SUMO-Nup53 69-90	129	161	no interaction
fig. S18C	Nup170	125	147	
	SUMO-Nup145N 606-683	15	19	
	Nup170+ SUMO-Nup145N 606-683	133	166	no interaction
fig. S18D	Nup192	176	196	
	SUMO-Nup53 69-90	14	14	
	Nup192 + SUMO-Nup53 69-90	174	210	no interaction
fig. S18E	Nup192	176	196	
	GST-Nup53 329-361	55	61	
	Nup192 + GST-Nup53 329-361	172	257	no interaction
fig. S18F	Nup192	176	196	
	GST-Nup145N 729-750	55	59	
	Nup192 + GST-Nup145N 729-750	172	255	no interaction
fig. S18G	Nic96 SOL	70	81	
	SUMO-Nup53 31-67	16	14	
	Nic96 + SUMO-Nup53 31-67	74	95	no interaction
fig. S18H	Nic96 SOL	70	81	
	GST-Nup53 329-361	55	61	
	Nic96 + GST-Nup53 329-361	66	142	no interaction
Fig. 2F	Nup188	181	204	
fig. S19B	Nup145N PreS	28 / 14	27 / 15	
	Nup188 Nup145N MID + Nup145N APD	213	231	stoichiometric
Fig. 2F	Nup188	178	204	
fig. S19C	SUMO-Nup145N 606-750	37	27	
	Nup188 SUMO-Nup145N 606-750	202	231	stoichiometric
Fig. 2F	Nup188	179	204	
fig. S19D	SUMO-Nup145N 606-683	21	19	
	Nup188 + SUMO-Nup145N 606-683	173	223	no interaction
Fig. 2F	Nup188	178	204	
fig. S19E	GST-Nup145N 729-750	55	59	
	Nup188 + GST-Nup145N 729-750	188	263	weak interaction
fig. S39A	Nic96 SOL	72	81	
	Nup53 RRM	13	13	
	Nic96 SOL + Nup53 RRM	73	94	no interaction
fig. S39B	Nup170	133	147	
	Nup53 RRM	12	13	
	Nup170 + Nup53 RRM	132	160	no interaction
fig. S39C	Nup192	183	196	
	Nup53 RRM	14	13	
	Nup192 + Nup53 RRM	182	209	no interaction
fig. S39D	Nup188	183	204	
	Nup53 RRM	14	13	
	Nup188 + Nup53 RRM	182	209	no interaction
fig. S44B	Nup120 Nup37 ELYS	209	239	
	Nup84 Nup133	241	258	
	Nup120 Nup37 ELYS Nup84 Nup133	264	497	weak interaction
fig. S44C	Nup120 Nup37 ELYS	209	239	
	Nup84 Nup133 NTE	233	246	
	Nup120 Nup37 ELYS +Nup84 Nup133 NTE	240	485	no interaction

Table S4.

Crystallization and cryoprotection conditions

Protein(s)	Concentration	Crystallization condition	Cryo protection condition
$Nup170^{NTD}$• $Nup53^{R3}$	20 mg/ml $Nup170^{NTD}$ 3-fold molar excess $Nup53^{R3}$	1.0 M Na/K phosphate (pH 6.9)	Mother liquor gradually supplemented to 25 % (v/v) ethylene glycol in 5 % steps
$Nup170^{CTD}$• $Nup145N^{R3}$	20 mg/ml $Nup170^{CTD}$ 3-fold molar excess $Nup145N^{R3}$	12 % (w/v) PEG 3,350 0.2 M lithium acetate	Mother liquor gradually supplemented to 25 % (v/v) ethylene glycol in 5 % steps
$Nup170^{CTD}$ SeMet	15 mg/ml	0.1 M MES (pH 6.3) 10 % PEG 20,000 10 % (v/v) ethylene glycol 0.2 M potassium thiocyanate	Mother liquor gradually supplemented to 25 % (v/v) ethylene glycol in 5 % steps
$Nup170^{CTD}$	30 mg/ml	0.1 M HEPES (pH 7.0) 1.0 M sodium acetate	Mother liquor gradually supplemented to 25 % (v/v) ethylene glycol in 5 % steps
$Nup170^{SOL}$	7 mg/ml	0.1 M HEPES (pH 7.1) 0.5 M ammonium sulfate	Mother liquor gradually supplemented to 25 % (v/v) ethylene glycol in 5 % steps
$Nic96^{SOL}$• $Nup53^{R2}$	12 mg/ml	0.1 M TRIS (pH 7.0) 12 % (w/v) PEG 6,000 1.2 M sodium chloride	Paratone N
$Nic96^{SOL}$	12 mg/ml	4 % Tacsimate (pH 7.4) 14 % (w/v) PEG 3,350	Paratone N
$Nup192^{\Delta HEAD}$	7 mg/ml	0.1 M TRIS (pH 7.9) 6 % (w/v) PEG 4,000 5 % (v/v) polypropylene glycol	Stabilized in 0.1 M TRIS (pH 7.9) 7 % (w/v) PEG 4,000 6 % (v/v) polypropylene glycol Solution gradually supplemented to 25% (v/v) polypropylene glycol in 5 % steps
$Nup53^{RRM}$	15 mg/ml	0.1 M HEPES (pH 7.0) 24 % (w/v) PEG 3,350 0.2 M potassium iodide	Mother liquor supplemented to 25 % ethylene glycol
$Nup53^{RRM}$	15 mg/ml	4 % Tacsimate (pH 4.1) 15 % (w/v) PEG 3,350	32 % Tacsimate (pH 4.1) 19 % (w/v) PEG 3,350
$Nup145N^{APD}$	10 mg/ml	0.1 M citric acid (pH 3.0) 25 % (w/v) PEG 3,350	Mother liquor supplemented to 25 % ethylene glycol
$Nup145N^{APD}$ •$Nup145C^{N}$	10 mg/ml	0.1 M TRIS (pH 8.8) 32 % (w/v) PEG 4,000 0.2 M lithium sulfate	Mother liquor supplemented to 10 % glycerol

Table S5.

X-ray crystallography analysis of $Nup170^{NTD}$•$Nup53^{R3}$ and $Nup170^{SOL}$

Data collection					
Protein	$Nup170^{NTD}$• $Nup53^{R3}$	$Nup170^{NTD}$• $Nup53^{R3}$	$Nup170^{SOL}$	$Nup170^{SOL}$ [g]	$Nup170^{SOL}$ [g]
PDB ID		5HAX		5HB1	
Synchrotron	APS[a]	APS	SSRL[b]	SSRL	SSRL
Beamline	23-ID-D	23-ID-D	BL12-2	BL12-2	BL12-2
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions					
a, *b*, *c* (Å)	68.7, 104.0, 119.0	69.2, 106.2, 120.1	93.2, 136.9, 89.0	96.0, 138.5, 89.2	96.0, 138.5, 89.2
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
	Se Peak		*Se Peak*		
Wavelength	0.9792	1.0332	0.9791	1.0330	
Resolution (Å)	50.0 – 2.5	50.0 – 2.1	50.0 – 4.5	40.0 – 4.0	40.0 – 4.3
R_{meas} (%)[c]	8.5 (127.7)	8.0 (188.5)	13.1 (290.0)	5.7 (353.7)	5.1 (121.5)
R_{pim} (%)[c]	1.7 (24.9)	2.2 (51.1)	3.7 (85.0)	1.7 (98.5)	1.5 (33.0)
$CC_{1/2}$[c]	100.0 (86.6)	100.0 (68.1)	99.9 (56.0)	100.0 (55.5)	100.0 (92.2)
$< I / \sigma I >$[c]	30.4 (2.4)	20.3 (1.8)	11.6 (1.0)	18.5 (0.8)	22.8 (2.5)
Completeness (%)[c]	98.9 (90.7)	100.0 (100.0)	99.3 (94.1)	99.7 (98.5)	99.8 (99.9)
No. of observations	776,896	676,933	91,230	135,558	108,324
No. of unique reflections[c,d]	30,229 (2,703)	52,358 (5,120)	7,200 (649)	10,614 (1,027)	8,524 (845)
Redundancy[c]	25.7 (25.2)	12.9 (13.5)	12.7 (11.1)	12.8 (12.4)	12.7 (13.4)
Refinement					
Resolution (Å)		50.0 – 2.1		40.0 – 4.0	
No. of reflections		52,342		7,867[e]	
No. of reflections test set		2,618 (5.0%)		789 (10.0%)	
R_{work} / R_{free}		19.9 / 23.0		30.4 / 35.2	
No. atoms		5,998		5,895	
Protein		5,581		5,895	
Water		387		0	
Ligand/Ions		30		0	
B-factors		62		222	
Protein		63		222	
Water		57		-	
Ligand/Ions		79		-	
RMSD					
Bond lengths (Å)		0.002		0.003	
Bond angles (°)		0.5		0.5	
Ramachandran plot[f]					
Favored (%)		97.6		96.8	
Additionally allowed (%)		2.4		3.2	
Outliers (%)		0.0		0.0	
MolProbity					
Clashscore[f]		0.18 (100^{th})		1.95 (100^{th})	
Molprobity score[f]		0.79 (100^{th})		1.46 (100^{th})	

[a]APS, Advanced Photon Source
[b]SSRL, Stanford Synchotron Radiation Lightsource
[c]Highest-resolution shell is shown in parentheses
[d]Friedel pairs were merged
[e]Refinement was performed with ellipsoidally truncated data
[f]As determined by MolProbity
[g]As a reference, two different high-resolution cutoffs are shown

Table S6.

X-ray crystallography analysis of *apo* Nup170^{CTD} and Nup170^{CTD}•Nup145N^{R3}

Data collection			
Protein	Nup170^{CTD} Y905M/L1007M /L1183M/V1292M	Nup170^{CTD}	Nup170^{CTD}•Nup145N^{R3}
PDB ID	5HAY	5HAZ	5HB0
Synchrotron	SSRL[a]	APS[b]	APS
Beamline	BL12-2	23-ID-D	23-ID-D
Space group	$P2_12_12_1$	$P2_12_12_1$	P1
Cell dimensions			
a, *b*, *c* (Å)	61.8, 101.7, 193.8	69.2, 106.2, 120.1	74.6, 111.47, 111.75
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	91.8, 92.5, 91.4
	Se Peak		
Wavelength	0.9795	1.0332	1.0332
Resolution (Å)	50.0 – 2.8	50.0 – 2.1	50.0 – 3.5
R_{meas} (%)[c]	11.6 (99.9)	5.4 (89.5)	41.4 (179.9)
R_{pim} (%)[c]	2.8 (25.2)	2.1 (33.8)	15.6 (67.6)
$CC_{1/2}$[c]	99.9 (81.8)	99.9 (85.0)	98.4 (51.7)
$< I / \sigma I >$[c]	22.4 (2.9)	21.0 (2.0)	6.2 (1.3)
Completeness (%)[c]	98.2 (82.2)	97.8 (93.9)	98.8 (94.9)
No. of observations	500,568	325,202	317,882
No. of unique reflections[c]	30,966 (2,527)	47,908 (4,589)	45,493 (4,380)
Redundancy[c]	16.2 (14.2)	6.8 (6.8)	7.0 (6.9)
Refinement			
Resolution (Å)	50.0 – 2.8	50.0 – 2.1	50.0 – 3.5
No. of reflections	30,817	47,882	45,466
No. of reflections test set	1,544 (5.0%)	2,395 (5.0%)	2,011 (4.4%)
R_{work} / R_{free}	20.9 / 25.1	20.6 / 23.1	21.0 / 26.0
No. atoms (non-hydrogen)	8,613	4,919	16,998
Protein	8,564	4,517	16,998
Water	48	378	0
Ligand/Ions	1	24	0
B-factors	65	69	89
Protein	65	69	89
Water	57	64	-
Ligand/Ions	62	117	-
RMSD			
Bond lengths (Å)	0.003	0.004	0.005
Bond angles (°)	0.6	0.6	0.6
Ramachandran plot[e]			
Favored (%)	97.2	97.5	95.1
Additionally allowed (%)	2.8	2.5	4.9
Outliers (%)	0.0	0.0	0.0
MolProbity			
Clashscore[e]	1.35 (100th)	1.44 (100th)	1.98 (100th)
Molprobity score[e]	1.29 (100th)	1.27 (100th)	1.55 (100th)

[a]SSRL, Stanford Synchotron Radiation Lightsource
[b]APS, Advanced Photon Source
[c]Highest-resolution shell is shown in parentheses
[d]Friedel pairs were merged
[e]As determined by MolProbity

Table S7.

X-ray crystallography analysis of *apo* Nic96^{SOL}, Nup96^{SOL}•Nup53^{R2} and Nup192$^{\Delta HEAD}$

Data collection				
Protein	Nic96^{SOL}	Nup96^{SOL}•Nup53^{R2}	Nup192$^{\Delta HEAD\ f}$	Nup192$^{\Delta HEAD\ f}$
PDB ID	5HB2	5HB3	5HB4	
Synchrotron	APS[a]	SSRL[b]	SSRL	SSRL
Beamline	23-ID-B	BL12-2	BL12-2	BL12-2
Space group	$P2_1$	P1	C2	C2
Cell dimensions				
a, *b*, *c* (Å)	53.0, 72.9, 122.7	59.5, 87.0, 98.1	190.7, 53.5, 171.8	190.7, 53.5, 171.8
α, β, γ (°)	90.0, 93.6, 90.0	100.7, 99.6, 95.7	90.0, 108,3, 90.0	90.0, 108,3, 90.0
	Se Peak		*Os Peak*	*Os Peak*
Wavelength	0.9792	1.0000	1.1250	1.1250
Resolution (Å)	50.0 – 3.3	50.0 – 2.65	50.0 – 3.2	50.0 – 3.5
R_{meas} (%)[c]	13.7 (104.2)	7.7 (83.7)	20.8 (325.0)	16.5 (151.7)
R_{pim} (%)[c]	5.0 (38.7)	3.9 (49.4)	5.7 (92.5)	4.5 (40.9)
$CC_{1/2}$[c]	99.8 (85.1)	99.9 (81.5)	99.9 (37.5)	99.9 (74.7)
$< I / \sigma I >$[c]	12.0 (2.5)	15.1 (2.1)	12.1 (0.8)	15.5 (2.1)
Completeness (%)[c]	100.0 (99.8)	88.6 (82.4)	98.6 (92.0)	99.4 (99.1)
No. of observations	107,569	189,415	367,182	280,883
No. of unique reflections[c,d]	14,334 (1,400)	48,566 (4,513)	27,800 (2,557)	21,186 (2,077)
Redundancy[c]	7.5 (7.0)	3.9 (3.8)	13.2 (11.8)	13.3 (13.7)
Refinement				
Resolution (Å)	50.0 – 3.3	50.0 – 2.65	50.0 – 3.2	
No. of reflections	14,284	48,536	27,512	
No. of reflections test set	1,410 (9.9%)	2,368 (4.9%)	1,377 (5.0%)	
R_{work} / R_{free}	23.1 / 27.8	21.1 / 24.9	23.3 / 26.5	
No. atoms (non-hydrogen)	5,674	11,569	11,035	
Protein	5,674	11,495	11,030	
Water	-	71	0	
Ligand/Ions	-	3	5	
B-factors	127	64	136	
Protein	127	65	136	
Water	-	45	-	
Ligand/Ions	-	57	167	
RMSD				
Bond lengths (Å)	0.003	0.004	0.002	
Bond angles (°)	0.6	0.9	0.5	
Ramachandran plot[e]				
Favored (%)	94.2	96.0	95.2	
Additionally allowed (%)	5.8	4.0	4.8	
Outliers (%)	0.0	0.0	0.0	
MolProbity				
Clashscore[e]	3.10 (100^{th})	3.71 (99^{th})	1.27 (100^{th})	
Molprobity score[e]	1.84 (100^{th})	1.43 (100^{th})	1.24 (100^{th})	

[a]APS, Advanced Photon Source
[b]SSRL, Stanford Synchotron Radiation Lightsource
[c]Highest-resolution shell is shown in parentheses
[d]Friedel pairs were merged
[e]As determined by MolProbity
[f]As a reference, two different high-resolution cutoffs are shown

Table S8.

X-ray crystallography analysis of Nup145N^{APD} and Nup145N^{APD}•Nup145C^{N}

Data collection		
Protein	Nup145N^{APD}	Nup145N^{APD}•Nup145C^{N}
PDB ID	5HB5	5HB6
Synchrotron	SSRL[a]	ALS[b]
Beamline	BL12-2	8.2.2
Space group	$P2_1$	P1
Cell dimensions		
a, *b*, *c* (Å)	46.1, 34.9, 78.2	43.0, 44.1, 45.3
α, β, γ (°)	90.0, 99.4, 90.0	99.8, 111.0, 105.9
Wavelength	0.9795	1.0000
Resolution (Å)	30.0 – 1.5	30.0 – 1.3
R_{meas} (%)[c]	9.4 (100.4)	5.9 (110.5)
R_{pim} (%)[c]	3.6 (39.1)	2.2 (56.0)
$CC_{1/2}$[c]	99.9 (87.6)	99.9 (61.7)
$< I / \sigma I >$[c]	12.4 (2.3)	17.2 (1.4)
Completeness (%)[c]	97.2 (92.8)	94.7 (88.9)
No. of observations	257,064	461,854
No. of unique reflections[c,d]	38,743 (3,651)	66,722 (6,234)
Redundancy[c]	6.6 (6.3)	6.9 (3.8)
Refinement		
Resolution (Å)	30.0 – 1.5	30.0 – 1.3
No. of reflections	38,683	66,707
No. of reflections test set	2,012 (5.2%)	1,792 (2.7%)
R_{work} / R_{free}	16.1 / 18.9	15.1 / 17.4
No. atoms (non-hydrogen)	2,753	2,848
Protein	2,331	2,351
Water	331	472
Ligand/Ions	91	25
B-factors	21	25
Protein	18	23
Water	32	36
Ligand/Ions	61	37
RMSD		
Bond lengths (Å)	0.005	0.008
Bond angles (°)	0.8	1.0
Ramachandran plot[e]		
Favored (%)	95.9	94.6
Additionally allowed (%)	4.1	5.1
Outliers (%)	0.0	0.3
MolProbity		
Clashscore[e]	0.42 (100th)	1.47 (99th)
Molprobity score[e]	0.99 (99th)	1.31 (91st)

[a]SSRL, Stanford Synchotron Radiation Lightsource
[b]ALS, Advanced Light Source
[c]Highest-resolution shell is shown in parentheses
[d]Friedel pairs were merged
[e]As determined by MolProbity

Table S9.

X-ray crystallographic analysis of $Nup53^{RRM}$

Data collection					
Protein	$Nup53^{RRM}$ [e]	$Nup53^{RRM}$ [e]	$Nup53^{RRM}$	$Nup53^{RRM}$	$Nup53^{RRM}$
PDB ID	5HB7			5HB8	
Synchrotron	SSRL[a]	SSRL[a]	SSRL	SSRL	SSRL
Beamline	BL12-2	BL12-2	BL12-2	BL12-2	BL12-2
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P3_121$	$P3_121$
Cell dimensions					
a, *b*, *c* (Å)	35.0, 50.4, 60.0	35.0, 50.4, 60.0	34.9, 50.4, 60.0	94.7, 94.7, 115.5	93.8, 97.8, 114.1
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
			Se Peak		*Os Peak*
Wavelength	0.7293	0.7293	0.9794	0.9794	1.1399
Resolution (Å)	40.0 – 0.80	40.0 – 1.0	40.0 – 1.1	50.0 – 1.7	20.0 – 3.1
R_{meas} (%)[b]	4.1 (94.5)	3.8 (12.1)	10.4 (111.0)	7.9 (155.6)	13.5 (154.4)
R_{pim} (%)[b]	1.3 (33.2)	1.2 (3.7)	2.1 (28.1)	1.3 (29.6)	3.7 (41.8)
$CC_{1/2}$[b]	99.9 (66.4)	99.9 (99.6)	99.9 (66.6)	100.0 (74.5)	99.8 (76.1)
$< I / \sigma I >$[b]	29.3 (2.3)	43.5 (18.8)	19.8 (2.1)	37.3 (2.2)	20.2 (1.8)
Completeness (%)[b]	88.4 (50.3)	97.1 (94.8)	95.4 (66.8)	91.1 (93.1)	99.9 (99.3)
No. of observations	890,036	550,262	984,014	2,169,719	147,214
No. of unique reflections[b,c]	92,712 (5,211)	56,349 (5,425)	42,372 (2,914)	60,735 (6,155)	10,957 (1,073)
Redundancy[b]	9.6 (7.6)	9.8 (10.3)	23.2 (14.6)	35.7 (25.9)	13.4 (13.2)
Refinement					
Resolution (Å)	40.0 – 0.8			50.0 – 1.7	
No. of reflections	92,709			59,806	
No. of reflections test set	4,650 (5.0%)			1,812 (3.0%)	
R_{work} / R_{free}	11.8 / 13.2			16.3 / 21.0	
No. atoms (non-hydrogen)	1,448			4,176	
Protein	1,192			3,632	
Water	253			466	
Ligand/Ions	3			78	
B-factors	15			38	
Protein	11			37	
Water	31			43	
Ligand/Ions	16			73	
RMSD					
Bond lengths (Å)	0.008			0.010	
Bond angles (°)	1.1			1.0	
Ramachandran plot[d]					
Favored (%)	98.6			99.3	
Additionally allowed (%)	1.4			0.7	
Outliers (%)	0.0			0.0	
MolProbity					
Clashscore[d]	0.43 (95^{th})			0.96 (99^{th})	
Molprobity score[d]	0.65 (98^{th})			0.88 (100^{th})	

[a]SSRL, Stanford Synchotron Radiation Lightsource
[b]Highest-resolution shell is shown in parentheses
[c]Friedel pairs were merged
[d]As determined by MolProbity
[e]As a reference, two different high-resolution cutoffs are shown

Table S10.

Molecular Weights and Stoichiometry in Composite Structure

C. thermophilum Protein	Stoichiometry	Molecular Weight (kDa)	Total Mass (kDa)
Symmetric Core (observed in composite structure)			
Nup120	32	141.8	4537.6
Nup37	32	80.3	2569.6
-	-	-	-
Nup85	32	128.1	4099.2
-	-	-	-
Sec13	32	33.8	1081.6
Nup145C	32	88.8	2841.6
Nup84	32	108.1	3459.2
Nup133	32	150.7	4822.4
Nup170	48	156.2	7497.6
Nup192	32	196.9	6300.8
Nic96	32	121.5	3888.0
Nup57	32	36.9	1180.8
Nup49	32	48.6	1555.2
Nsp1	32	67.1	2147.2
Nup53	32	47.0	1504.0
Nup145N	32	99.8	3193.6
Nup188	16	203.8	3260.8
TOTAL			**53939.2**
Cytoplasmic Filaments			
-	-	-	
Nup159	-	156.2	
Nup82	-	98.1	
Gle1	-	58.0	
Gle2	-	39.4	
Nup42	-	58.0	
Nuclear Basket			
Mlp1	-	231.1	
Nup152	-	152.2	
Nup56	-	56.2	
ELYS	-	33.1	
POMs/Other			
Pom152	-	141.5	
-	-	-	
Ndc1	-	71.8	
Pom34	-	34.2	
-	-	-	

H. sapiens Protein	Stoichiometry	Molecular Weight (kDa)	Total Mass (kDa)
Symmetric Core (observed in composite structure)			
Nup160	32	162.1	5187.2
Nup37	32	36.7	1174.4
Nup43	32	42.2	1350.4
Nup85	32	75.0	2400.0
Seh1	32	39.7	1270.4
Sec13	32	35.5	1136.0
Nup96	32	105.9	3388.8
Nup107	32	106.4	3404.8
Nup133	32	129.0	4128.0
Nup155	48	155.2	7449.6
Nup205	32	227.9	7292.8
Nup93	32	93.5	2992.0
Nup54	32	55.4	1772.8
Nup58	32	53.2	1702.4
Nup62	32	53.3	1705.6
Nup53	32	34.8	1113.6
Nup98	32	91.7	2934.4
Nup188	16	196.0	3136.0
TOTAL			**53539.0**
Cytoplasmic Filaments			
Nup358	-	358.2	
Nup214	-	213.6	
Nup88	-	83.5	
Gle1	-	79.8	
Rae1	-	41.0	
Nupl2	-	44.9	
Nuclear Basket			
TPR	-	267.3	
Nup153	-	153.9	
Nup50	-	50.1	
ELYS	-	252.5	
POMs/Other			
Gp210	-	205.1	
POM121	-	127.7	
NDC1	-	76.3	
ALADIN	-	59.6	

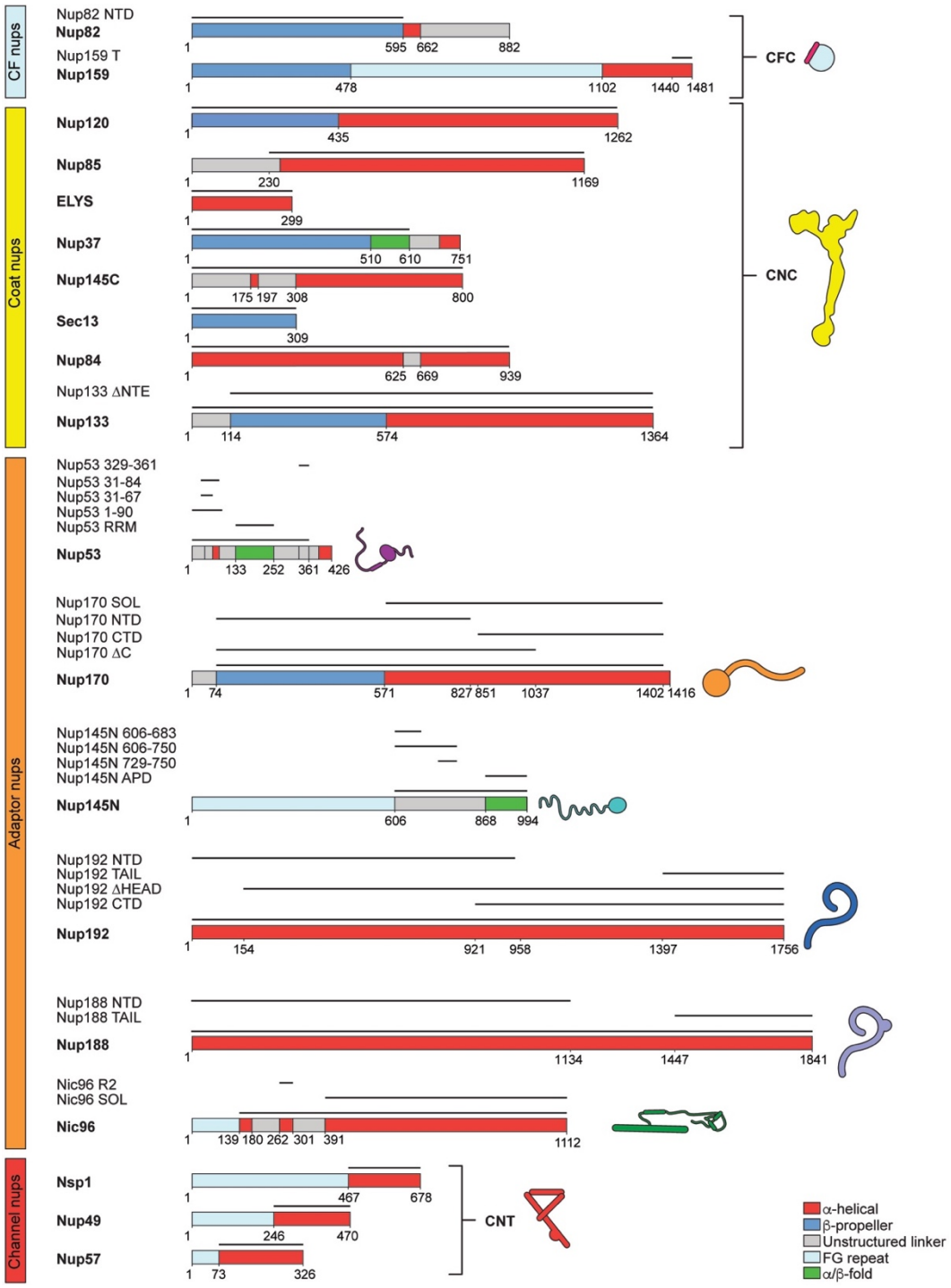


Fig.S1.

Nucleoporin fragments. Domain boundaries of all nucleoporin fragments used throughout the text are indicated by black lines.

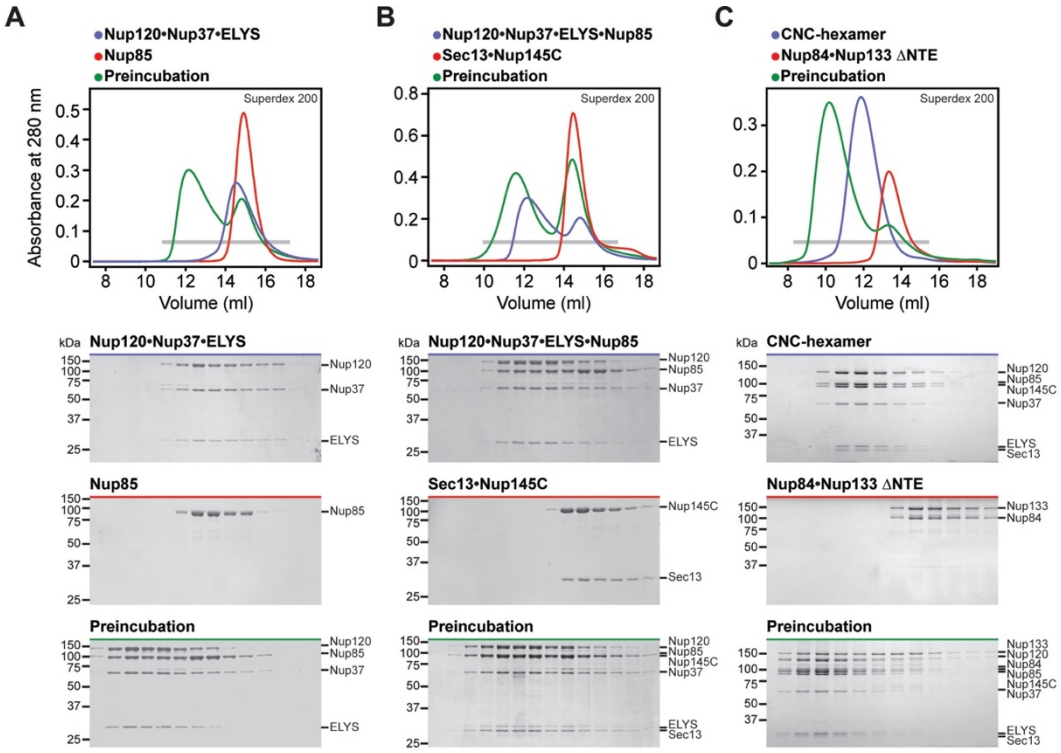


Fig. S2.

Reconstitution of the CNC. SEC and SDS-PAGE analysis for the reconstitution of (**A**) the Nup120•Nup37•ELYS•Nup85 hetero-tetramer, (**B**) the CNC-hexamer, and (**C**) the CNC-octamer. SEC profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. The CNC-octamer could only be reconstituted with a Nup133 construct lacking the NTE.

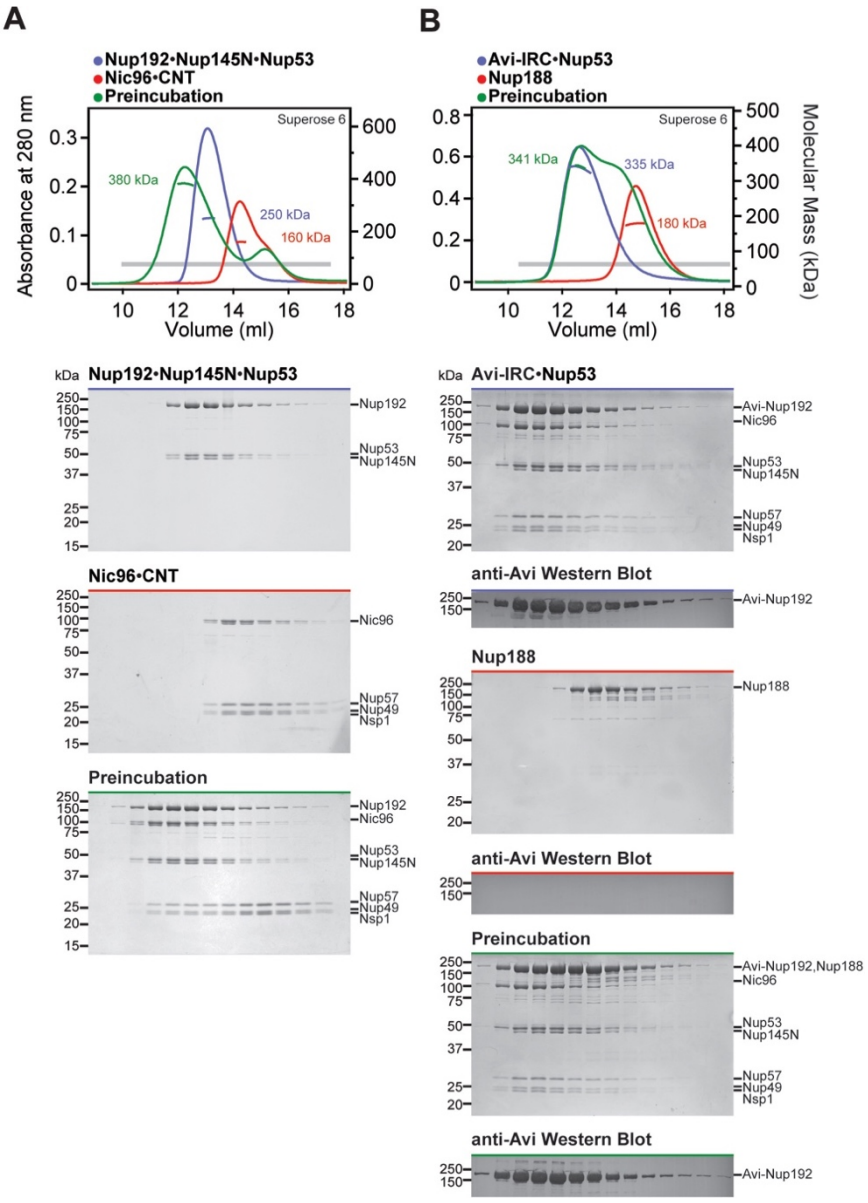


Fig. S3.

Reconstitution of the IRC•Nup53 complex. SEC-MALS and SDS-PAGE analysis of (**A**) the reconstitution of the IRC•Nup53 hetero-heptamer and (**B**) the interaction of IRC•Nup53 with Nup188. SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining or by immunoblotting with a mouse anti-AviTag antibody. As Nup188 and Nup192 cannot be distinguished by SDS-PAGE analysis, lack of Nup192 displacement from the IRC by Nup188 was confirmed by western blotting against Avi-tagged Nup192.

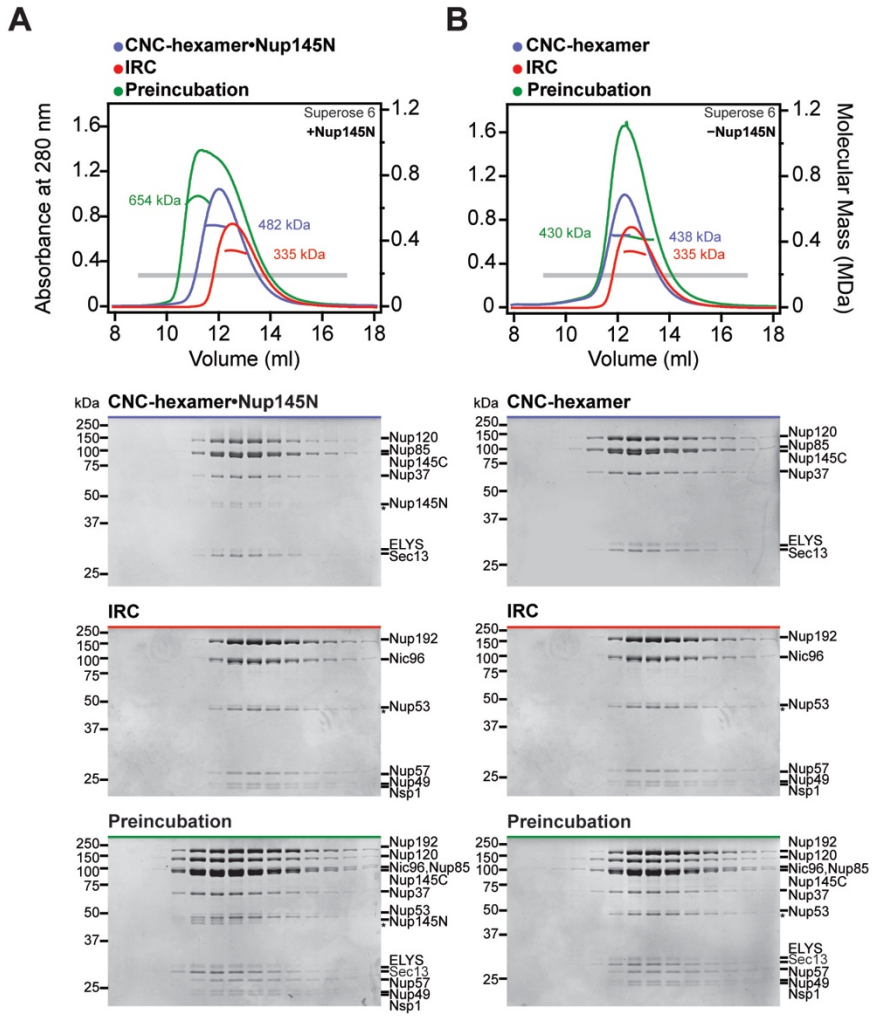


Fig. S4.

Reconstitution of NPC core protomers. (**A**-**D**) SEC-MALS and SDS-PAGE analysis corresponding to Fig. 1, D to G. SEC-MALS profiles of nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

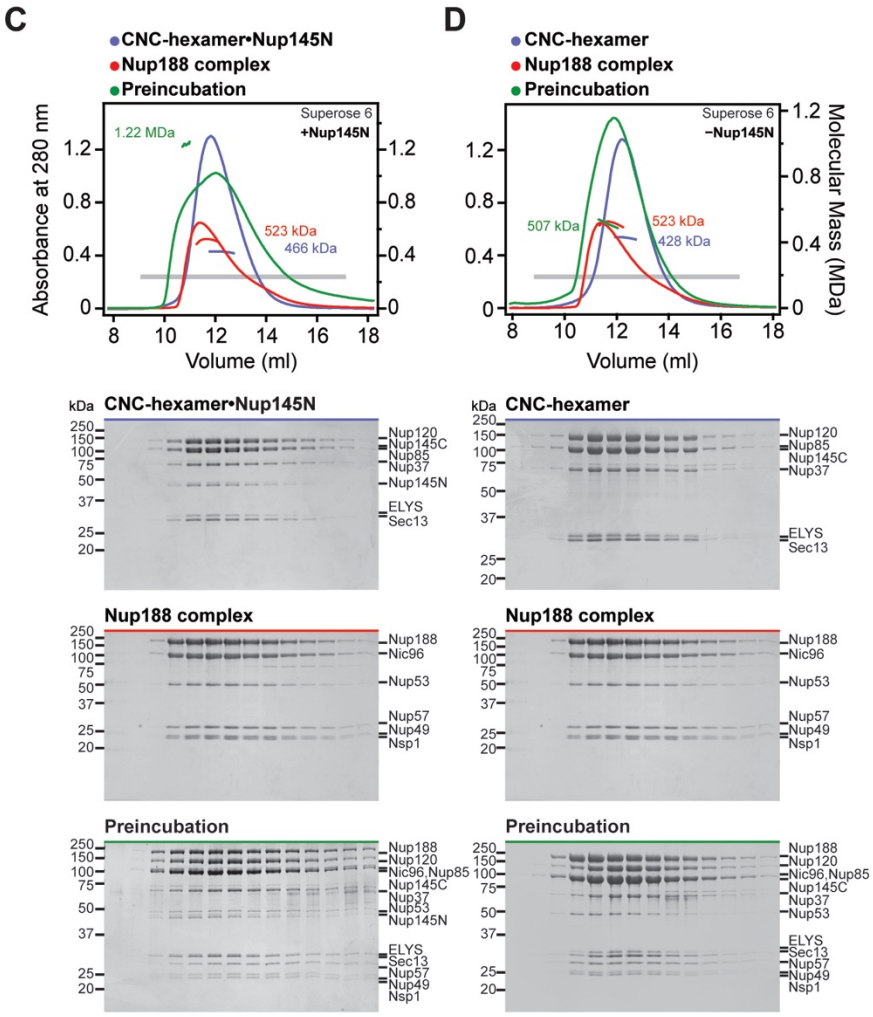


Fig. S4 continued.

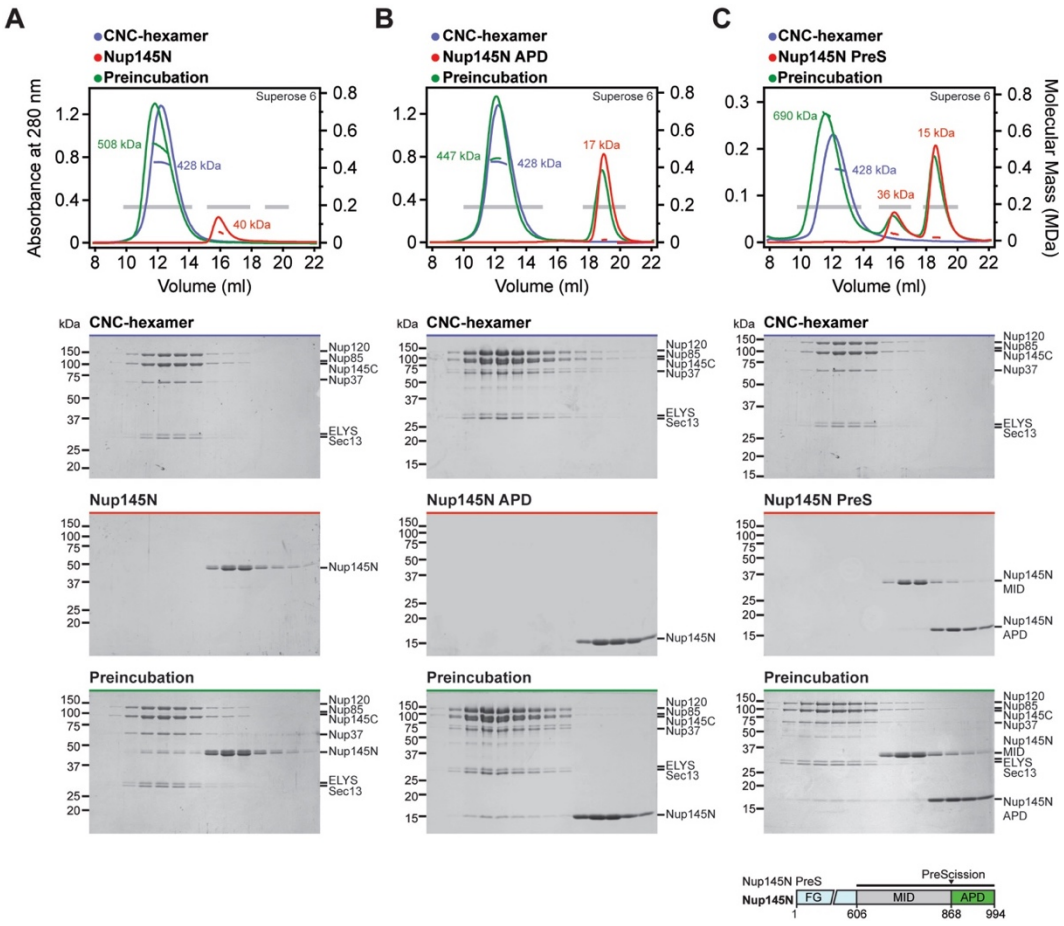


Fig. S5.

The CNC-hexamer interacts with Nup145N^{APD}. SEC-MALS and SDS-PAGE analysis of the CNC-hexamer interaction with (**A**) Nup145N, (**B**) Nup145N^{APD}, and (**C**) Nup145N PreS preincubated with PreScission protease. SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. As reference, the domain structure of Nup145N PreS is shown, indicating the construct boundaries and PreScission cleavage site with a black bar and a black triangle, respectively. After cleavage with PreScission protease, Nup145N^{APD} co-elutes with the CNC-hexamer, whereas Nup145N^{MID} does not incorporate into the complex.

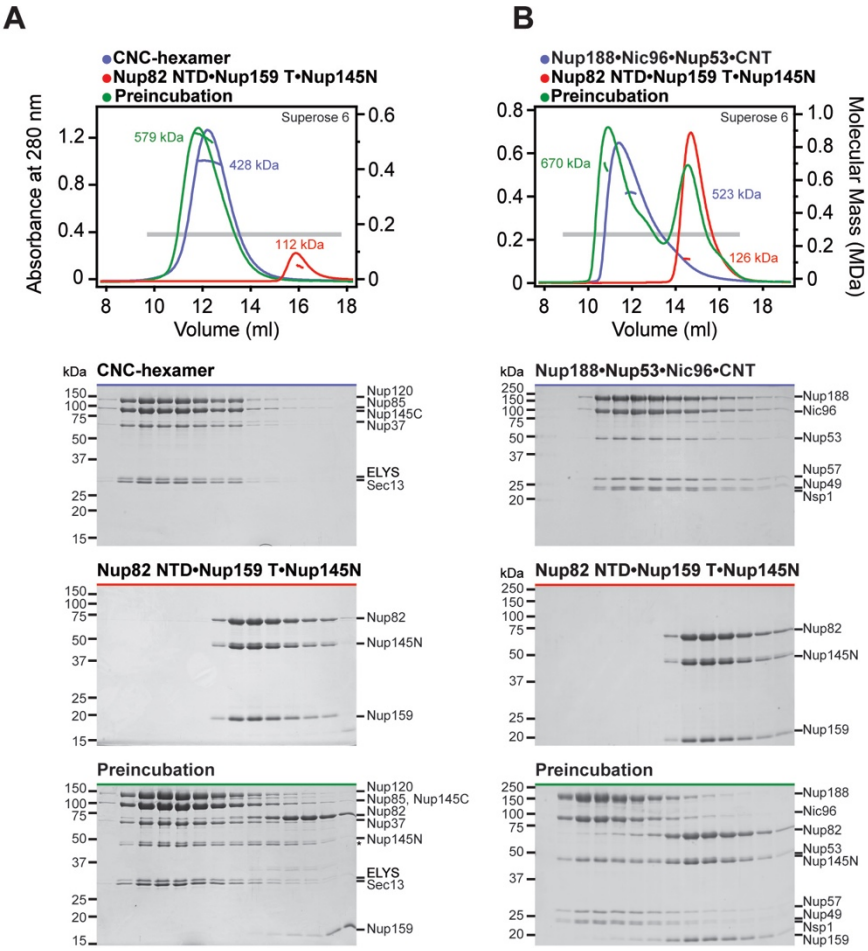


Fig. S6.

Incorporation of Nup145N into complexes with CNC-hexamer or Nup188•Nic96•Nup53•CNT is exclusive of Nup145N binding to Nup82^{NTD}•Nup159^{T}. SEC-MALS and SDS-PAGE analysis of the CFC interaction with (**A**) the CNC-hexamer or (**B**) the Nup188•Nic96•Nup53•CNT complex. SEC-MALS profiles of nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. In both preincubations, Nup145N incorporates stoichiometrically into the complexes, but the Nup82^{NTD}•Nup159^{T} hetero-dimer does not. However, Nup82^{NTD}•Nup159^{T} interacts weakly with the Nup188-containing complex.

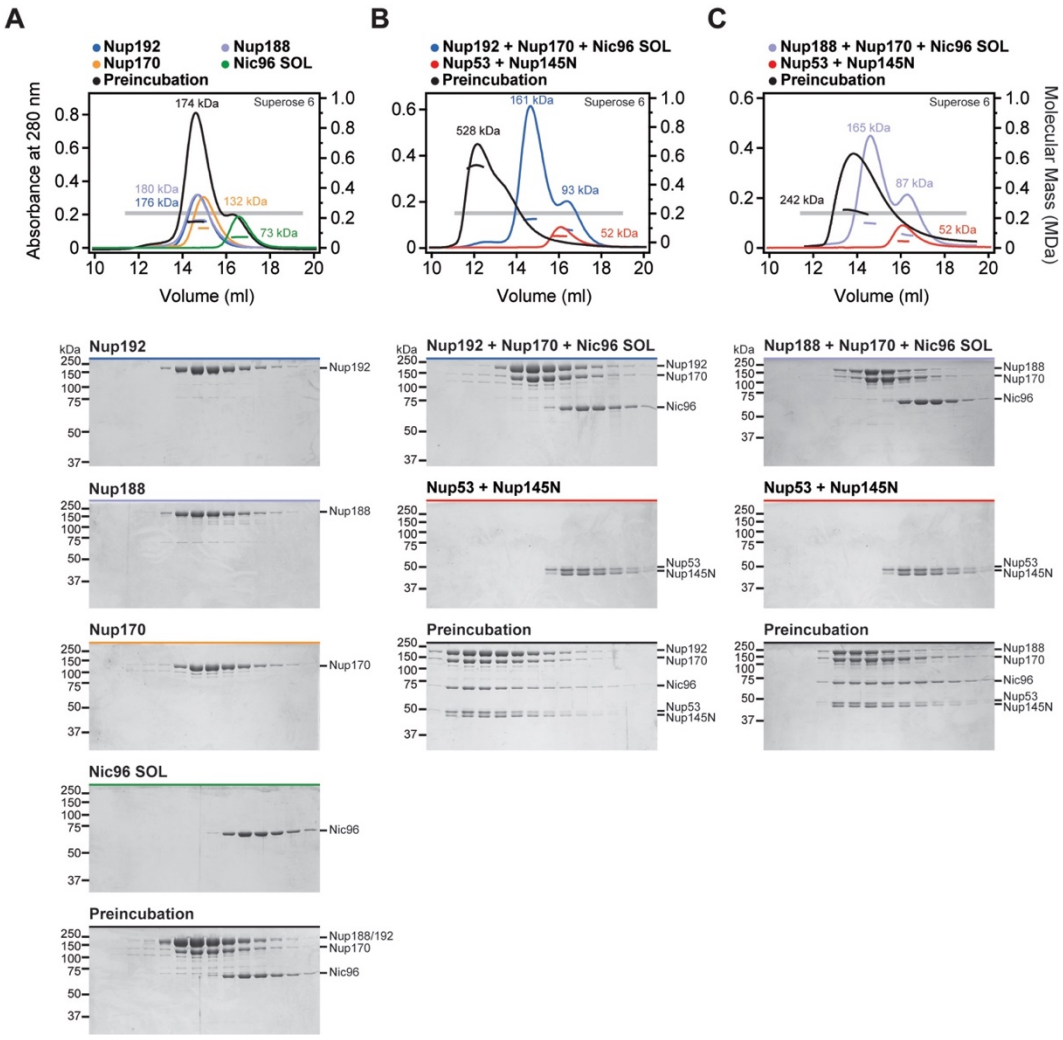


Fig. S7.

Scaffold nucleoporins are assembled by the linker nucleoporins. SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2, A to C. (**A**) For interaction analysis between scaffold nucleoporins, SEC-MALS profiles of nucleoporins are shown individually (blue, purple, orange, or green) and after their preincubation (black). (**B**) To analyze the assembly of Nup192, Nup170, and Nic96^{SOL} with the linker nucleoporins Nup53 and Nup145N, SEC-MALS profiles of the scaffold mixture (blue), the linker mixture (red), or after their preincubation (black) are shown. (**C**) To analyze the assembly of Nup188, Nup170 and Nic96^{SOL} with the linker nucleoporins Nup53 and Nup145N, SEC-MALS profiles of the scaffold mixture (purple), the linker mixture (red), or after their preincubation (black) are shown. All SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Because Nup188 could not be incorporated into the IRC, Nup188 and Nup192 were not simultaneously included in preincubations.

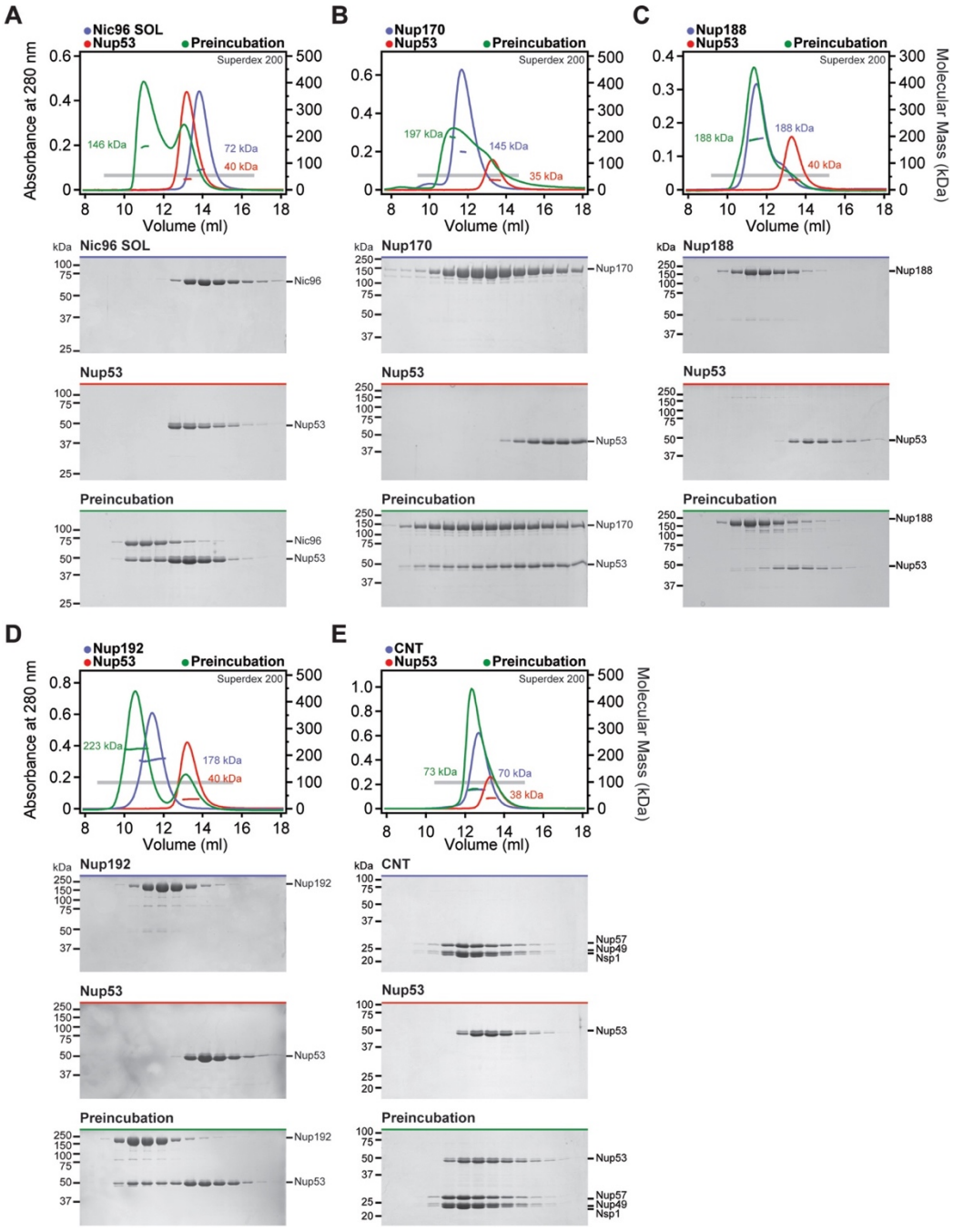


Fig. S8.

Identification of interactions between Nup53 and the scaffold nucleoporins. (**A**-**E**) SEC-MALS and SDS-PAGE analysis for Fig. 2D. SEC-MALS profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Robust complex formation with Nup53 was observed for Nic96^{SOL}, Nup170, and Nup192. Interactions of Nup53 with Nup188 and the CNT were only barely detectable.

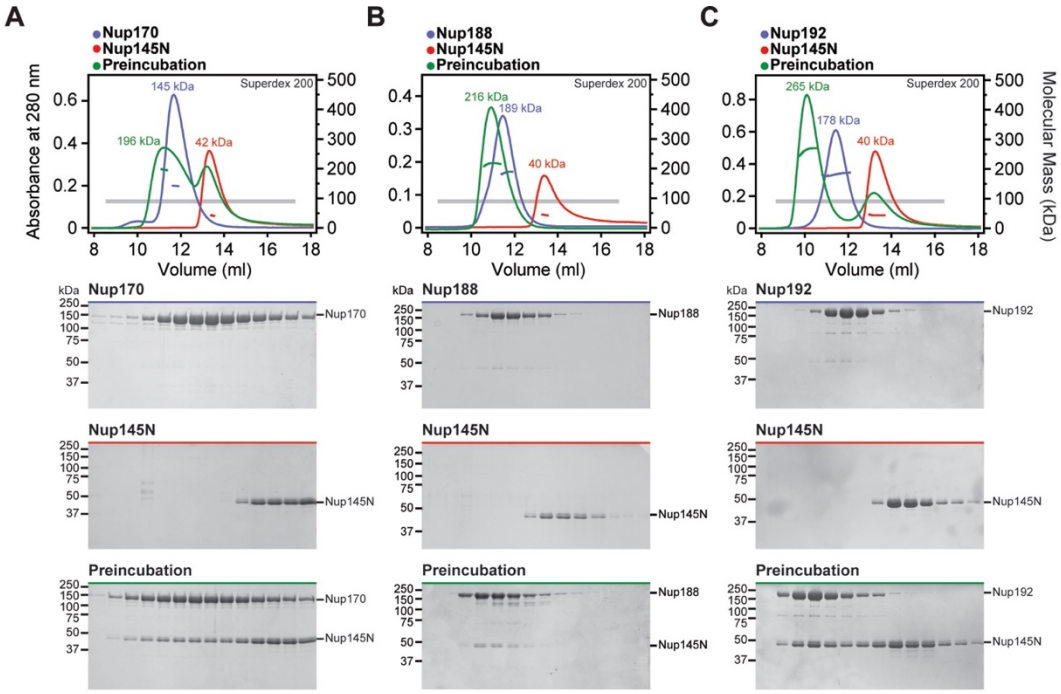


Fig. S9.

Identification of interactions between Nup145N and scaffold nucleoporins. (**A-C**) SEC-MALS and SDS-PAGE analysis for Fig. 2E. SEC-MALS profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Robust complex formation with Nup145N was observed for Nup170, Nup188, and Nup192. We previously demonstrated interactions between Nup145N and Nic96 and the CNT (*15*).

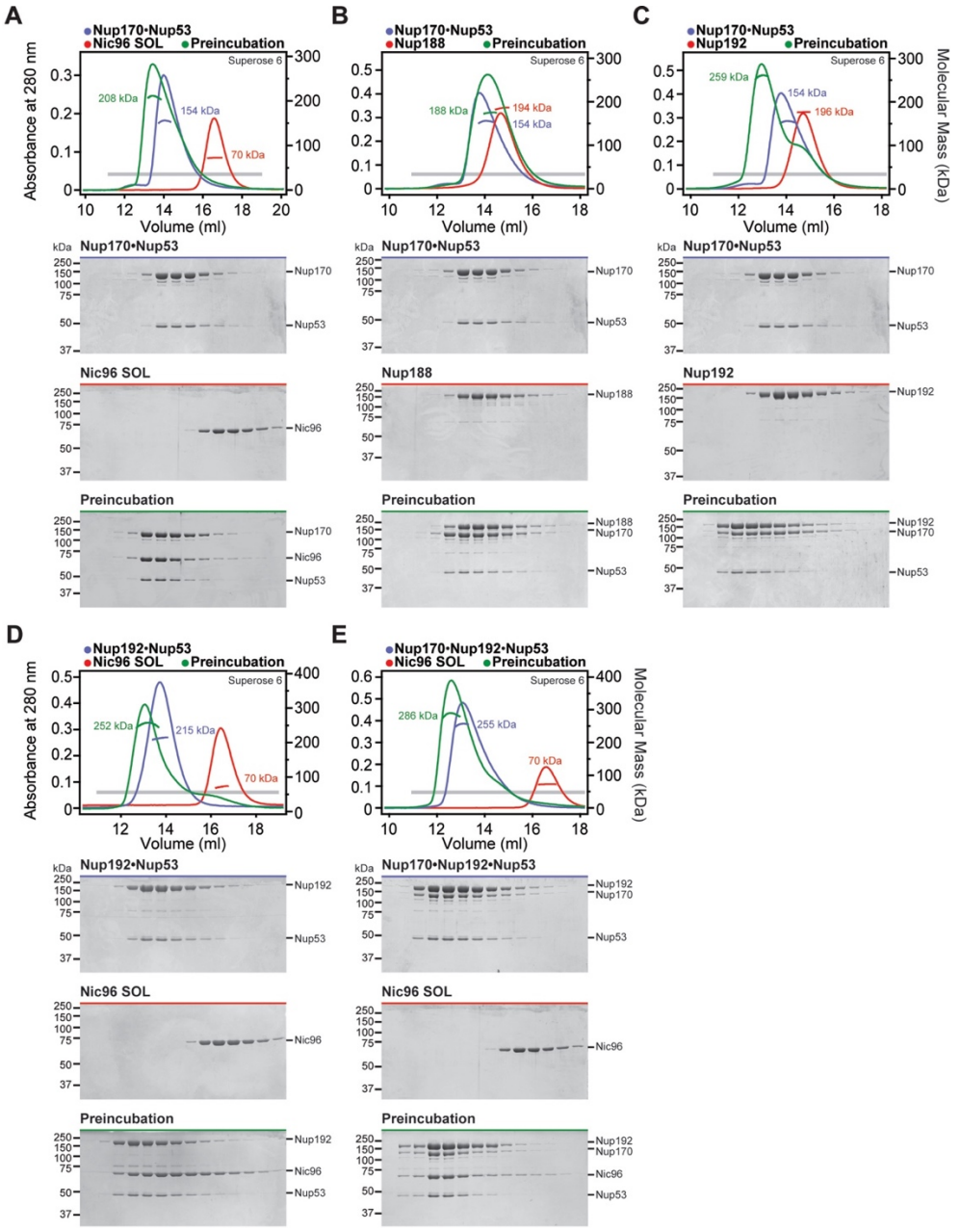


Fig. S10.

Identification of hetero-trimeric complexes assembled by Nup53. (**A-E**) SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2D. Purified hetero-dimeric Nup53 complexes were tested for their ability to form hetero-trimers with an additional scaffold nucleoporin. SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Nup53 is able to bind simultaneously to Nup170, Nic96^{SOL}, and Nup192. The interaction of the Nup170•Nup53 hetero-dimer with Nup188 was only barely detectable.

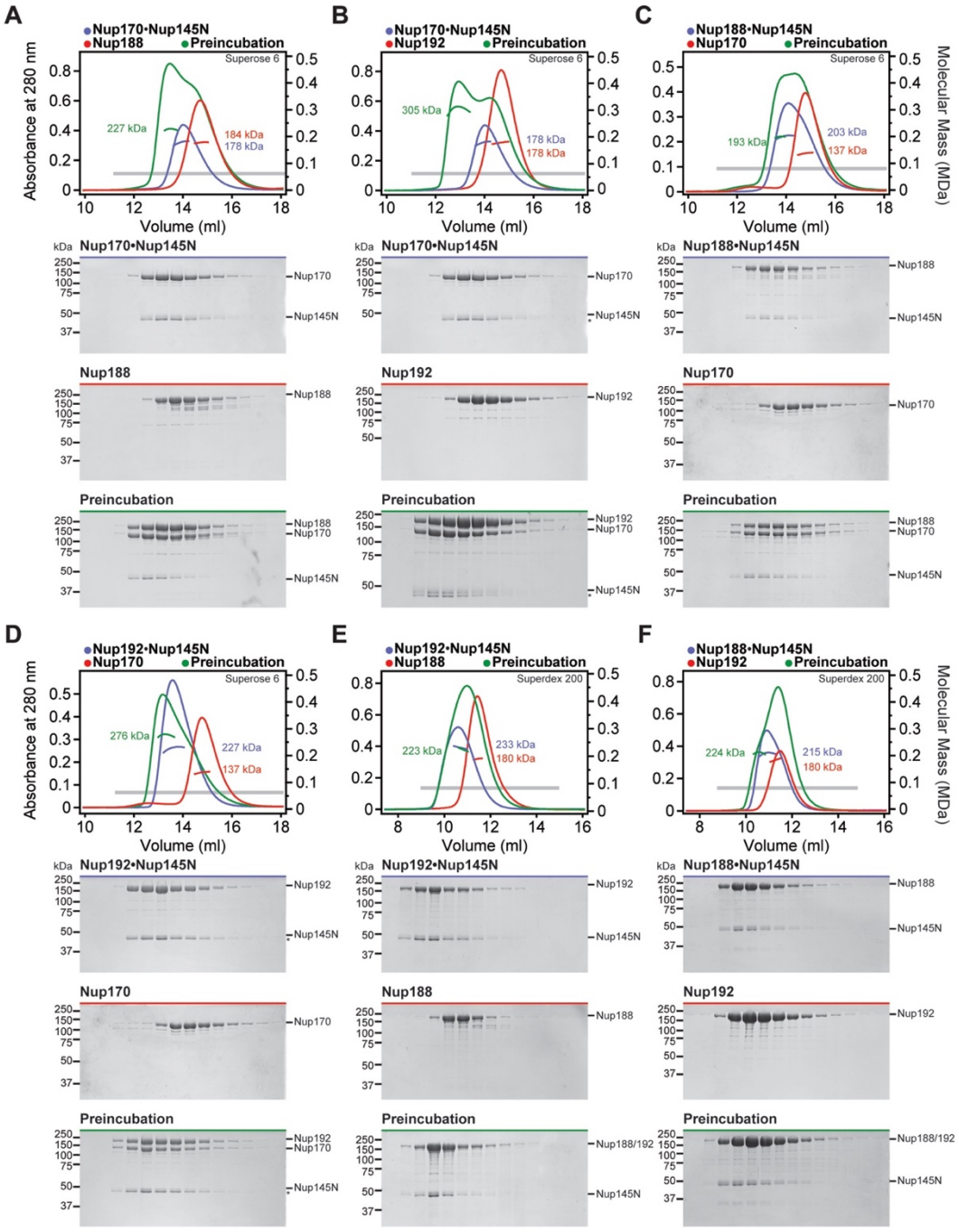


Fig. S11.

Identification of hetero-trimeric complexes assembled by Nup145N. (**A-G**) SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2E. Purified hetero-dimeric Nup145N complexes were tested for their ability to form hetero-trimers with an additional scaffold

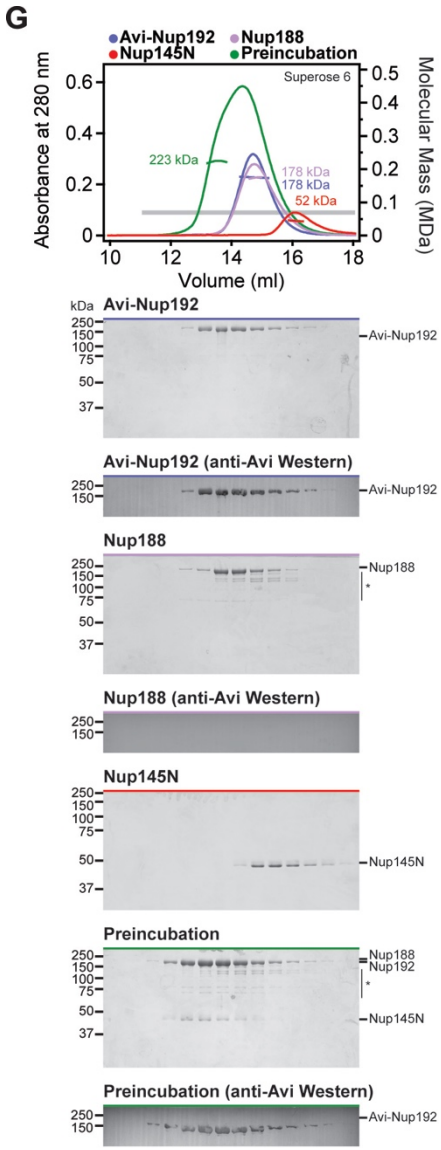


Fig. S11 continued.
nucleoporin. SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue, red, and purple) and after their preincubation (green). SEC profiles were obtained using either a Superose 6 10/300 GL column or a Superdex 200 10/300 GL column as indicated. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining or by western blot analysis with a mouse anti-AviTag antibody. Hetero-trimeric complex formation was determined on the basis of a shift in elution volume, an increase in measured molecular mass, and the presence of all three proteins in the higher molecular mass fractions. For preincubations with Nup188 or Nup188•Nup145N, only minor shifts in elution volume and no increases in the measured molecular masses were observed, indicating the formation of a mixture of species rather than stable hetero-trimeric complexes as seen between Nup170, Nup192, and Nup145N. As Nup188 and Nup192 cannot be distinguished by SDS-PAGE, western blot analysis was employed to confirm that Nup188 and Nup192 bind to Nup145N in a mutually exclusive fashion.

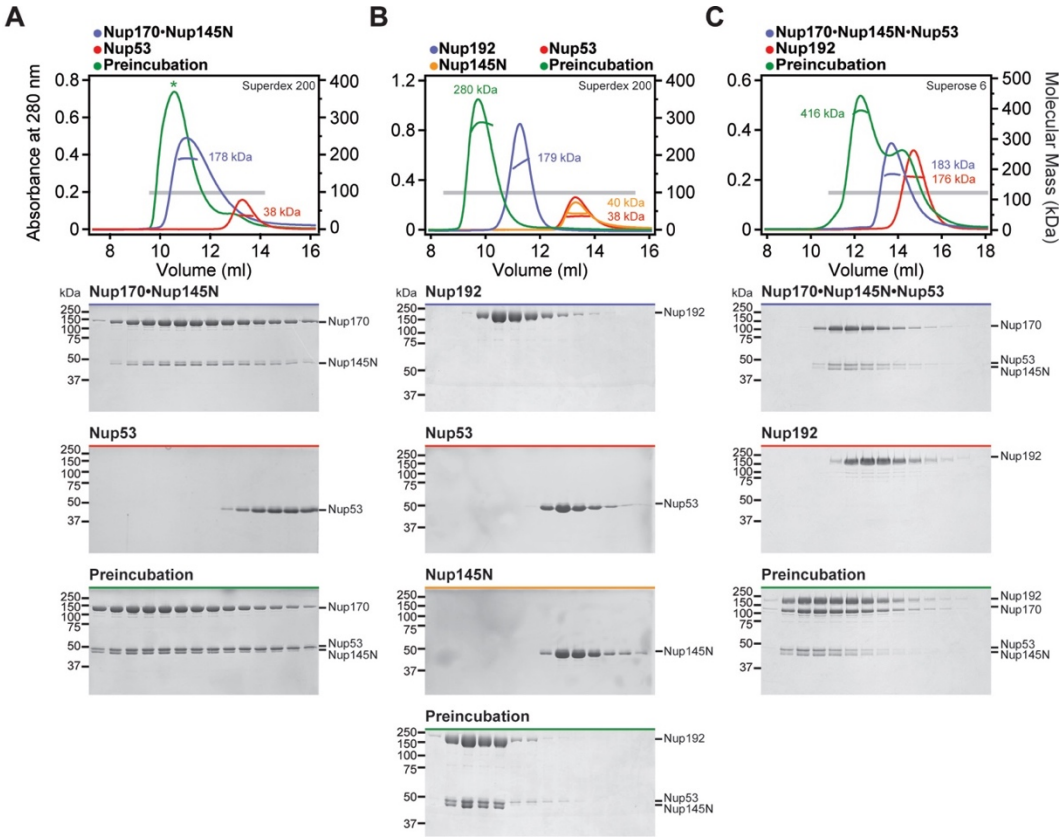


Fig. S12.

Binding sites for linker nucleoporins on Nup170 and Nup192 are distinct and compatible. (**A**-**C**) SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue, red, or yellow) and after their preincubation (green). SEC profiles were obtained using a Superdex 200 10/300 GL column for panels (A, B) and a Superose 6 10/300 GL column for panel (C). Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. We were unable to obtain an accurate molecular mass measurement for the preincubation of Nup53 with Nup170•Nup145N, as indicated by an asterisk above the chromatogram.

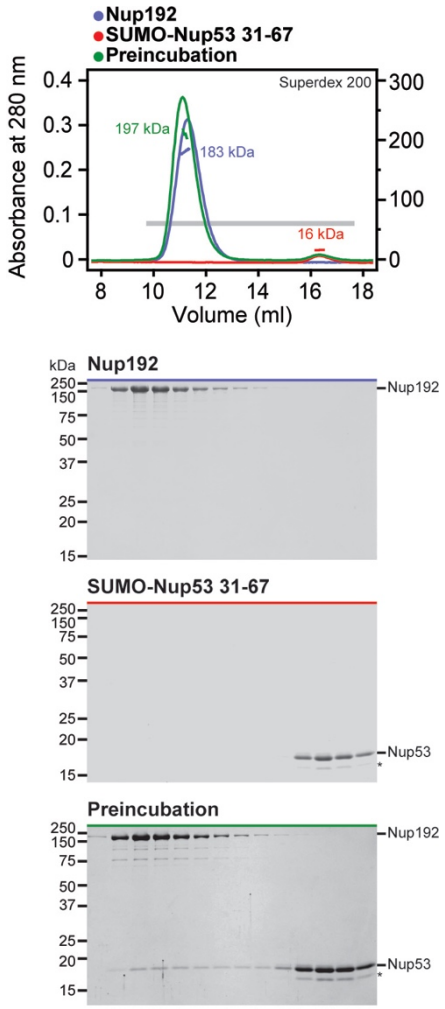


Fig. S13.
Nup192 recognizes Nup53$^{31-67}$. SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2F. SEC-MALS profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

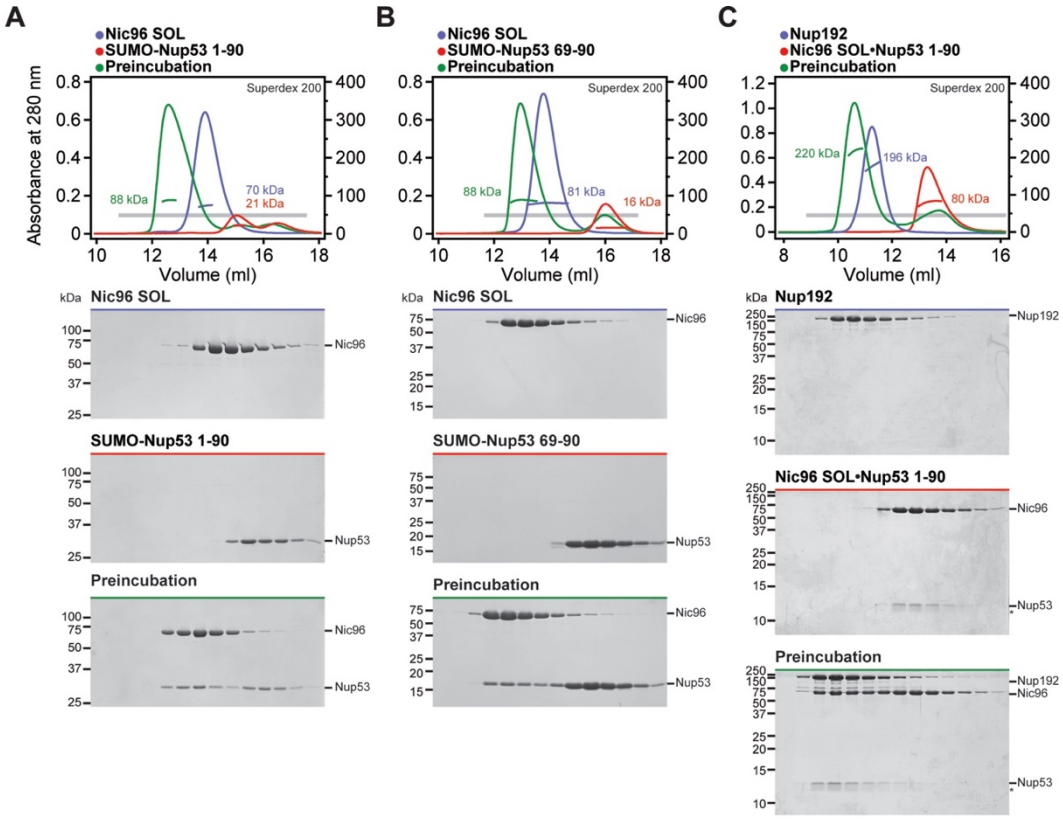


Fig. S14.
Nic96^{SOL} recognizes Nup53$^{69-90}$ and Nup53$^{1-90}$ binds simultaneously to Nup192 and Nic96^{SOL}. (**A-C**) SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2F. SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. The Nup53 binding sites for Nic96^{SOL} and Nup192 map to directly adjacent sequence fragments and are not mutually exclusive.

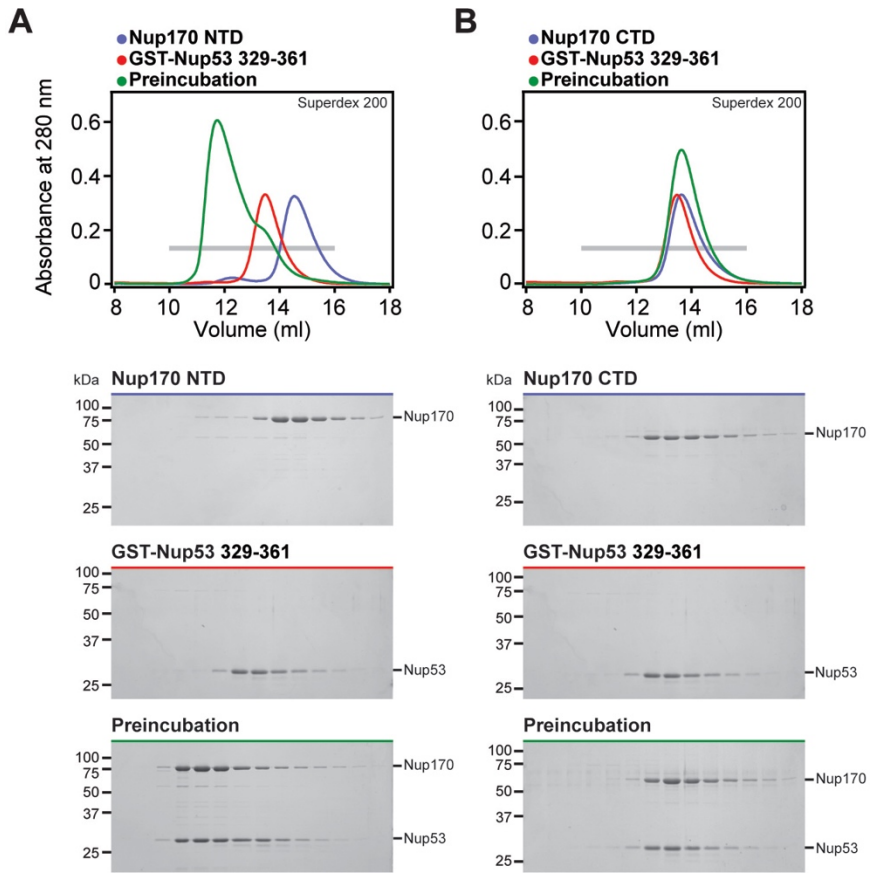


Fig. S15.
Nup170^{NTD} recognizes Nup53$^{329-361}$. (**A**, **B**) SEC and SDS-PAGE analysis corresponding to Fig. 2F. SEC profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

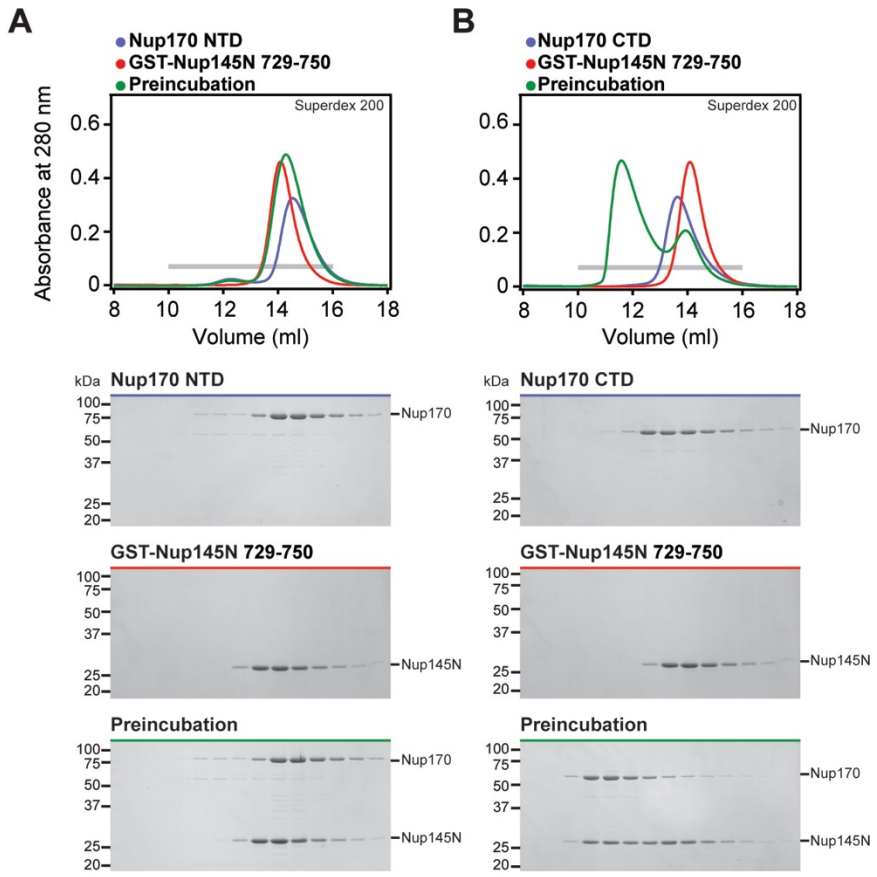


Fig. S16.
Nup170^{CTD} recognizes Nup145N$^{729-750}$. (**A**, **B**) SEC and SDS-PAGE analysis corresponding to Fig. 2F. SEC profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

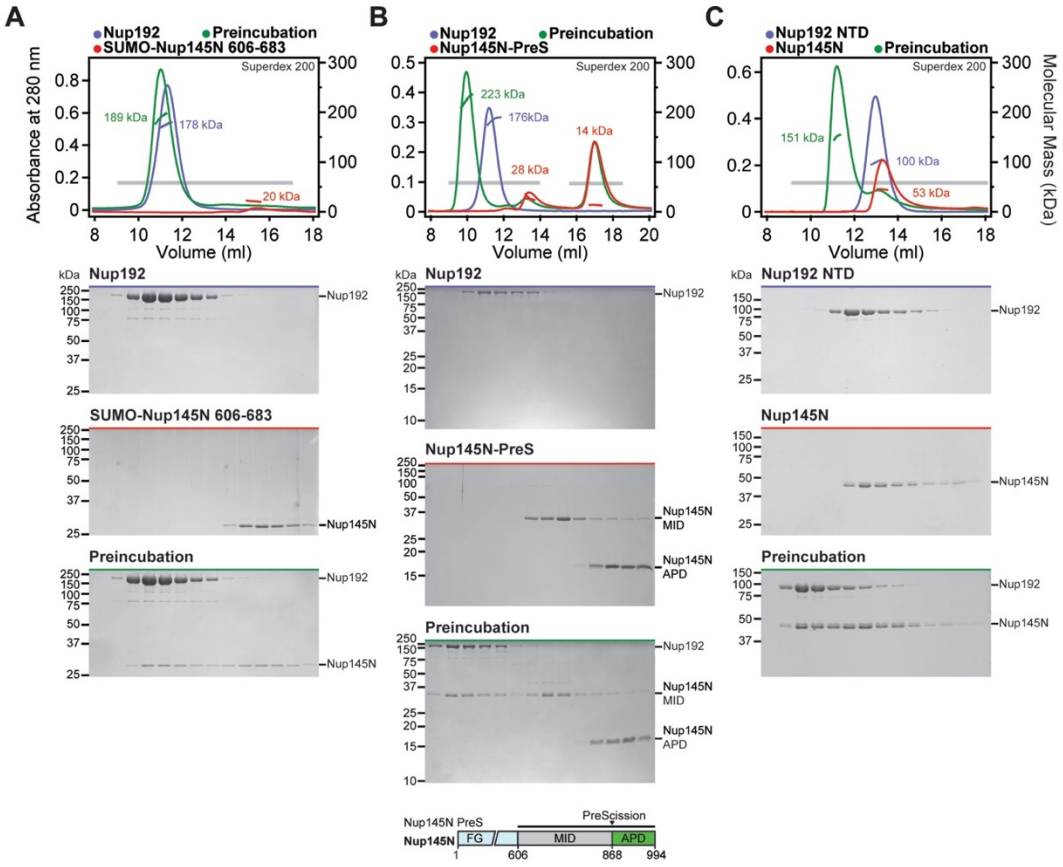


Fig. S17.

Nup192 recognizes Nup145N$^{606-683}$ primarily with its NTD. (**A**-**E**) SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2F. SEC-MALS profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Nup192 binds to a minimal sequence on Nup145N between residues 606-683. After cleavage of Nup145N^{PreS} with PreScission protease, Nup145N^{MID} co-elutes with Nup192, whereas Nup145N^{APD} does not interact with Nup192. As a reference, the domain structure of Nup145N^{PreS} is shown, indicating the construct boundaries and PreScission cleavage site with a black bar and a black triangle, respectively. Nup145N forms a stoichiometric complex with Nup192^{NTD}, but Nup145N binding to Nup192^{CTD} and Nup192^{TAIL} were barely detectable and undetectable, respectively.

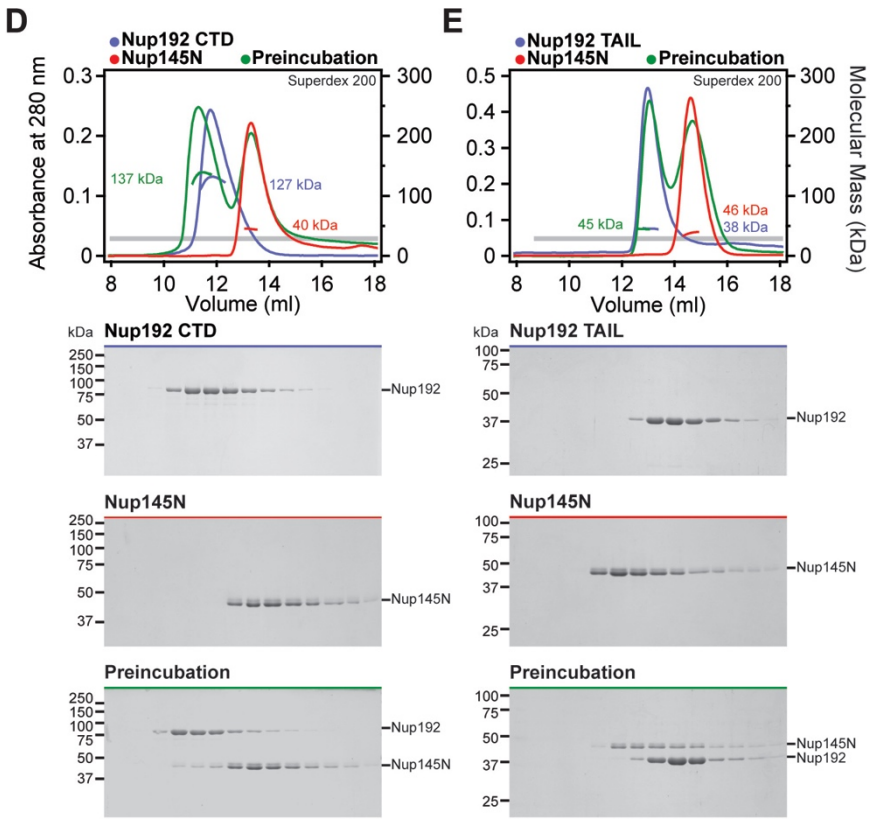


Fig. S17 continued.

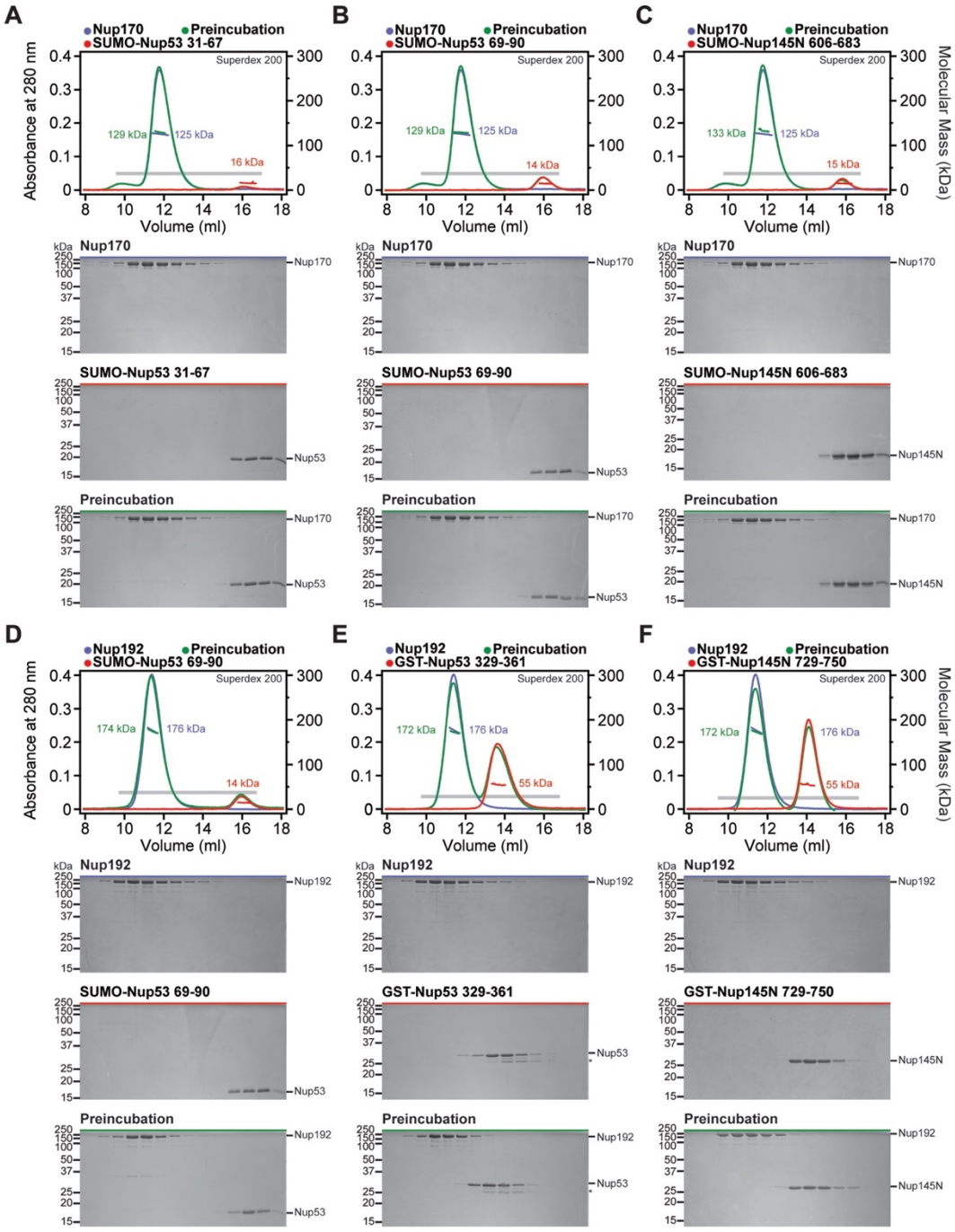

A
Nup170
Preincubation
SUMO-Nup53 31-67
Superdex 200
129 kDa
125 kDa
16 kDa
Absorbance at 280 nm
Molecular Mass (kDa)
Volume (ml)
B
Nup170
Preincubation
SUMO-Nup53 69-90
Superdex 200
129 kDa
125 kDa
14 kDa
C
Nup170
Preincubation
SUMO-Nup145N 606-683
Superdex 200
133 kDa
125 kDa
15 kDa
kDa
Nup170
SUMO-Nup53 31-67
Nup53
Preincubation
SUMO-Nup53 69-90
SUMO-Nup145N 606-683
Nup145N
D
Nup192
Preincubation
SUMO-Nup53 69-90
Superdex 200
174 kDa
176 kDa
14 kDa
E
Nup192
Preincubation
GST-Nup53 329-361
Superdex 200
172 kDa
176 kDa
55 kDa
F
Nup192
Preincubation
GST-Nup145N 729-750
Superdex 200
172 kDa
176 kDa
55 kDa
Nup192
SUMO-Nup53 69-90
GST-Nup53 329-361
GST-Nup145N 729-750

Fig. S18.

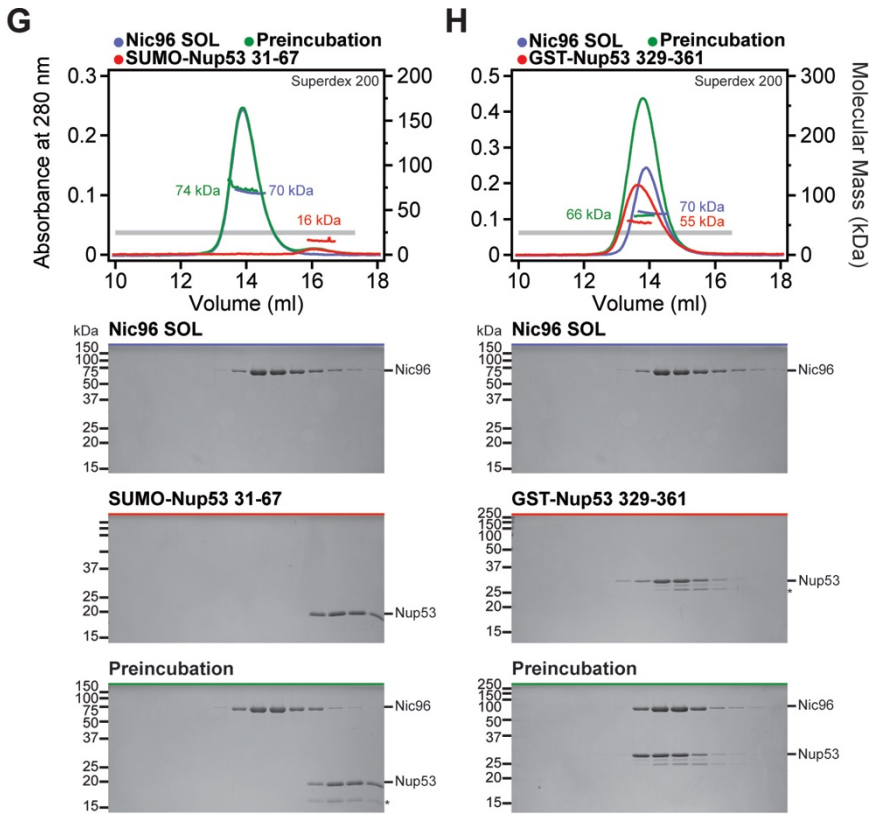


Fig. S18 continued.

Mapped minimal linker nucleoporin fragments are specific for the identified binding partners. (**A-H**) SEC-MALS profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. The minimal sequence fragments of the linker nucleoporins Nup53 and Nup145N did not interact with other scaffold nucleoporins, demonstrating their specificity for a single scaffold nucleoporin.

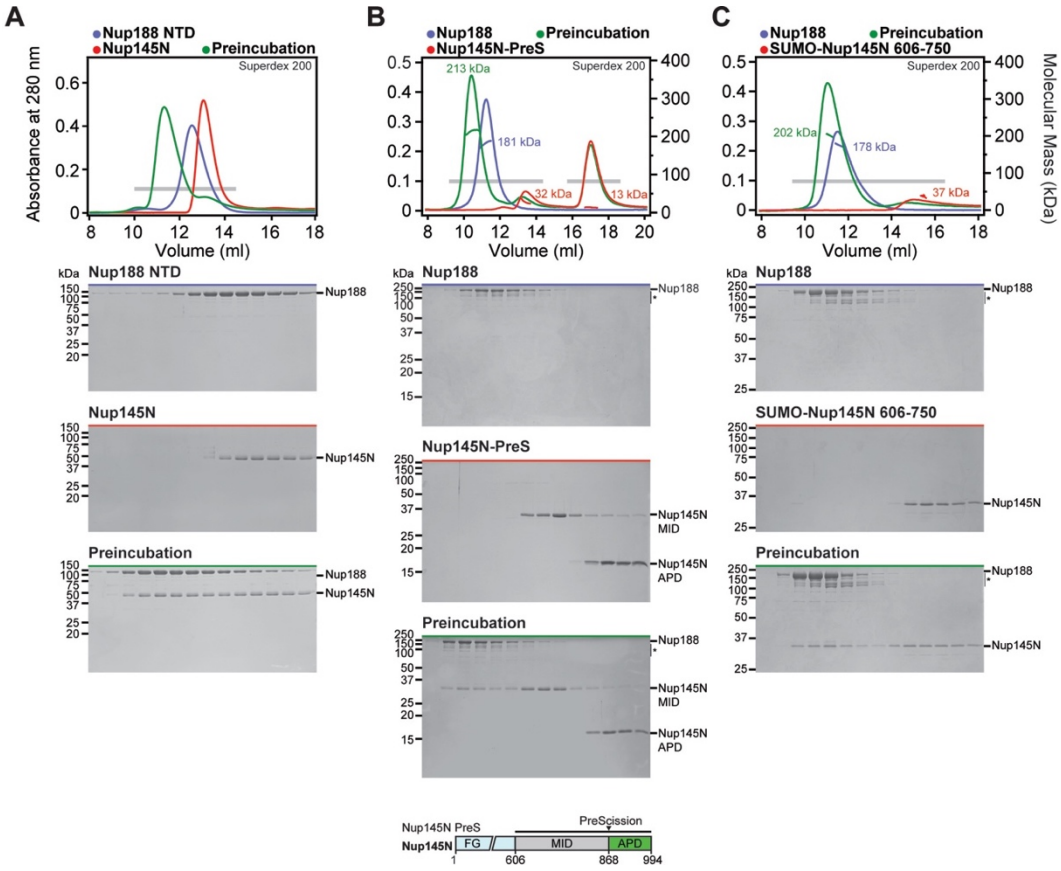


Fig. S19.

Nup188 recognizes Nup145N$^{606\text{-}750}$ primarily via its NTD. (**A**-**E**) SEC-MALS and SDS-PAGE analysis corresponding to Fig. 2F. SEC-MALS profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Nup188^{NTD} is responsible for binding Nup145N. After Nup145N^{PreS} cleavage with PreScission protease, Nup145N^{MID} co-elutes with Nup188, whereas Nup145N^{APD} does not interact with Nup188. As reference, the domain structure of Nup145N^{PreS} is shown, indicating the construct boundaries and PreScission cleavage site with a black bar and a black triangle, respectively. Nup188 binds to a minimal sequence on Nup145N between residues 606-750, but does not interact with the minimal Nup145N binding sequences for Nup192 or Nup170.

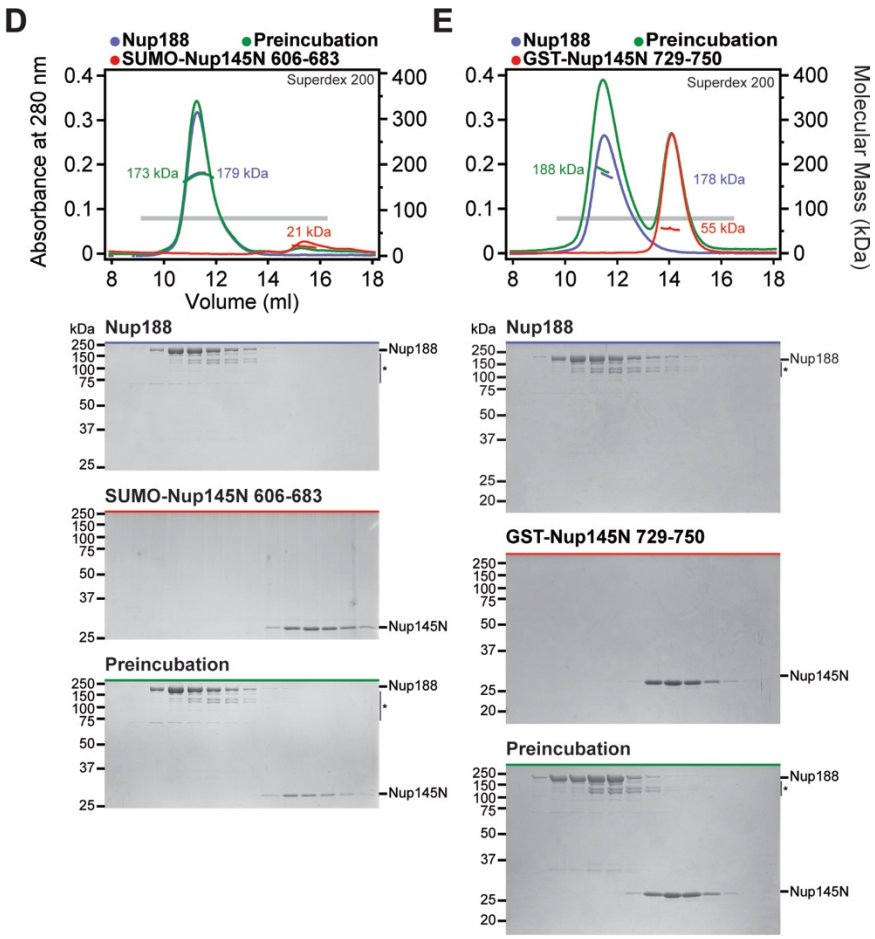


Fig. S19 continued.

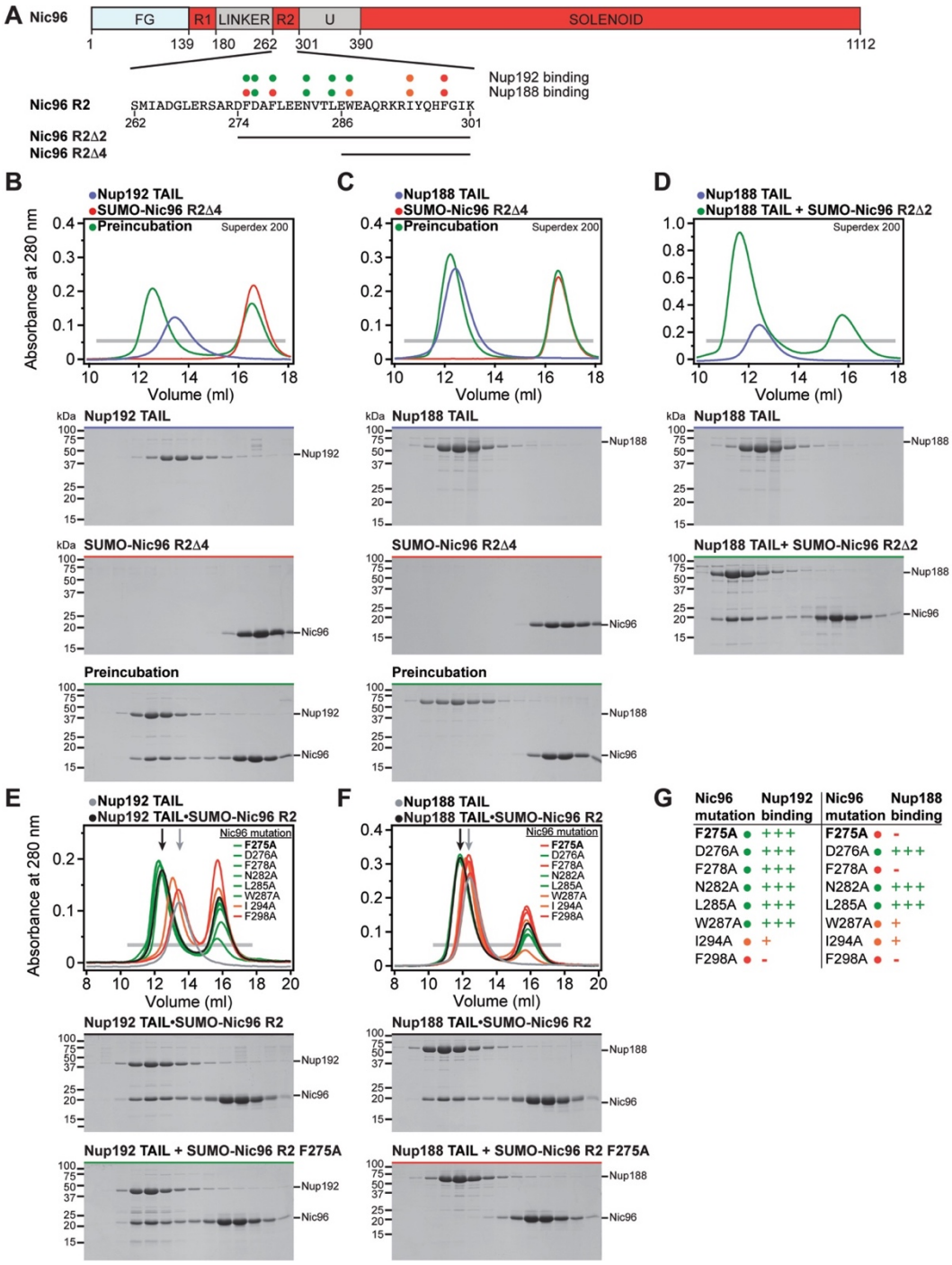


Fig. S20.
Nup192^{TAIL} and Nup188^{TAIL} bind to partially overlapping sites in Nic96^{R2}. (**A**) Domain boundaries of Nic96 with the protein sequence corresponding to the R2 region shown below. Black lines indicate construct boundaries for Nic96^{R2}, Nic96$^{R2\Delta2}$, and Nic96$^{R2\Delta4}$. The effects of Nic96^{R2} alanine mutations on Nup192^{TAIL} and Nup188^{TAIL} binding are indicated by colored dots; no effect (green), reduced binding (orange), and complete disruption (red). (**B-D**) SEC and SDS-PAGE analysis corresponding to Fig. 2F. SEC profiles of nucleoporins are

shown individually (blue and red) and after their preincubation (green). Nic96$^{R2\Delta2}$ is insoluble when expressed alone and therefore was expressed and purified in the presence of Nup188^{TAIL}. (**E**, **F**) Mutational analysis of the Nic96^{R2} interaction with Nup192^{TAIL} and Nup188^{TAIL}. SEC profiles of Nup192^{TAIL} or Nup188^{TAIL} are shown individually (gray) and after preincubation with wild type Nic96^{R2} (black) and are annotated with arrows of the same color. SEC profiles of mutant proteins preincubated with Nic96^{R2} are colored according to measured effect as in panel (A). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. (**G**) Table summarizing the measured effects of Nic96^{R2} mutation on Nup192^{TAIL} and Nup188^{TAIL} binding colored according to panel (A): no effect (+++), moderate effect (+), abolished binding (-). The Nic96^{R2} F275A mutant abolished binding to Nup188^{TAIL} but not to Nup192^{TAIL}.

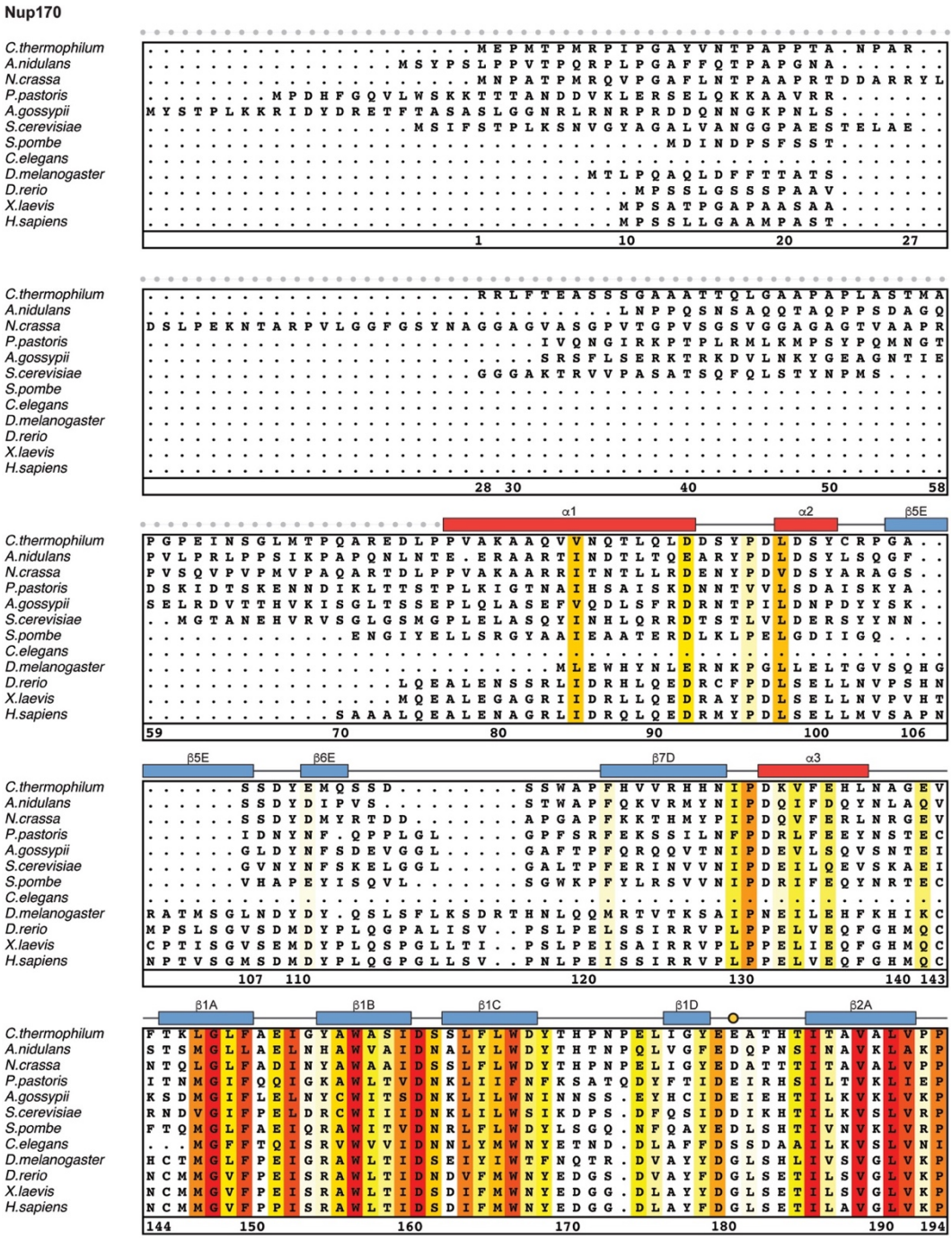
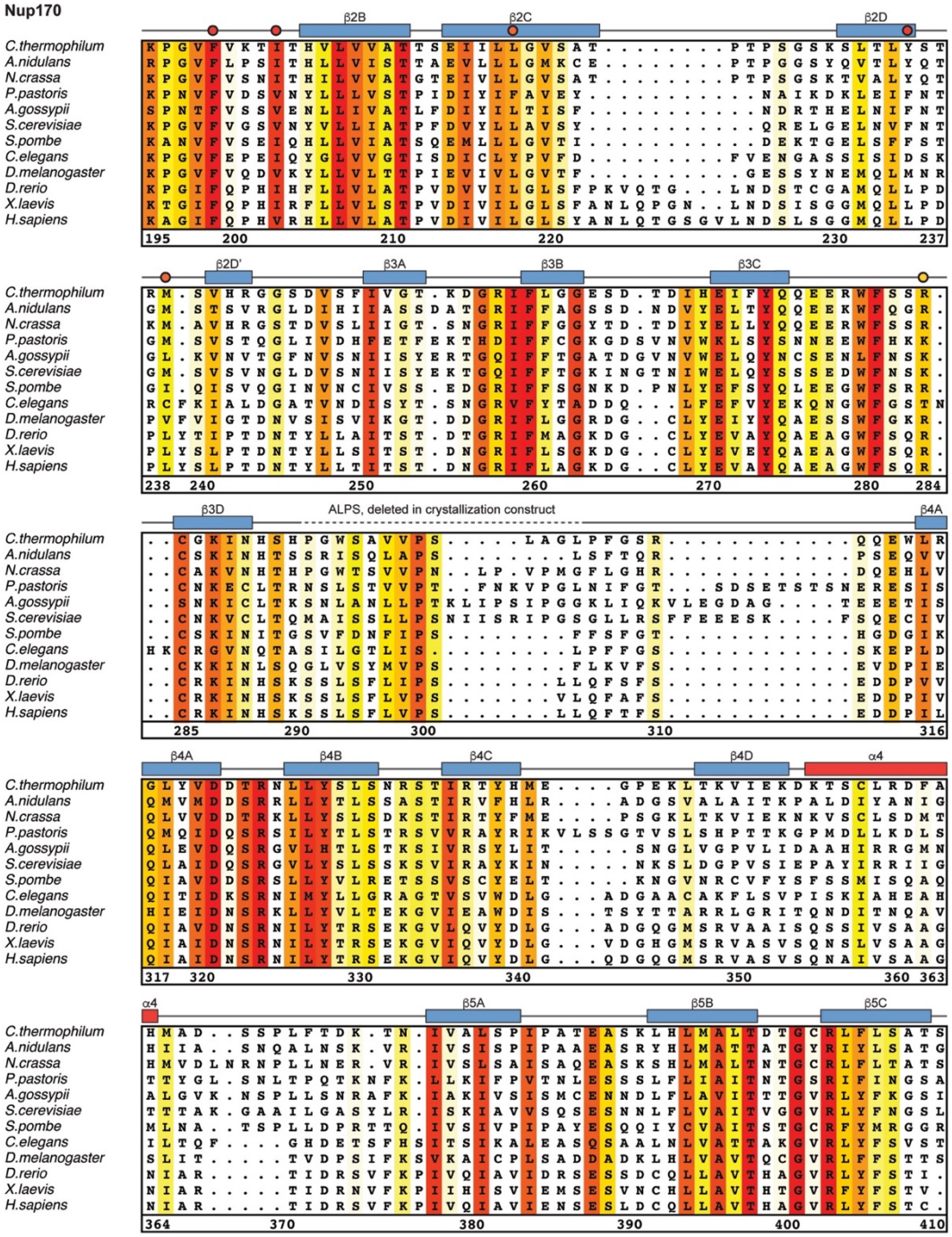
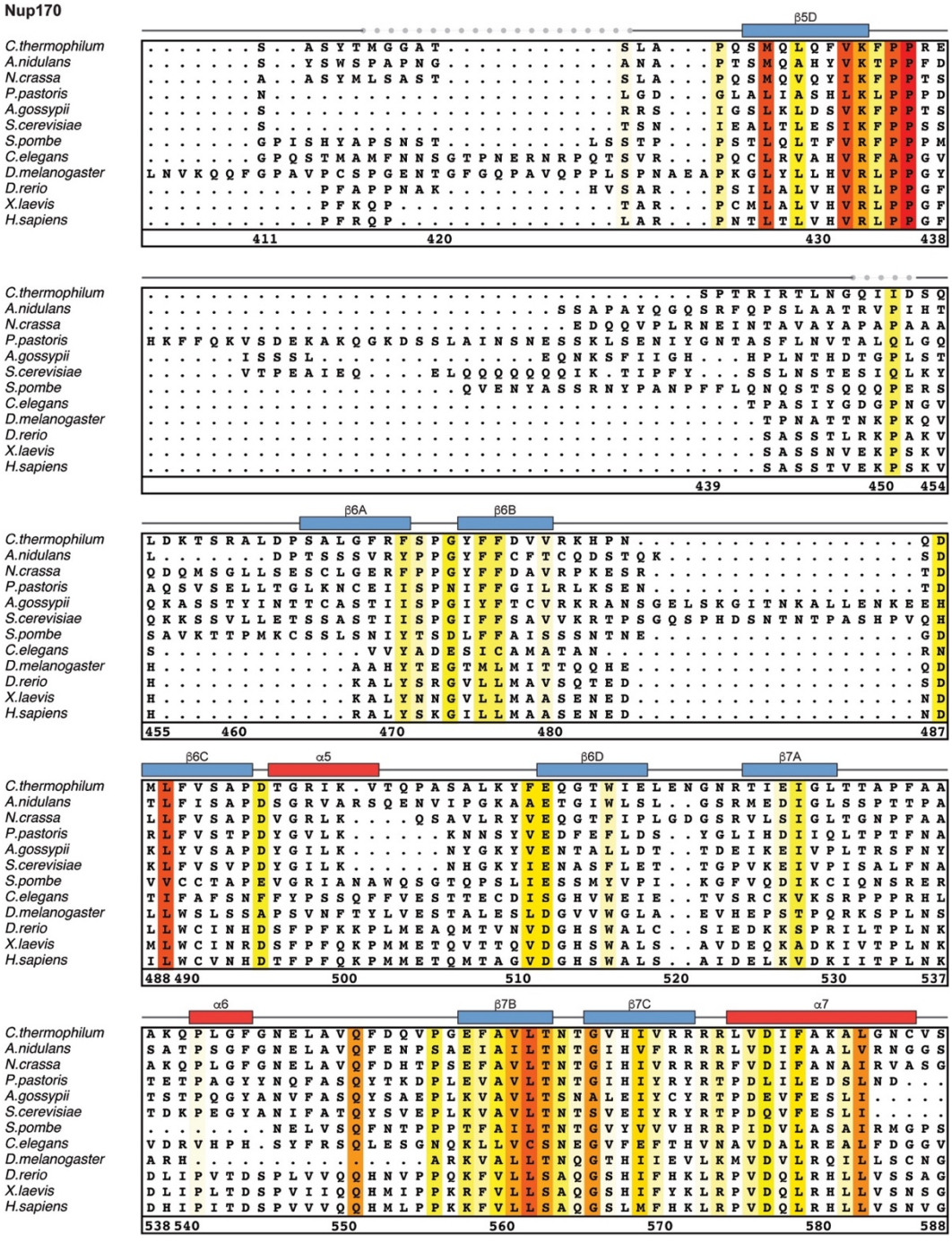


Fig. S21.

Multispecies sequence alignment of Nup170. Sequences from twelve diverse species were aligned and colored by sequence similarity according to the BLOSUM62 matrix from white (less than 55 % similarity), to yellow (55 % similarity), to red (100 % identity). Numbering below alignment is relative to the *C. thermophilum* sequence. Secondary structure observed in the Nup170 structures is shown above the alignment: α-helices (red bars), β-sheets (blue bars), and unstructured regions (black lines). Mutations that affect Nup53 or Nup145N binding identified by a mutational analysis (Fig. 3; figs. S22; fig. S23) are indicated by circles



above the alignment and colored according to the measured effect; weak effect (yellow), moderate effect (orange), abolished binding (red). Dashed lines indicate loops that were deleted in crystallization constructs. Disordered regions are indicated by gray dots.


Nup170
β5D
C.thermophilum
A.nidulans
N.crassa
P.pastoris
A.gossypii
S.cerevisiae
S.pombe
C.elegans
D.melanogaster
D.rerio
X.laevis
H.sapiens
411
420
430
438
439
450
454
β6A
β6B
455
460
470
480
487
β6C
α5
β6D
β7A
488 490
500
510
520
530
537
α6
β7B
β7C
α7
538 540
550
560
570
580
588

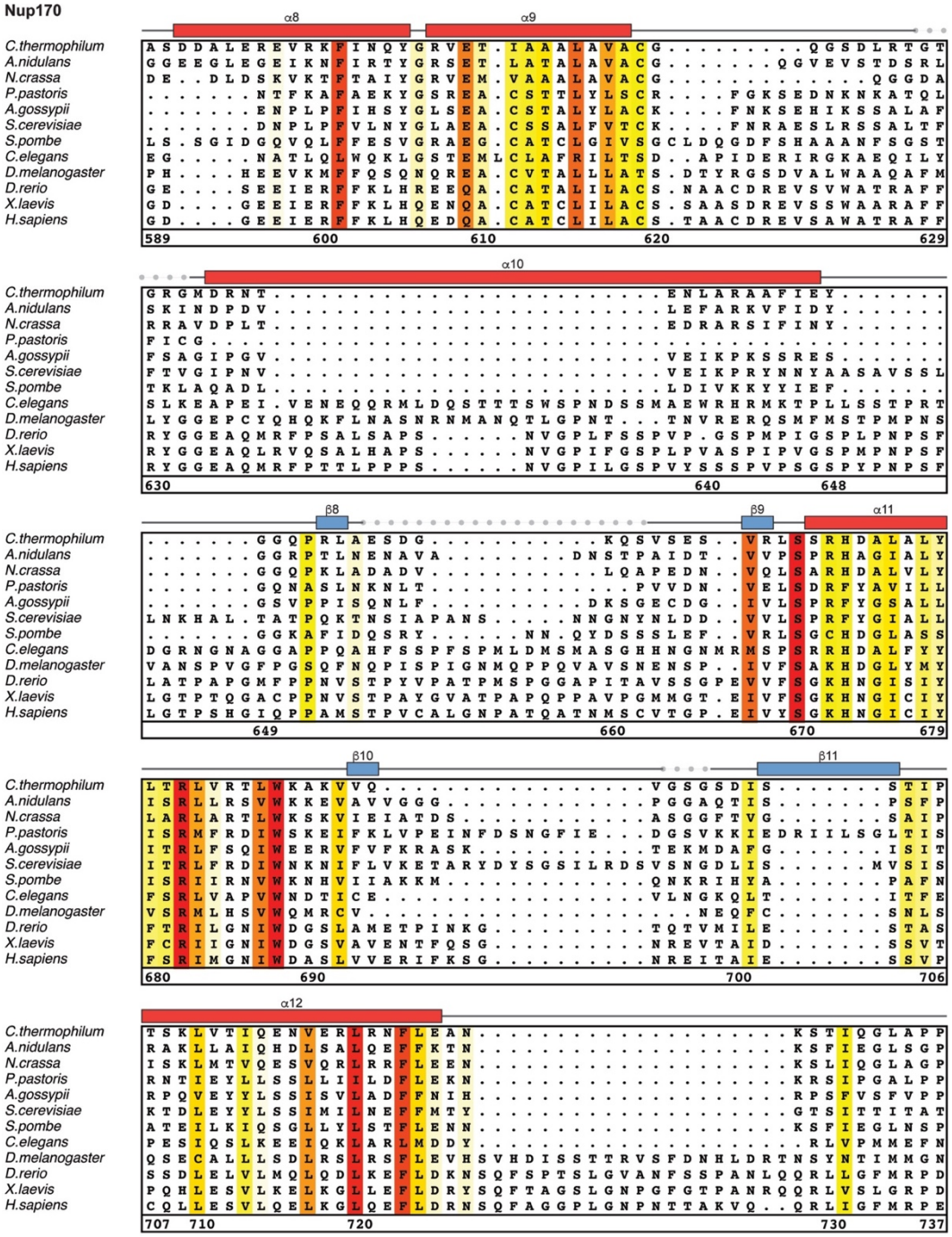


Fig. S21 continued.

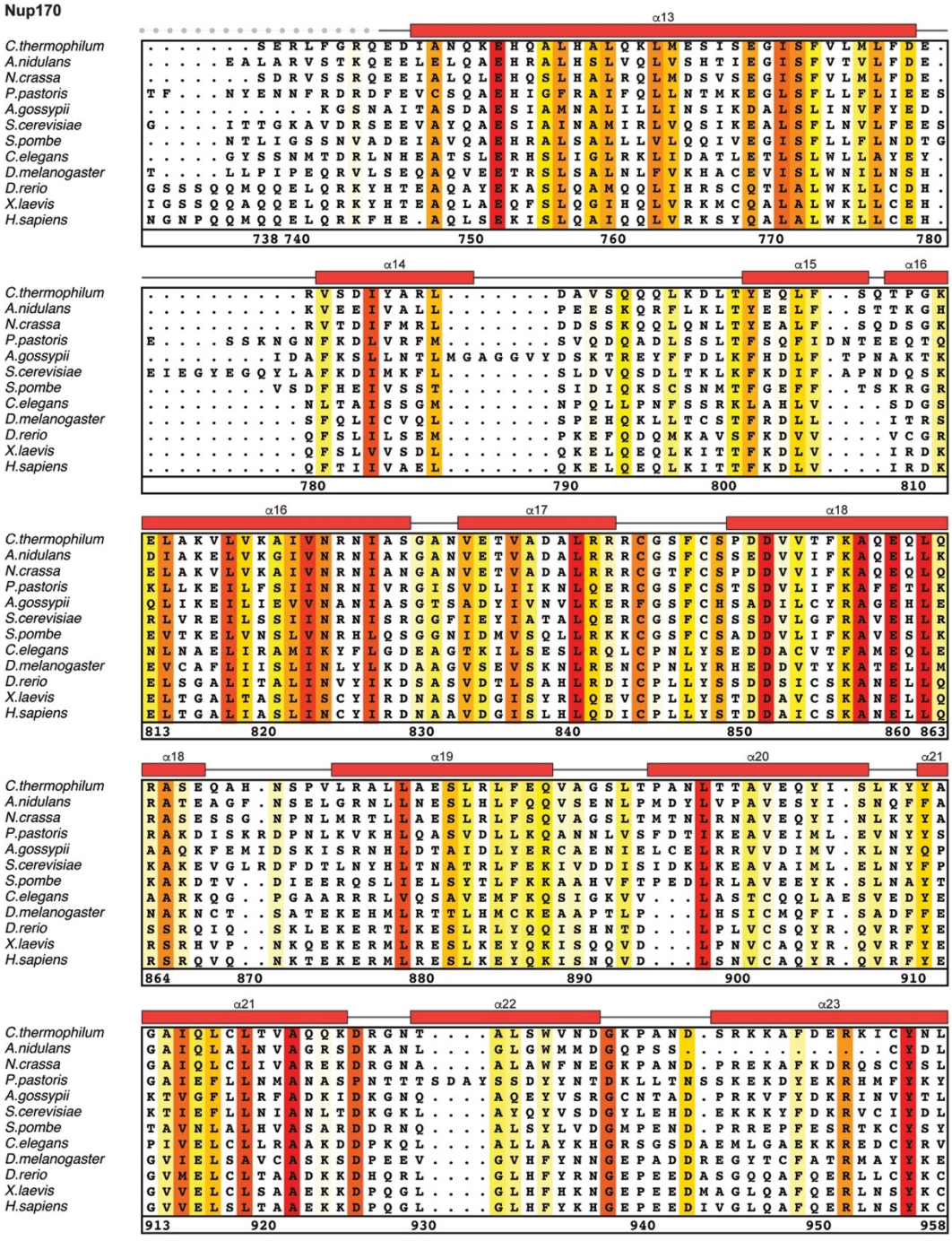


Fig. S21 continued.

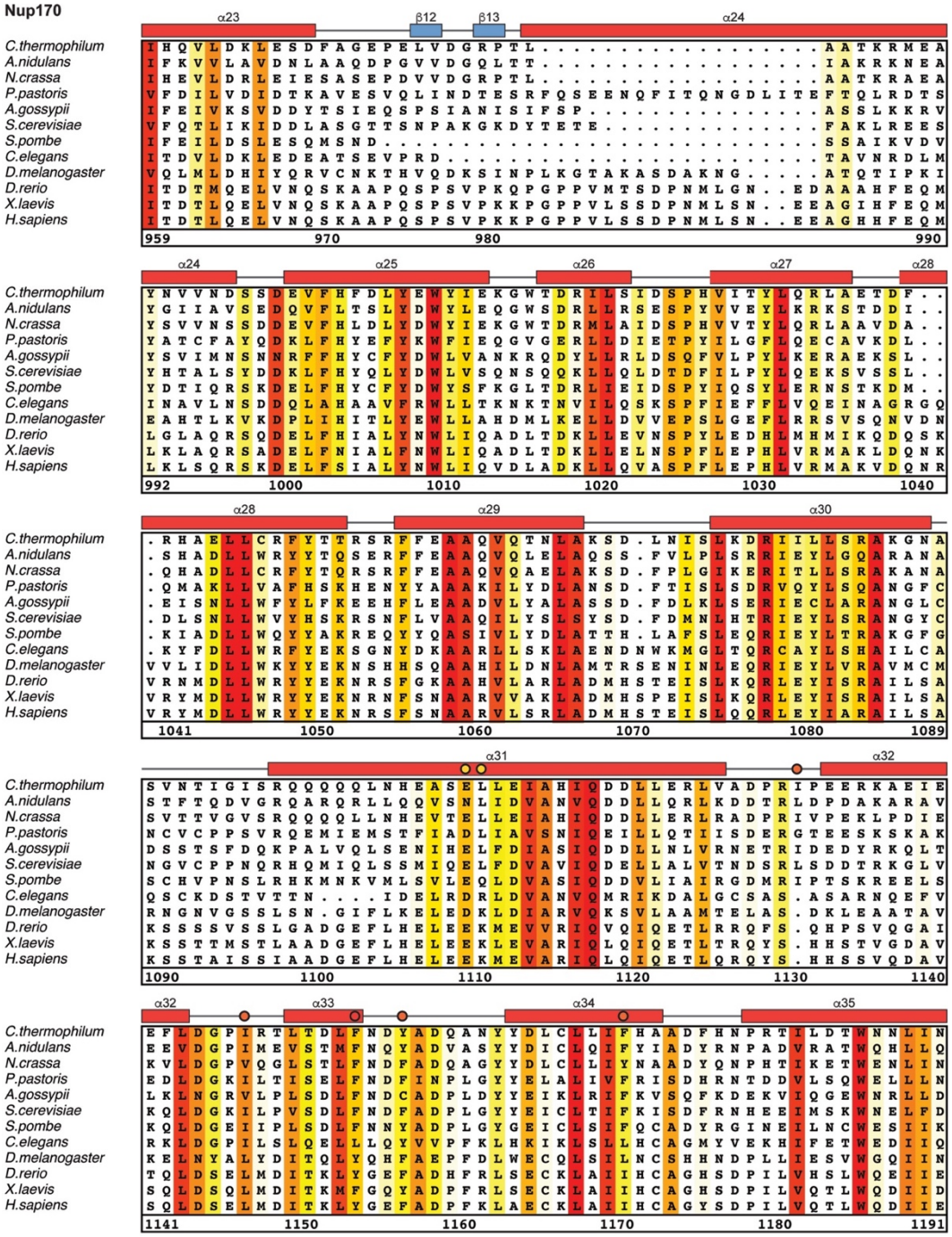


Fig. S21 continued.

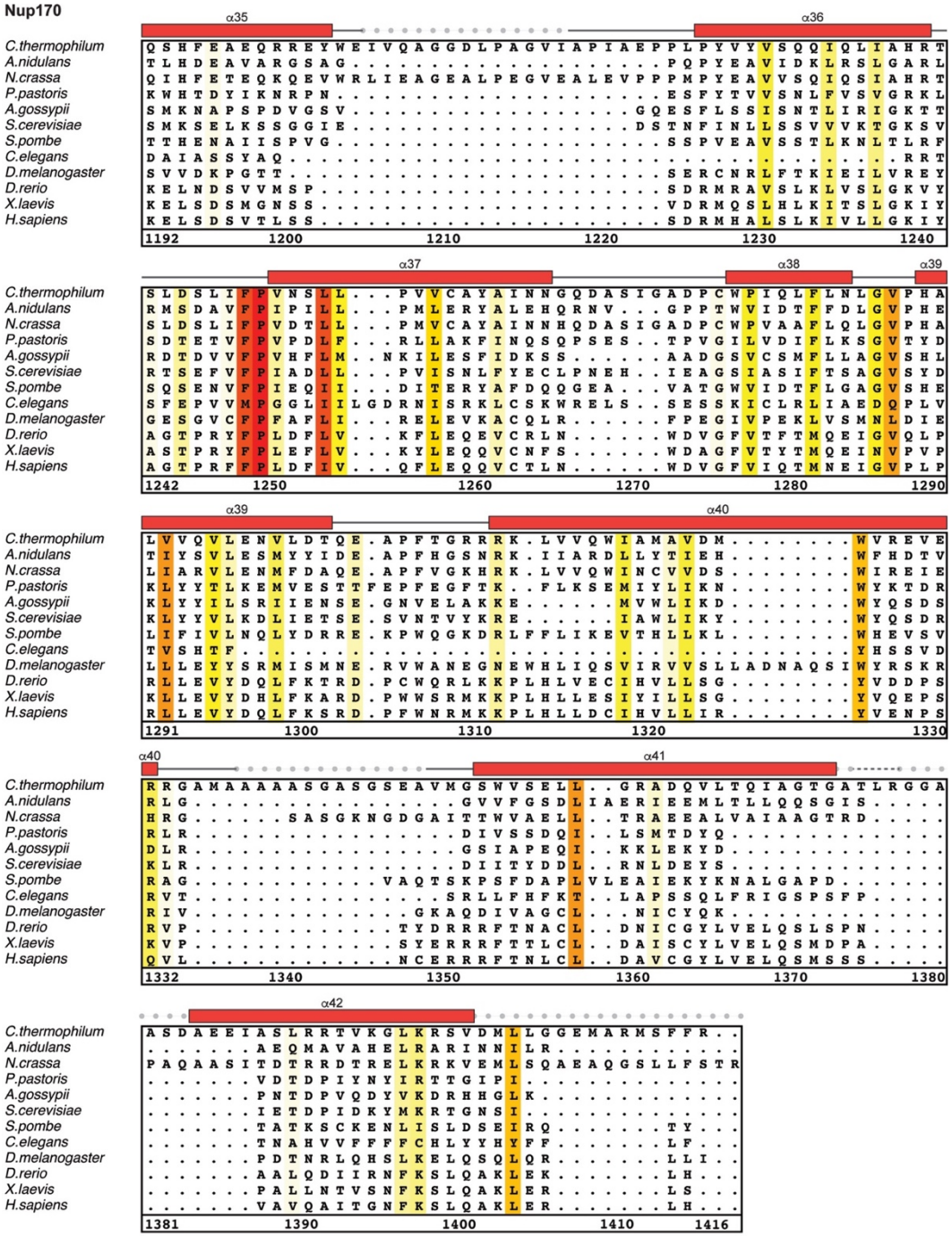


Fig. S21 continued.

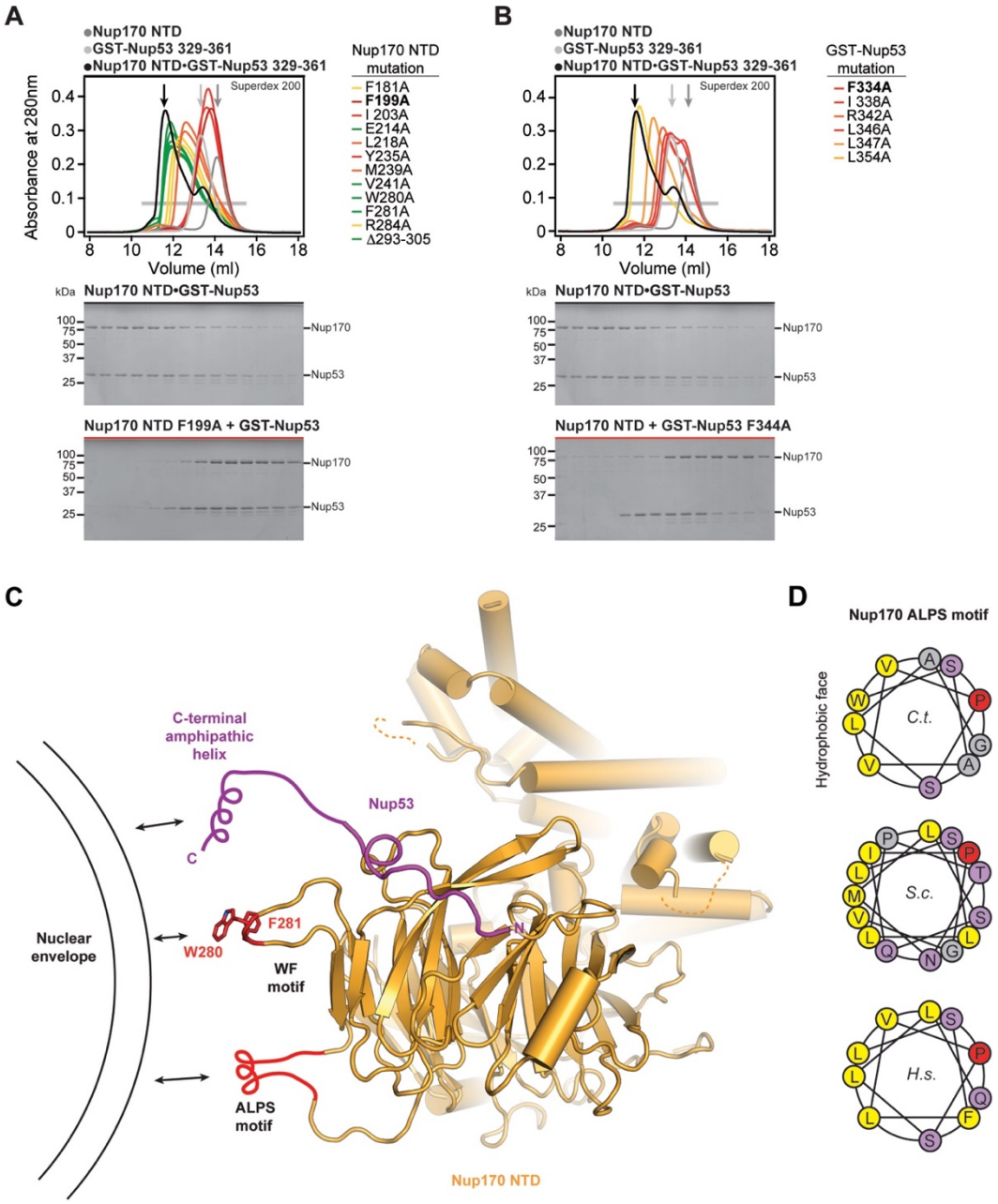


Fig. S22.

Mutational analysis of Nup170^{NTD}•Nup53 and proposed interactions with the nuclear envelope. SEC interaction experiments performed with mutants of (**A**) Nup170^{NTD} and (**B**) GST-Nup53$^{329-361}$. SEC profiles of wild type Nup170^{NTD} (dark gray) and GST-Nup53$^{329-361}$ (light gray) are shown individually and after preincubation (black) and are annotated with arrows of the same color. Mutant SEC profiles are colored according to the measured effect: no effect (green), weak effect (yellow), moderate effect (orange), and abolished binding (red). All SEC profiles were obtained using a Superdex 200 10/300 GL column. (**C**) Nup170^{NTD} is shown in cartoon representation and possible membrane interaction motifs are highlighted. The Nup170 β-propeller domain contains two sequence motifs, WF and ALPS, which are located on the same face directly adjacent to a C-terminal amphipathic helix of

Nup53 that has previously been shown to anchor Nup53 to the nuclear envelope (*18, 19*). (**D**) Helical wheel diagrams of the Nup170 ALPS motif in *C. thermophilum*, *S. cerevisiae*, and *H. sapiens* are shown with hydrophobic and polar residues colored in yellow and purple, respectively. The universally conserved proline residue on the polar face of the helix, a feature reminiscent of antimicrobial membrane-destabilizing peptides, is colored in red (*34*).

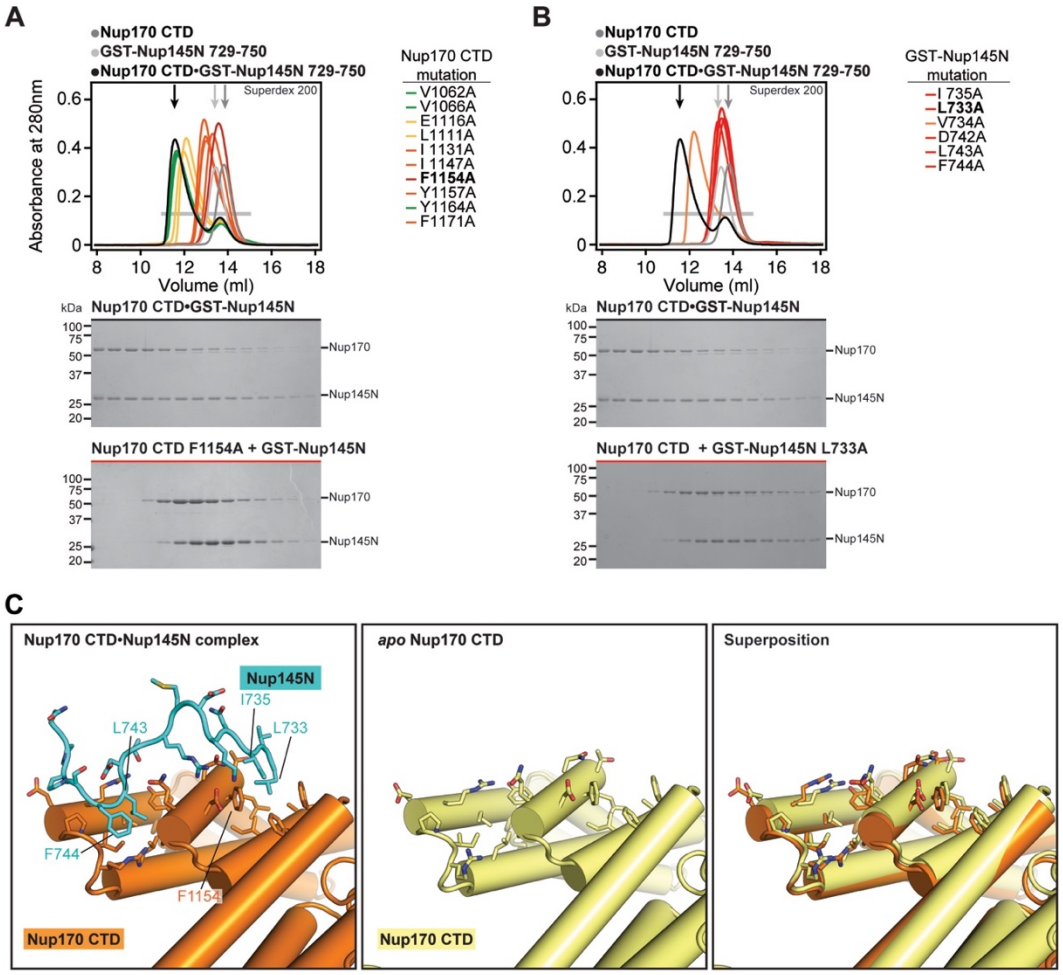


Fig. S23.

Mutational and structural analyses of the Nup170^{CTD}-Nup145N interaction. SEC interaction experiments performed with mutants of (**A**) Nup170^{CTD} and (**B**) GST-Nup145N$^{729-750}$. SEC profiles of wild type Nup170^{CTD} (dark gray) and GST-Nup145N$^{729-750}$ (light gray) are shown individually and after preincubation (black) and are annotated with arrows of the same color. Mutant SEC profiles colored according to measured effect: no effect (green), weak effect (yellow), moderate effect (orange), and abolished binding (red). All SEC profiles were obtained using a Superdex 200 10/300 GL column. (**C**) Comparison of the Nup145N binding pocket in Nup170 in the bound and *apo* states reveals minimal conformational changes upon binding.

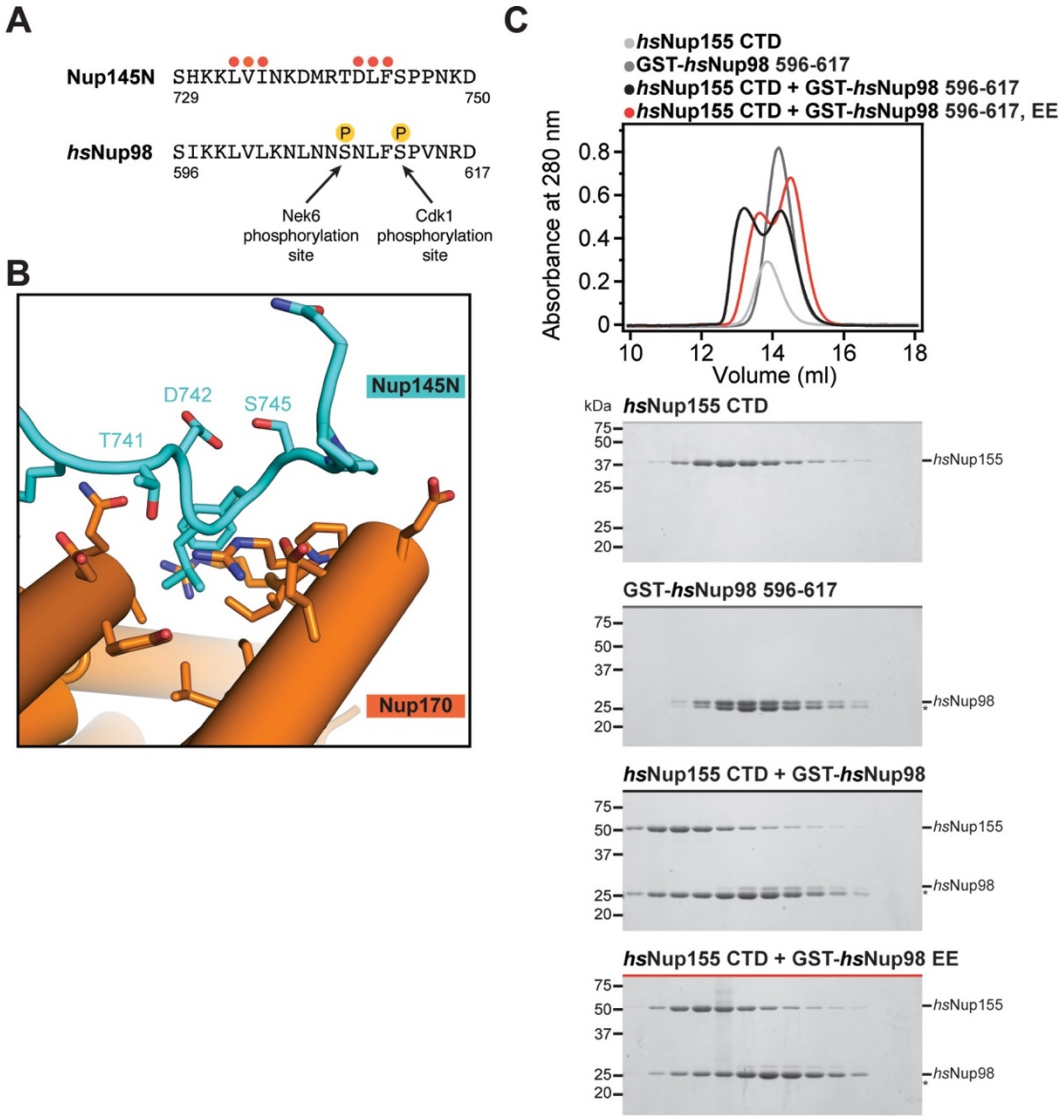


Fig. S24.
The Nup170^{CTD}-Nup145N interaction is conserved in humans and partially disrupted by phosphomimetic mutations. (**A**) Sequence alignment of Nup145N$^{729-750}$ and *hs*Nup98$^{596-617}$. Nup145N$^{729-750}$ mutants are indicated by circles and colored as in Fig. 3J. *hs*Nup98 residues S608 and S612 that are phosphorylated during mitosis are indicated (*35*). (**B**) Close-up view of the Nup170^{CTD}-Nup145N interaction indicating the positioning of T741 and S745 in Nup145N, which correspond to the mitotically phosphorylated *hs*Nup98 residues S608 and S612. (**C**) SEC interaction analysis of the *hs*Nup155^{CTD}-*hs*Nup98$^{596-617}$ interaction. SEC profiles of *hs*Nup155^{CTD} (light gray) and *hs*Nup98$^{596-617}$ (dark grey), the preincubation of the wild type proteins (black), and the preincubation of *hs*Nup155^{CTD} with the phosphomimetic *hs*Nup98$^{596-617}$ S608E/S612E double mutant (red) are shown. All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

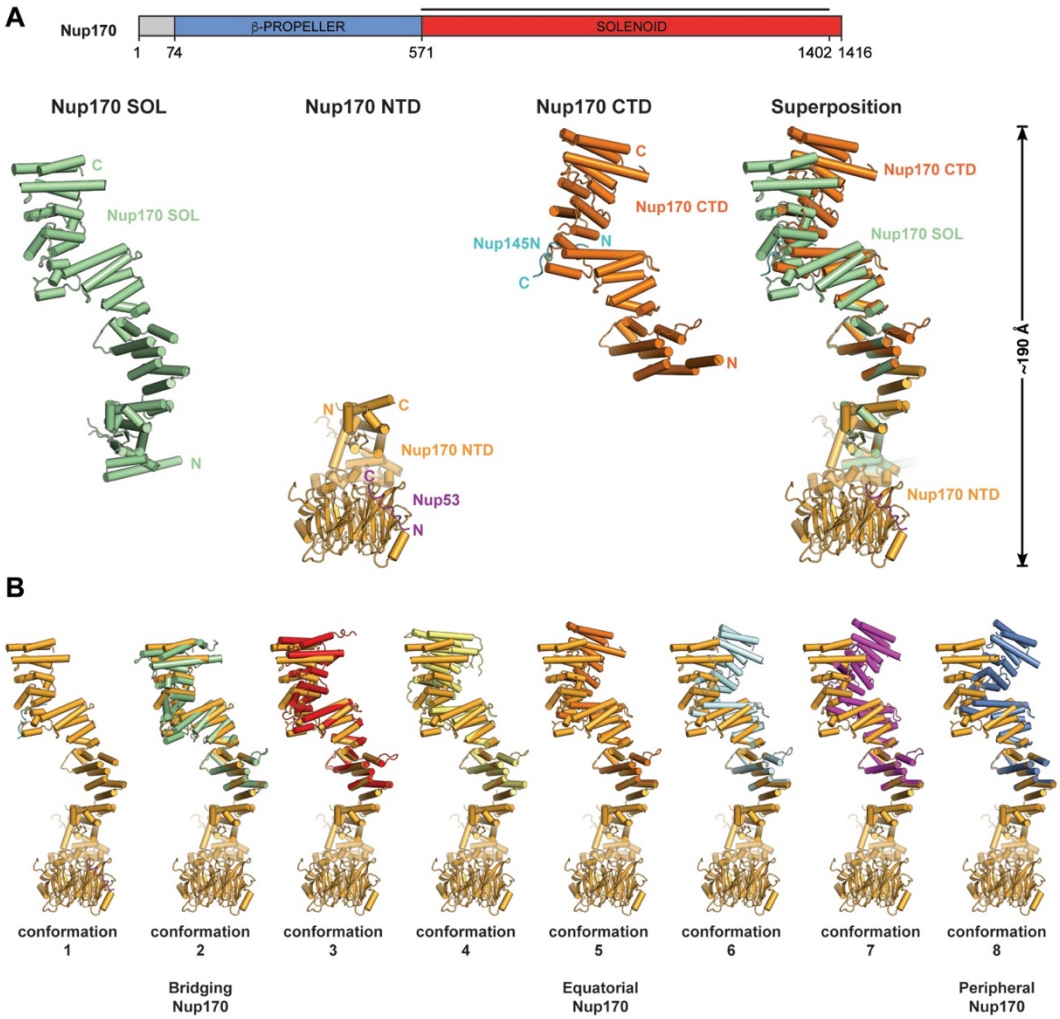


Fig. S25.

Superposition-generated structures of full-length Nup170. (**A**) The construct boundaries of Nup170^{SOL} are indicated by a black line above the domain structure. The crystal structures of Nup170^{SOL} (green), Nup170^{NTD} (light orange), and Nup170^{CTD} (dark orange), and their superposition are shown in cartoon representation. (**B**) Complete conformational range of Nup170 generated by superposing different conformations of Nup170^{CTD} with the structure of Nup170^{SOL}. Conformations were obtained from the Nup170^{SOL} structure (conformation 1), different molecules in the asymmetric unit of the *apo* Nup170^{CTD} structures (conformations 2, 3, and 4), and different molecules in the asymmetric unit of the Nup170^{CTD}•Nup145N complex structure (conformations 5, 6, 7, and 8). The conformations that were docked into cryoET reconstruction of the NPC are indicated below.

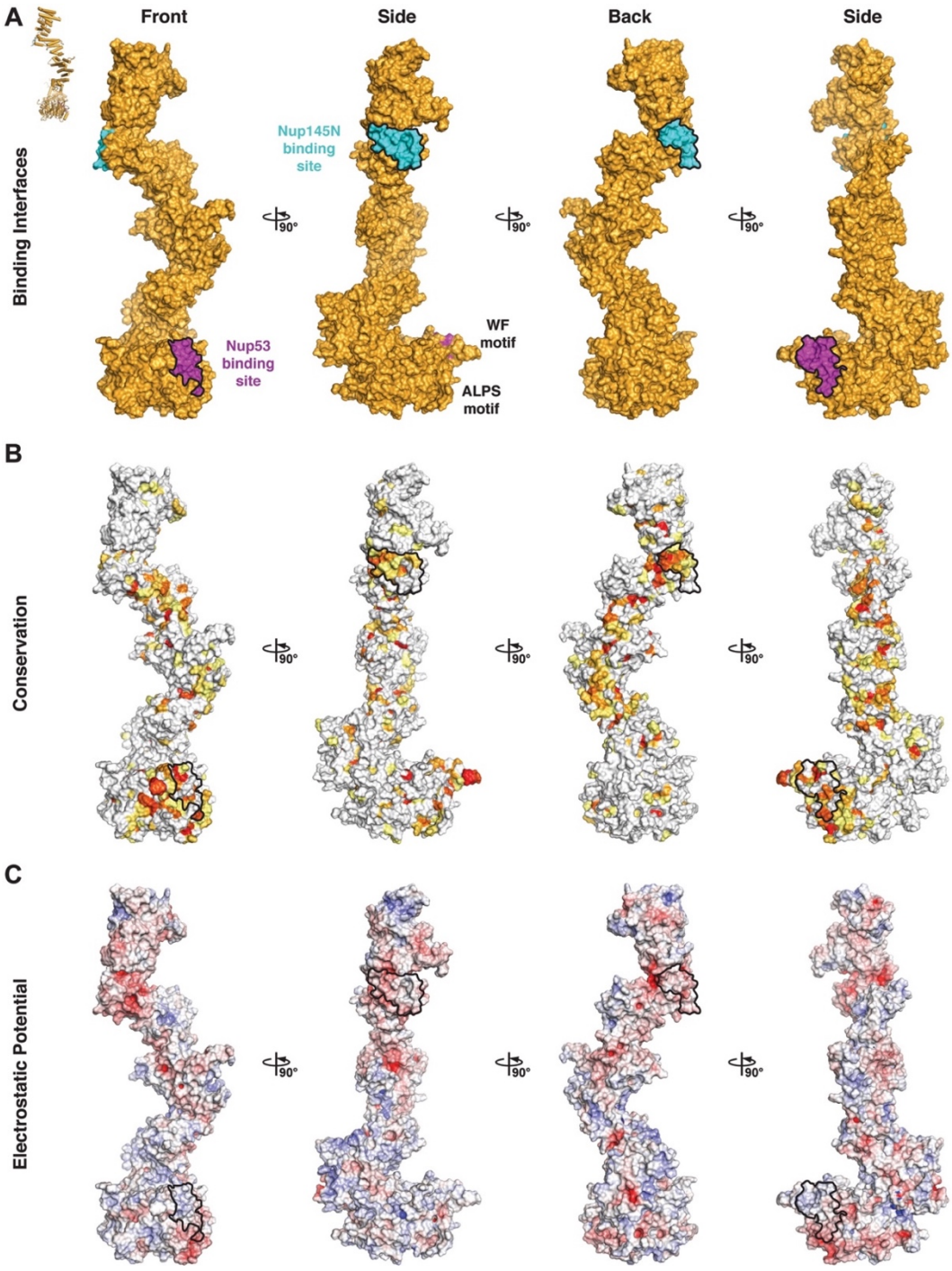


Fig. S26.
Surface properties of Nup170. Surface representations of Nup170 is shown in four different orientations related by 90° rotations. Nup53 and Nup145N binding interfaces are outlined in black. (**A**) Surface representation with the Nup53 and Nup145N binding sites colored in purple and cyan, respectively. (**B**) Surface representation colored according to sequence identity based on the alignment in fig. S21. (**C**) Surface representation colored according to electrostatic potential from -10 k_BT/e (red) to +10 k_BT/e (blue).

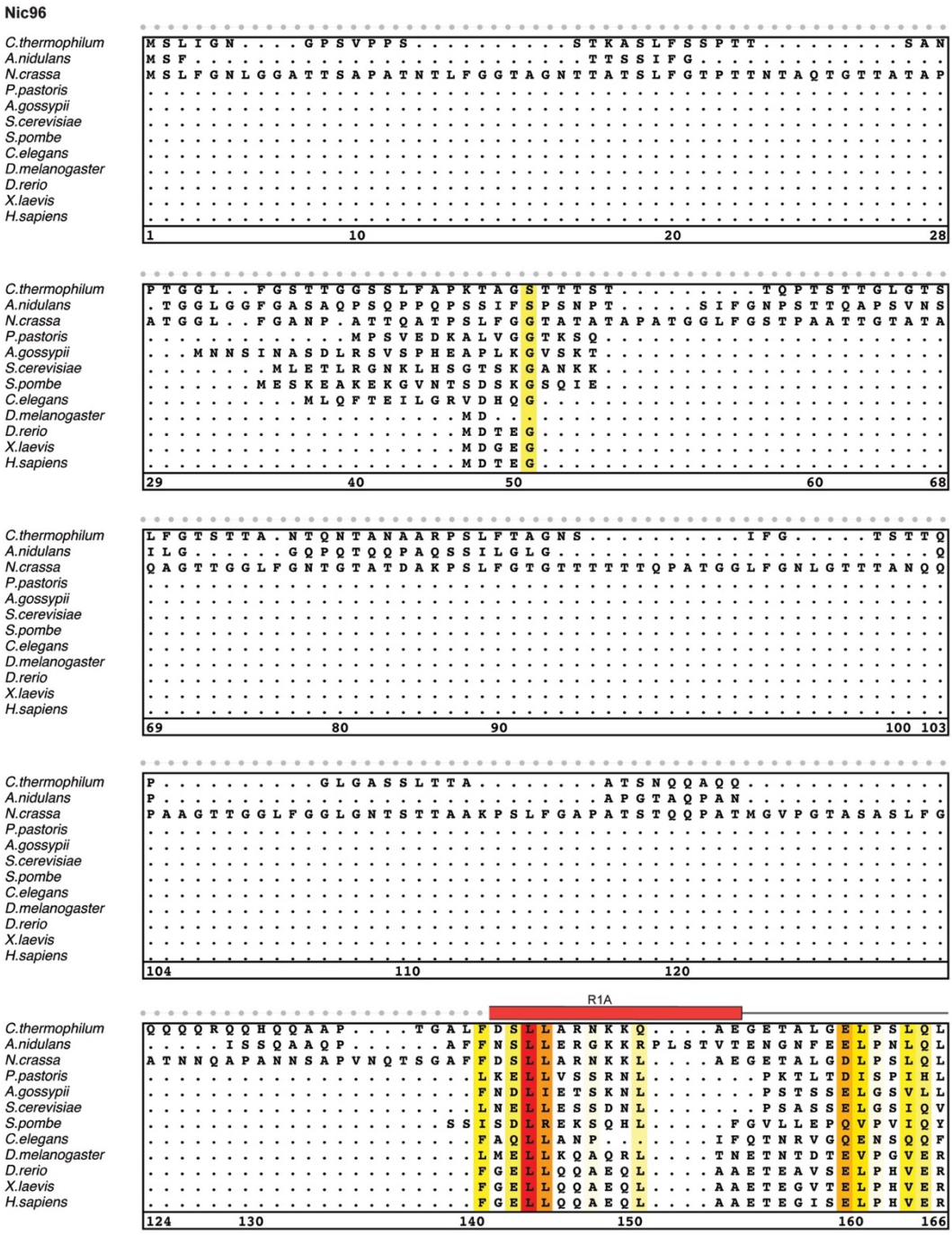
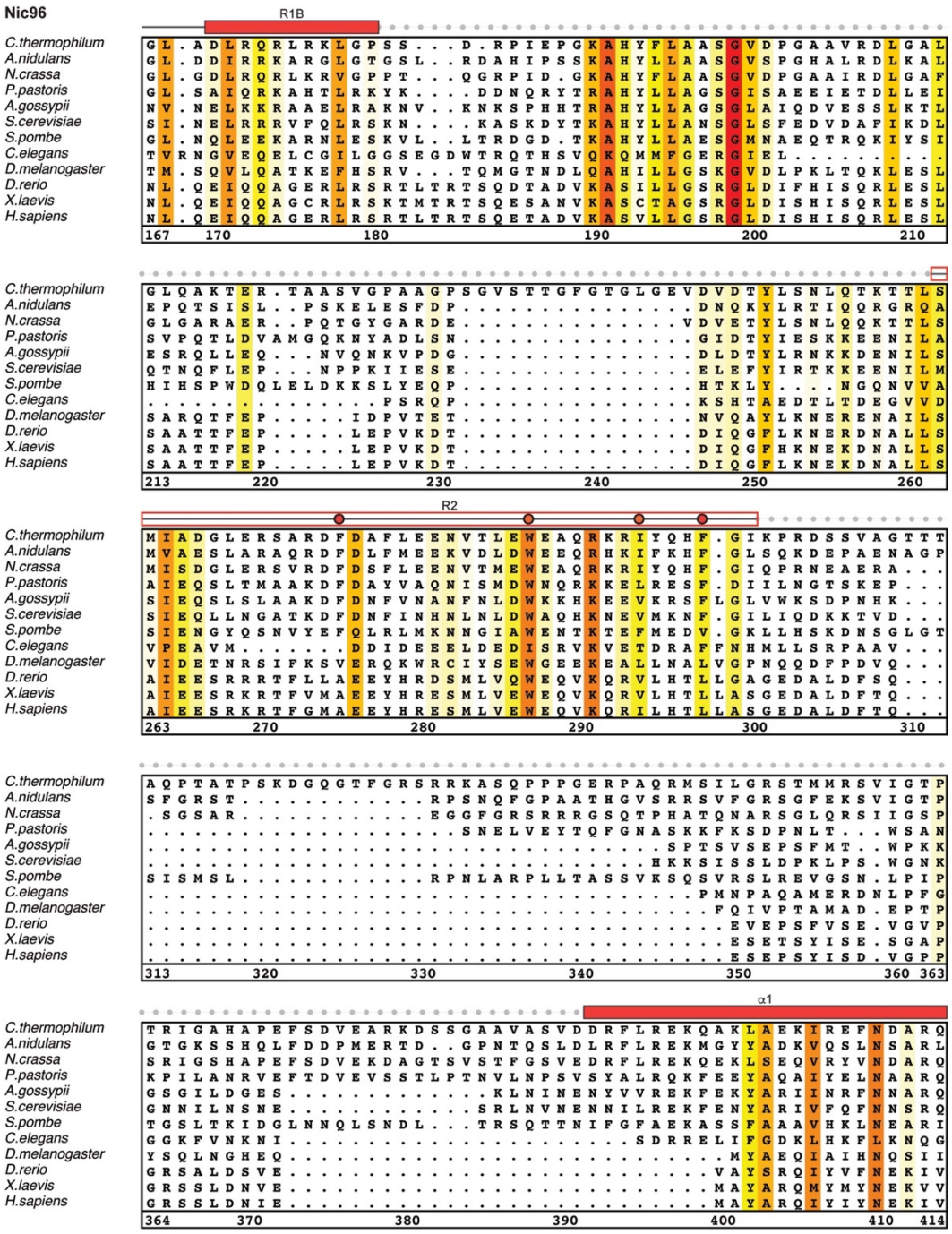
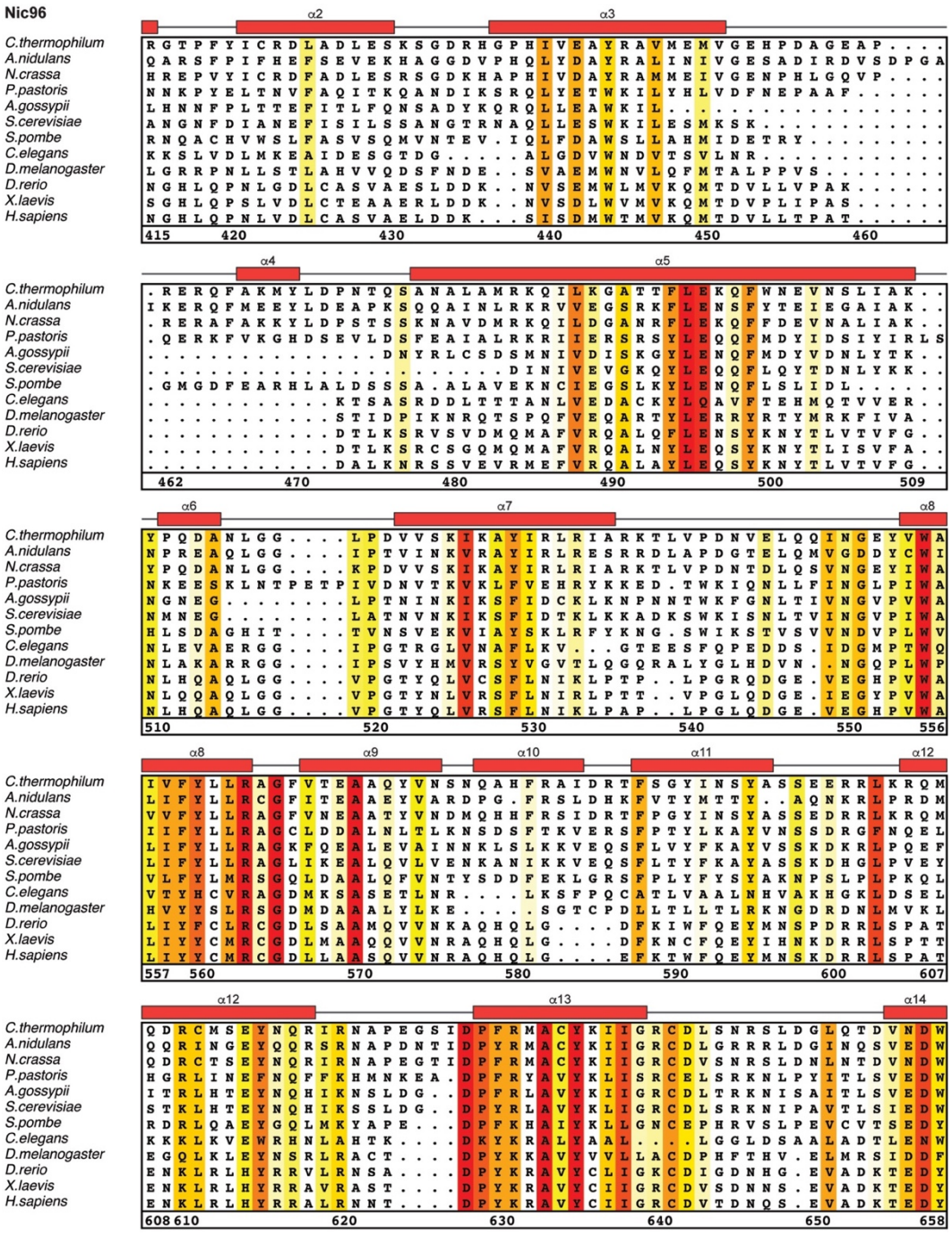


Fig. S27.

Multispecies sequence alignment of Nic96^{SOL}. Sequences from twelve diverse species were aligned and colored by sequence similarity according to the BLOSUM62 matrix from white (less than 55 % similarity), to yellow (55 % similarity), to red (100 % identity). Numbering below alignment is relative to the *C. thermophilum* sequence. Secondary structure observed in the Nic96 structure is shown above the alignment: α-helices (red bars), β-sheets (blue bars), and unstructured regions (black lines). The secondary structure of R2 is unknown, but predicted to be helical, which is indicated by a red outline. Mutations that affect binding to



Nup53, Nup188, or Nup192 found by mutational analysis (figs. S20 and S28) are indicated by circles above the alignment and colored according to the measured effect; weak effect (yellow), moderate effect (orange), and abolished binding (red). Disordered regions are indicated by gray dots.


Nic96
α2
α3
C.thermophilum
A.nidulans
N.crassa
P.pastoris
A.gossypii
S.cerevisiae
S.pombe
C.elegans
D.melanogaster
D.rerio
X.laevis
H.sapiens
415 420 430 440 450 460
α4
α5
462 470 480 490 500 509
α6
α7
α8
510 520 530 540 550 556
α8
α9
α10
α11
α12
557 560 570 580 590 600 607
α12
α13
α14
608 610 620 630 640 650 658

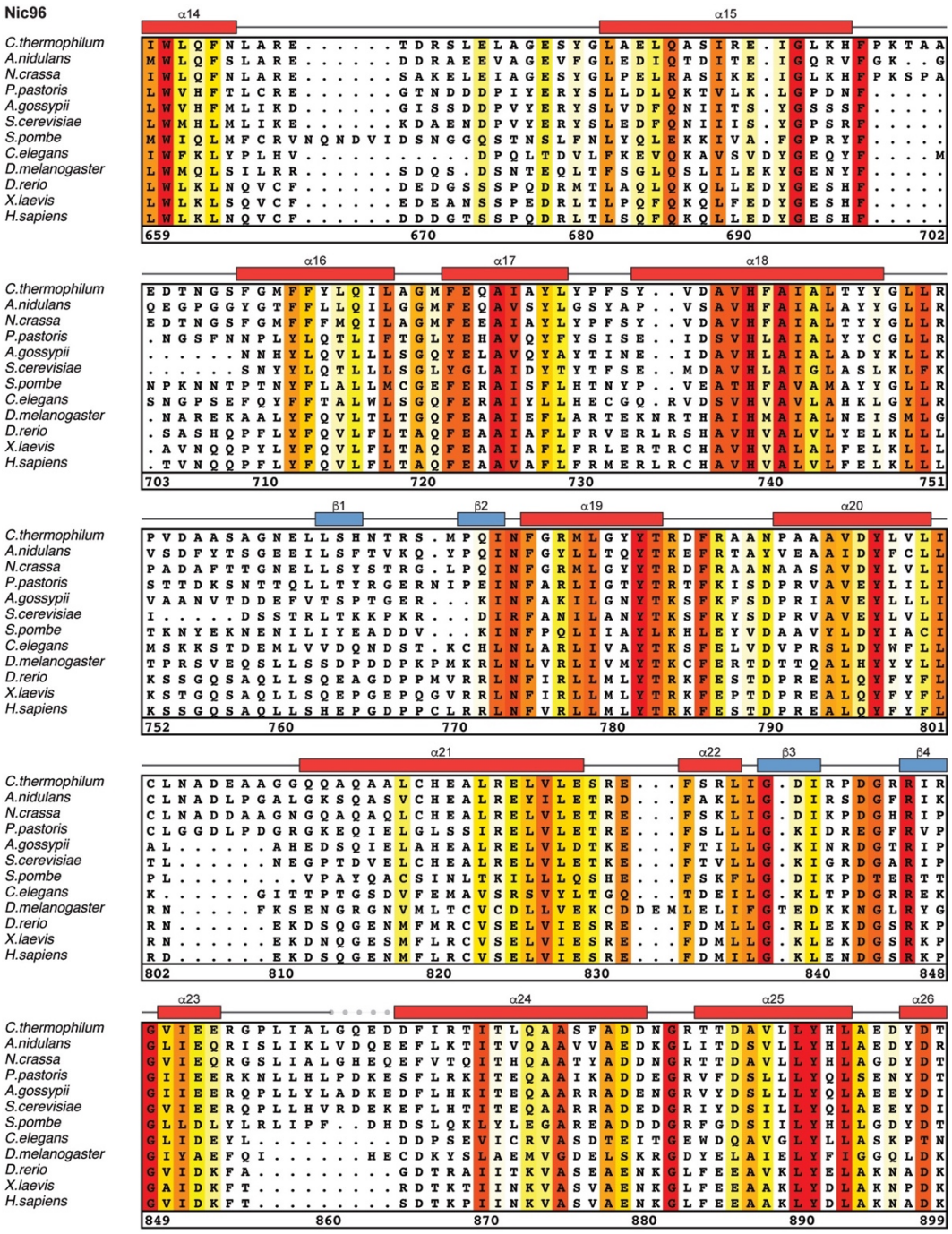


Fig. S27 continued.

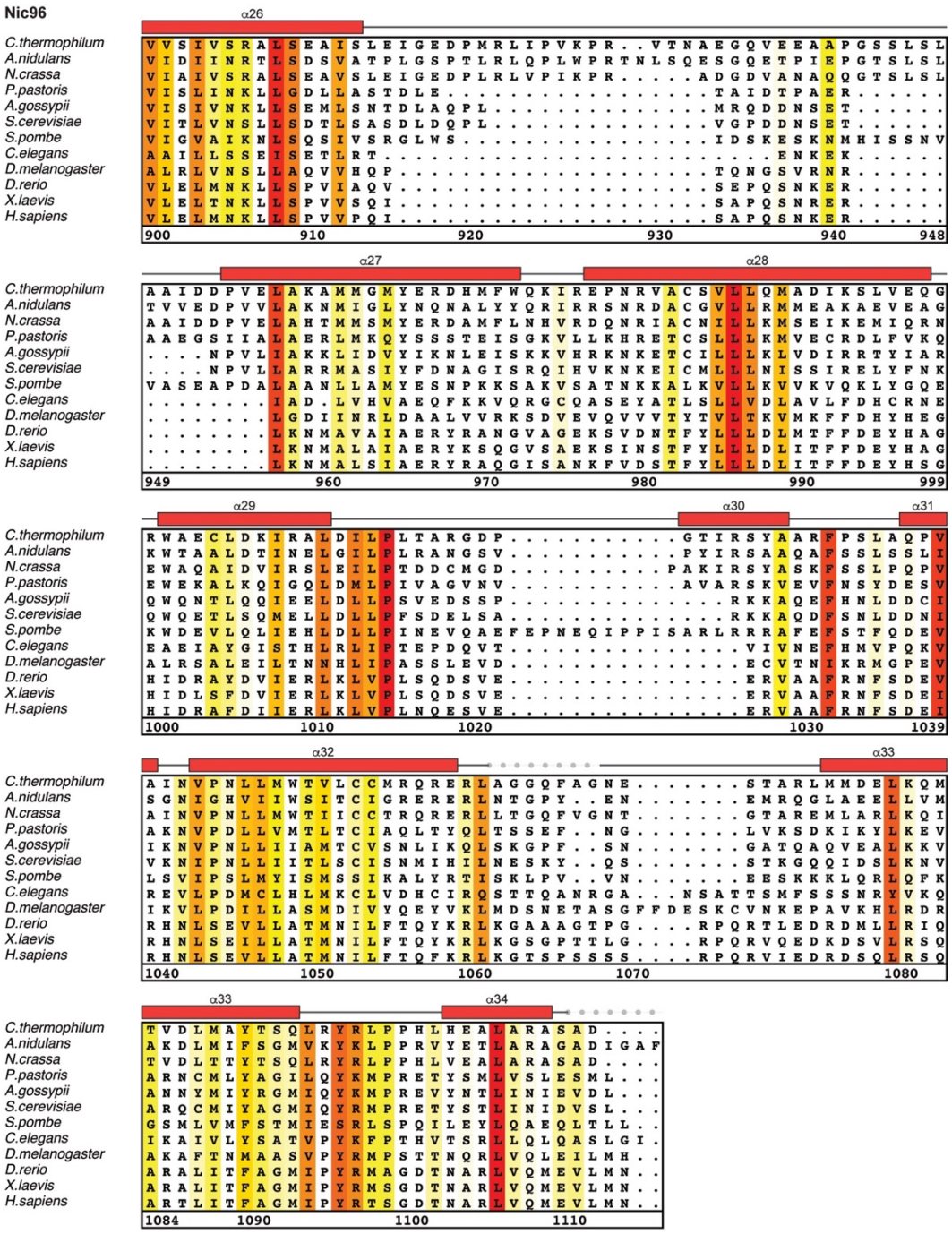


Fig. S27 continued.

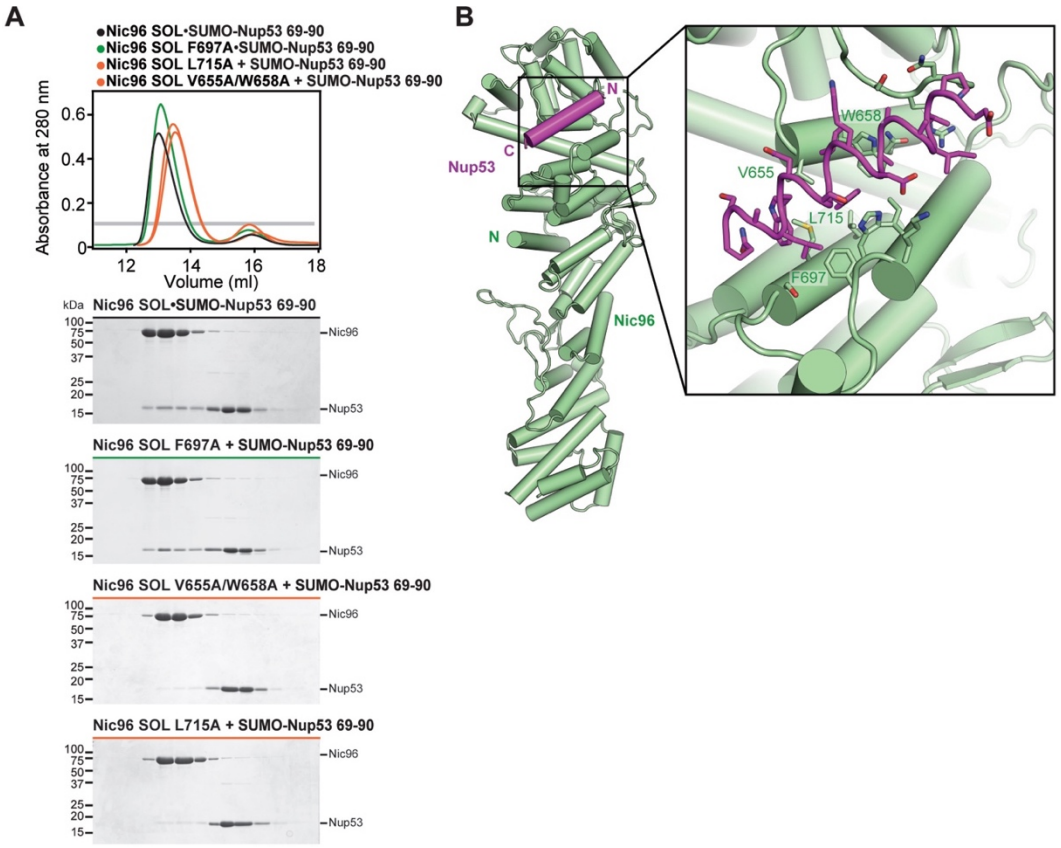


Fig. S28.

Identification of Nic96^{SOL} mutants that disrupt Nup53 binding. (**A**) SEC interaction analysis between Nic96^{SOL} mutants and SUMO-Nup53^{R2}. SEC profiles are colored according to the effect on SUMO-Nup53^{R2} binding; no effect (green), moderate effect (orange), or abolished binding (red). For reference, the SEC profile of the wild type Nic96^{SOL}•SUMO-Nup53^{R2} hetero-dimer (black) is shown. All SEC profiles were obtained using a Superdex 200 10/300 GL column. A gray bar indicates fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. (**B**) The structure of Nic96^{SOL}•Nup53^{R2} is shown in cartoon representation. The inset illustrates the region that is expanded on the right with mutated residues labeled in green.

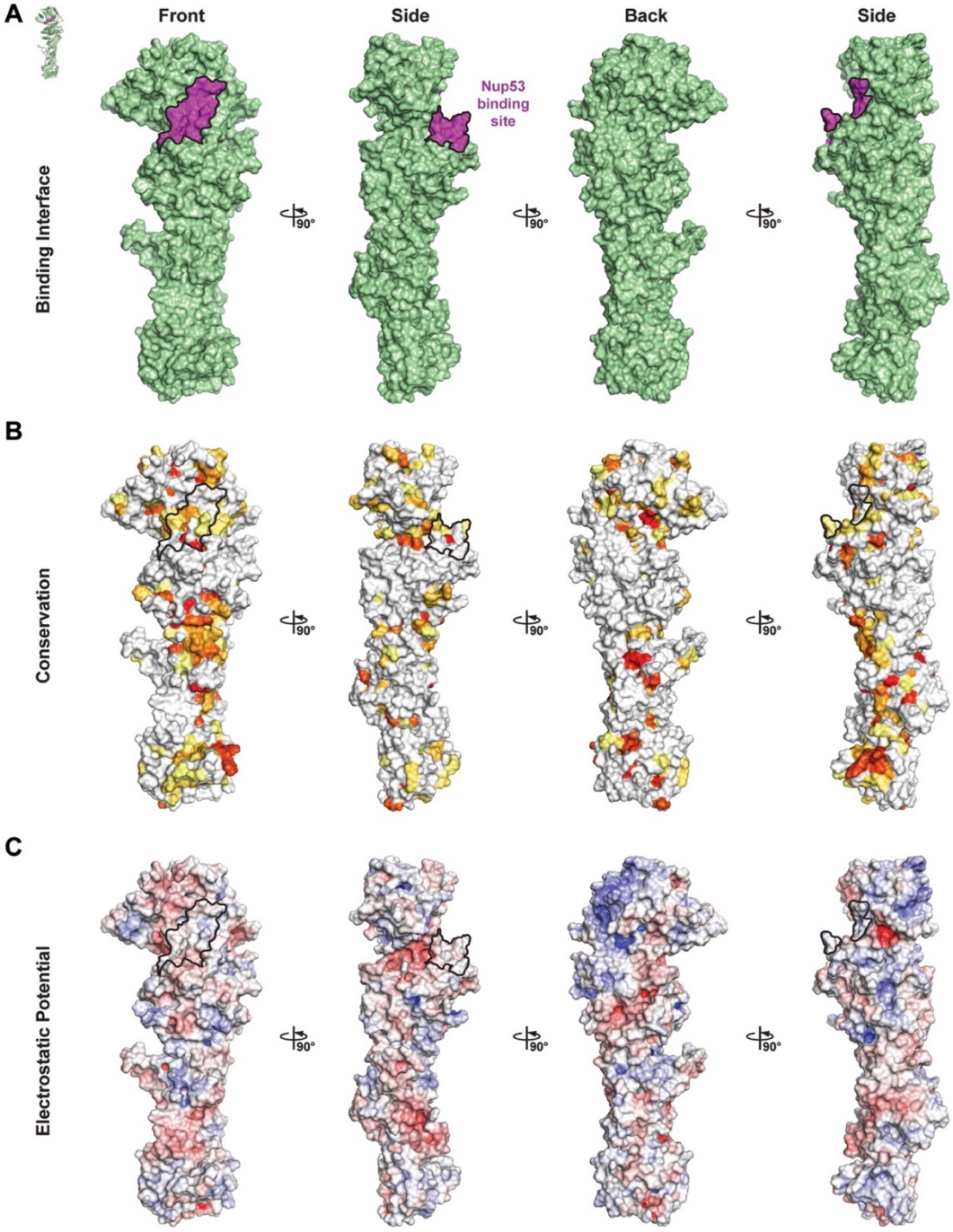


Fig. S29.
Surface properties of Nic96^{SOL}. Surface representations of Nic96^{SOL} are shown in four different orientations related by 90° rotations. The Nup53 binding interface is outlined in black. (**A**) Surface representation with the Nup53 binding site colored in purple. (**B**) Surface representation colored according to sequence identity based on the alignment in fig. S27. (**C**) Surface representation colored according to electrostatic potential from -10 k_BT/e (red) to +10 k_BT/e (blue).

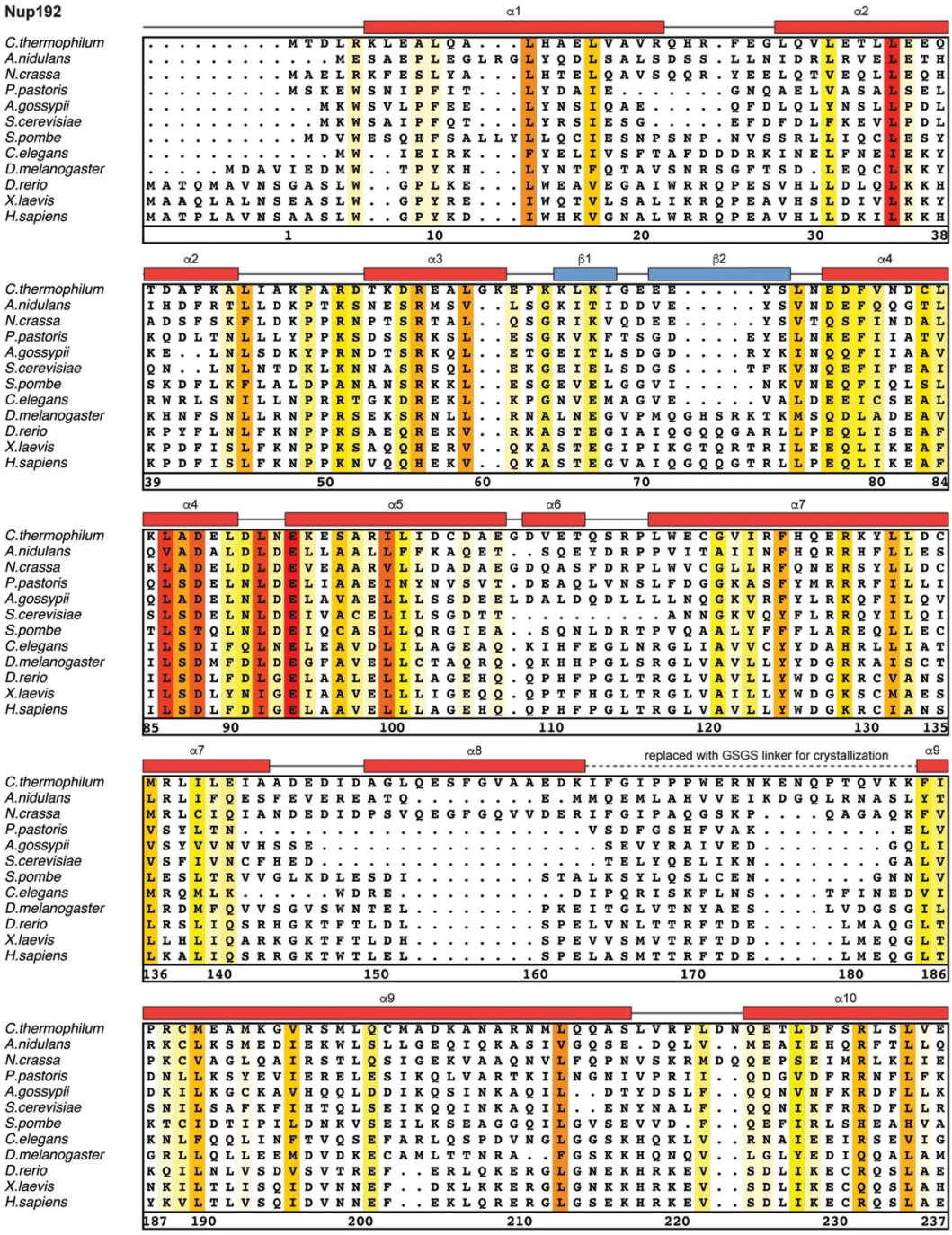
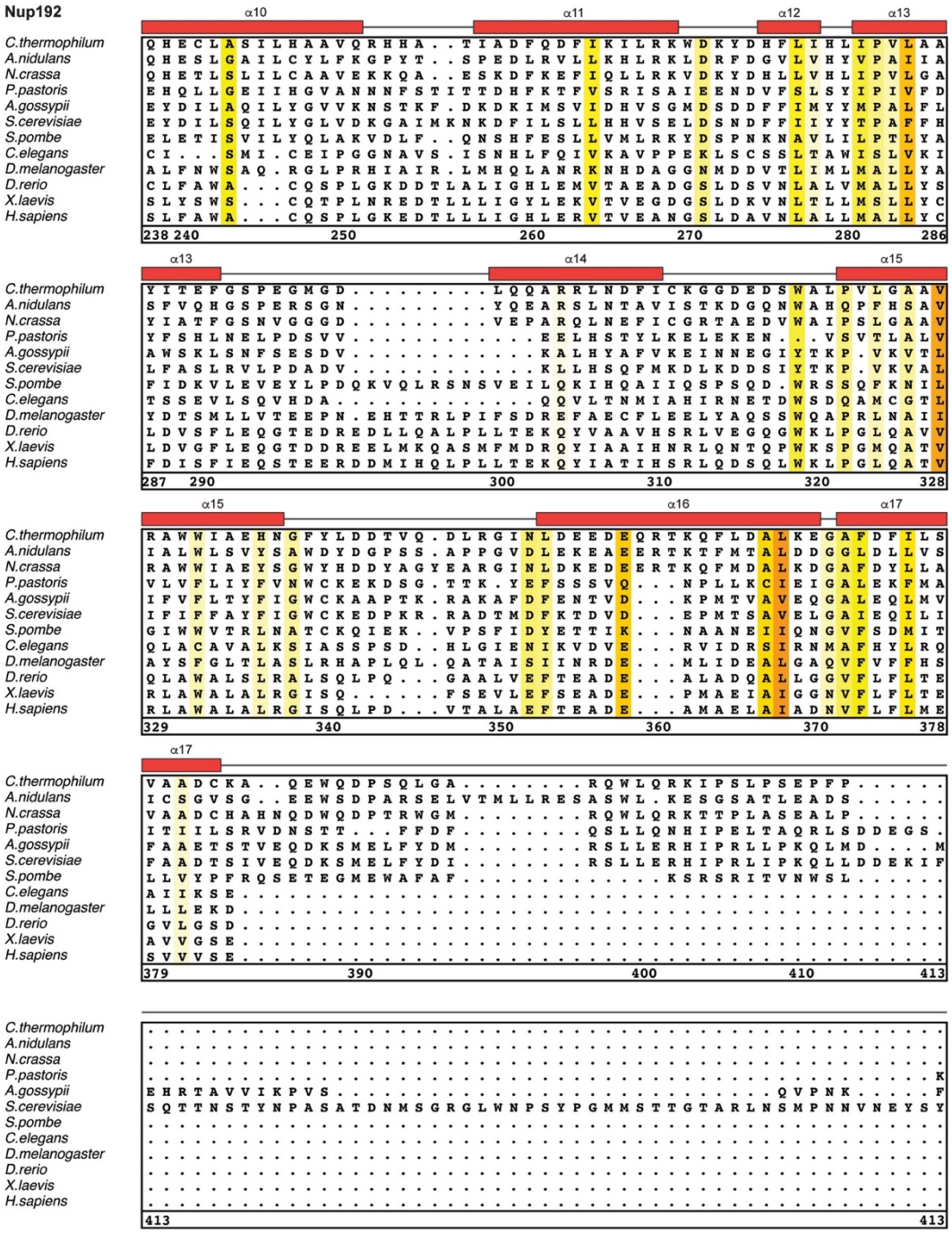


Fig. S30.

Multispecies sequence alignment of Nup192. Sequences from twelve diverse species were aligned and colored by sequence similarity according to the BLOSUM62 matrix from white (less than 55 % similarity), to yellow (55 % similarity), to red (100 % identity). Numbering below alignment is relative to the *C. thermophilum* sequence. Secondary structure observed in the Nup192 structure is shown above the alignment: α-helices (red bars), β-sheets (blue bars), and unstructured regions (black lines). A dashed line indicates the loop that was deleted for crystallization. Disordered regions are indicated by gray dots.


Nup192
α10
α11
α12
α13
C.thermophilum
A.nidulans
N.crassa
P.pastoris
A.gossypii
S.cerevisiae
S.pombe
C.elegans
D.melanogaster
D.rerio
X.laevis
H.sapiens
238 240 250 260 270 280 286
α13
α14
α15
287 290 300 310 320 328
α15
α16
α17
329 340 350 360 370 378
α17
379 390 400 410 413
413 413

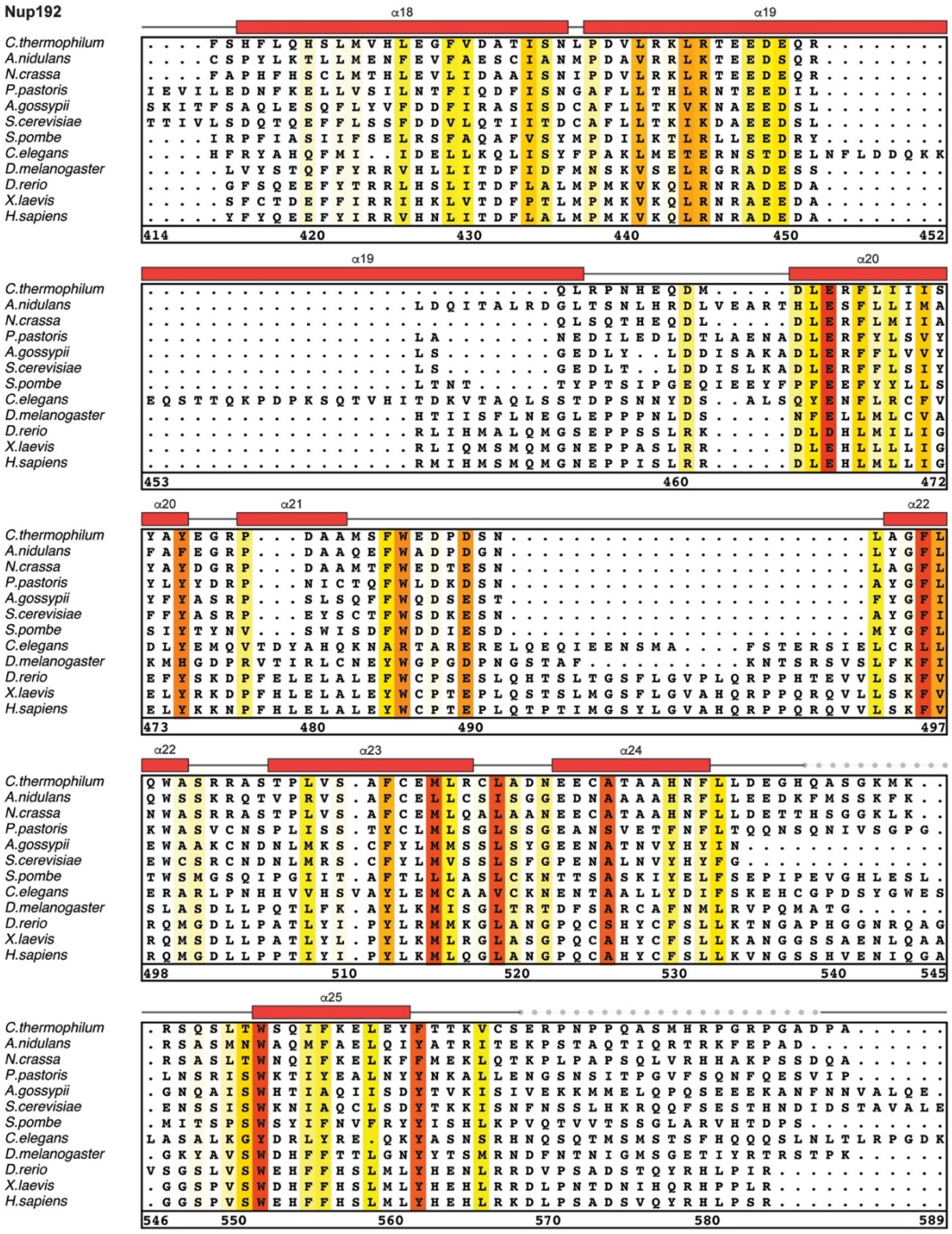


Fig. S30 continued.

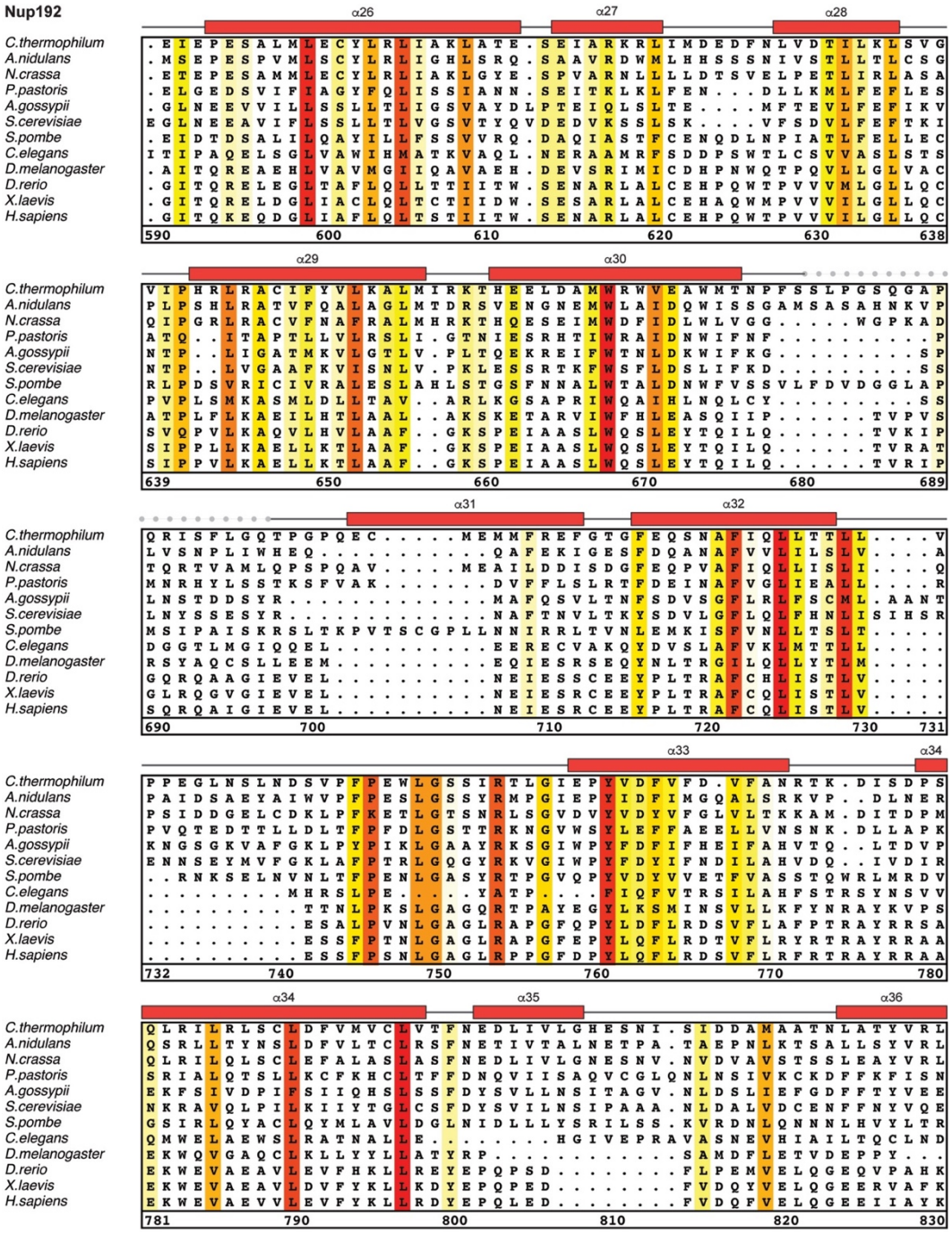


Fig. S30 continued.

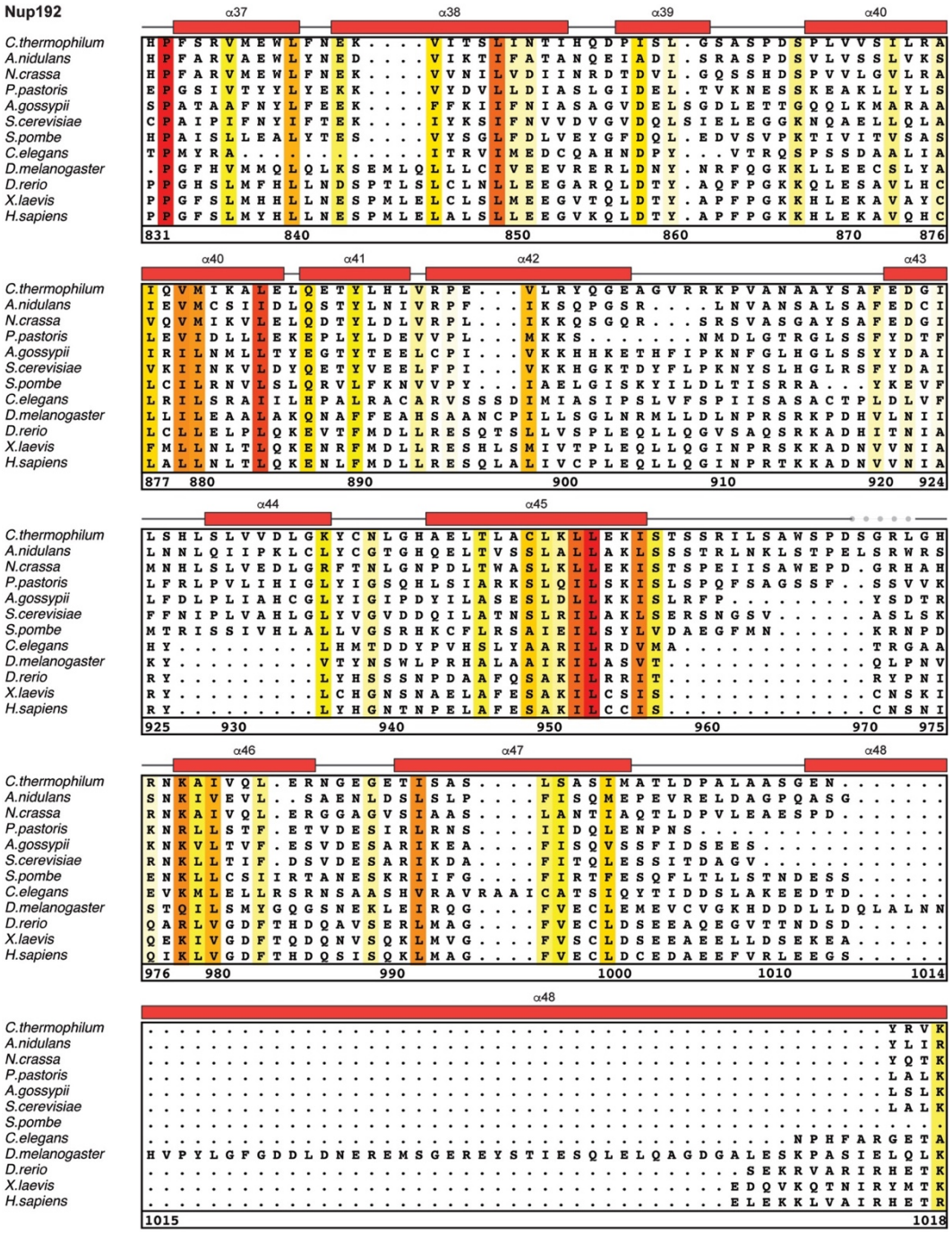


Fig. S30 continued.

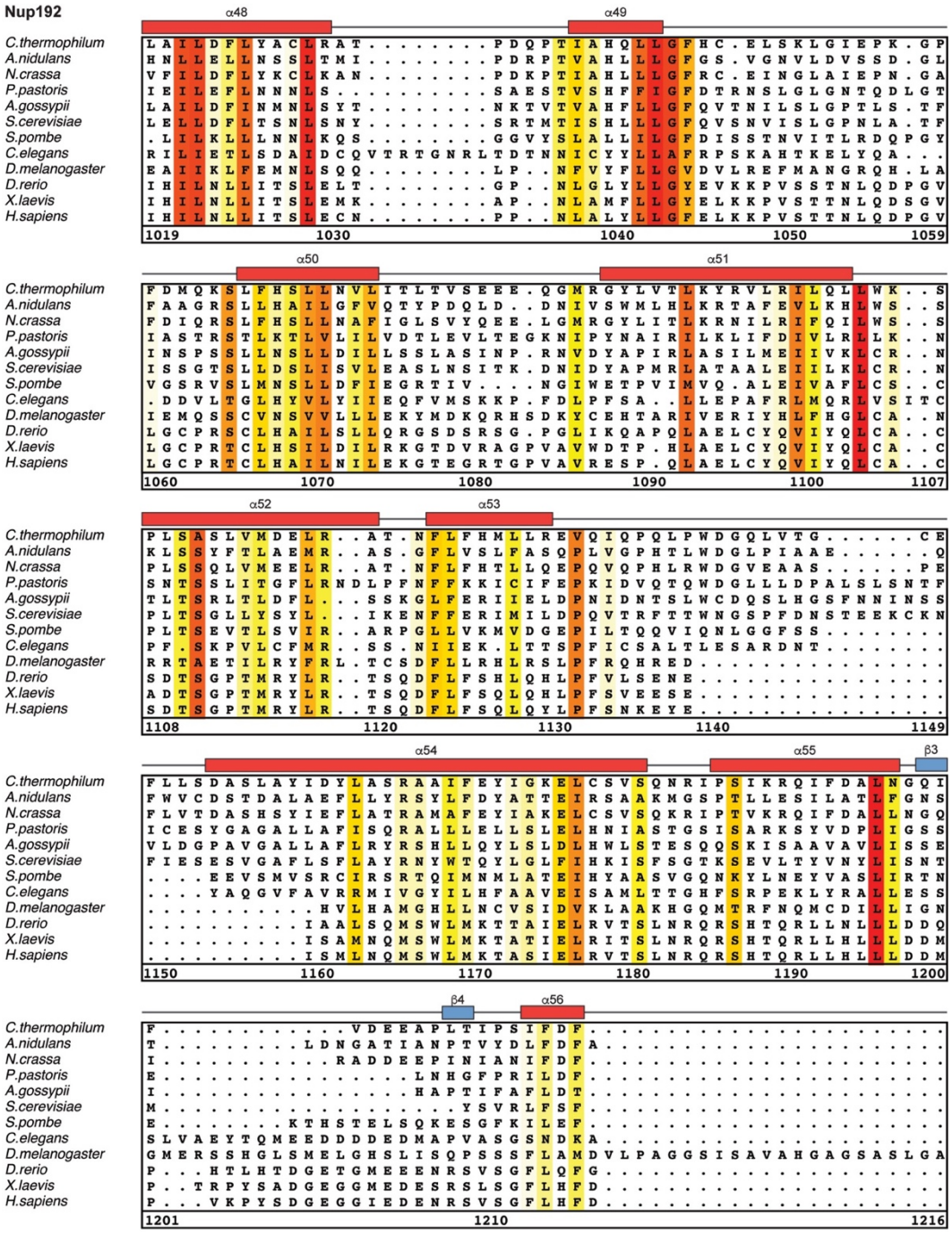


Fig. S30 continued.

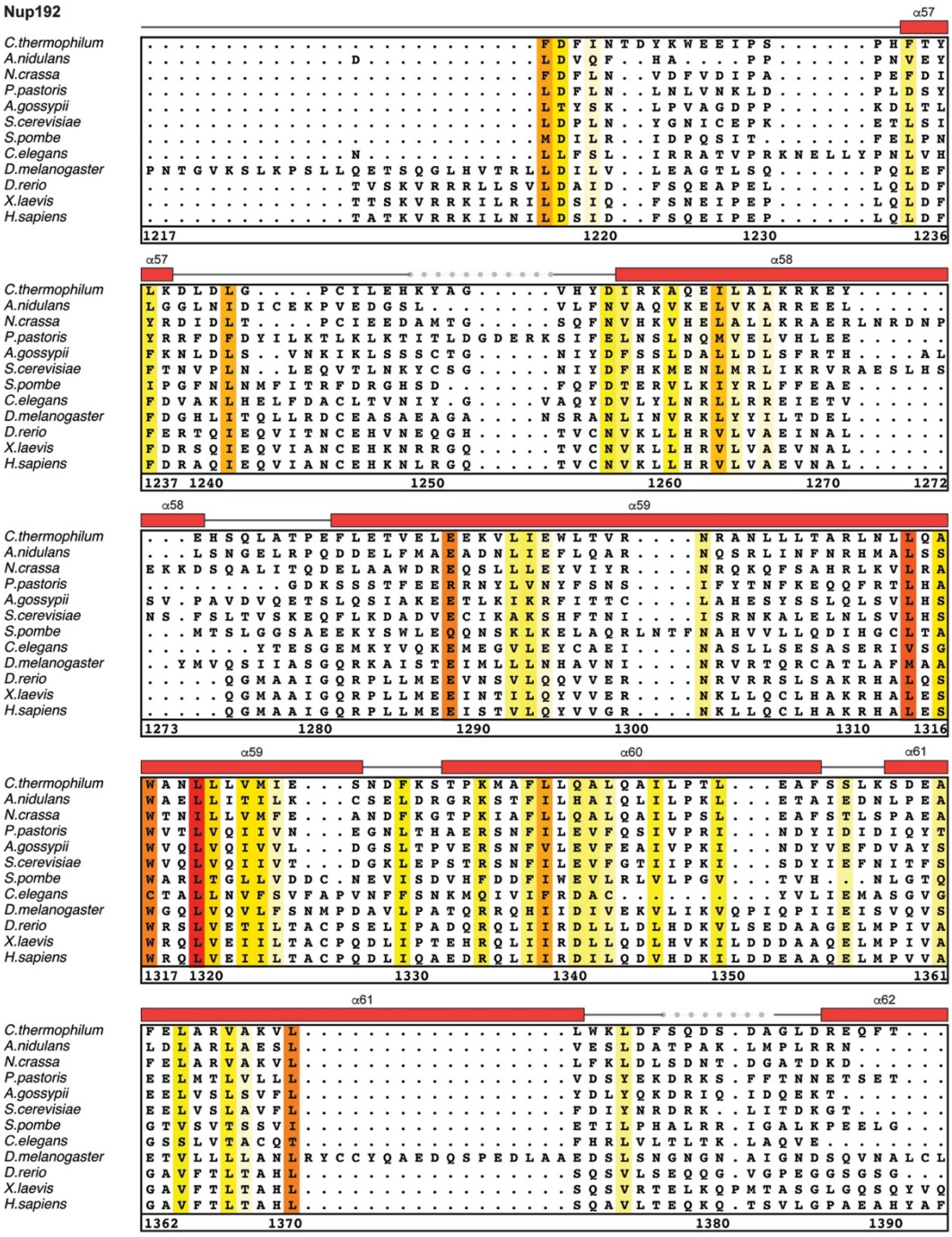


Fig. S30 continued.

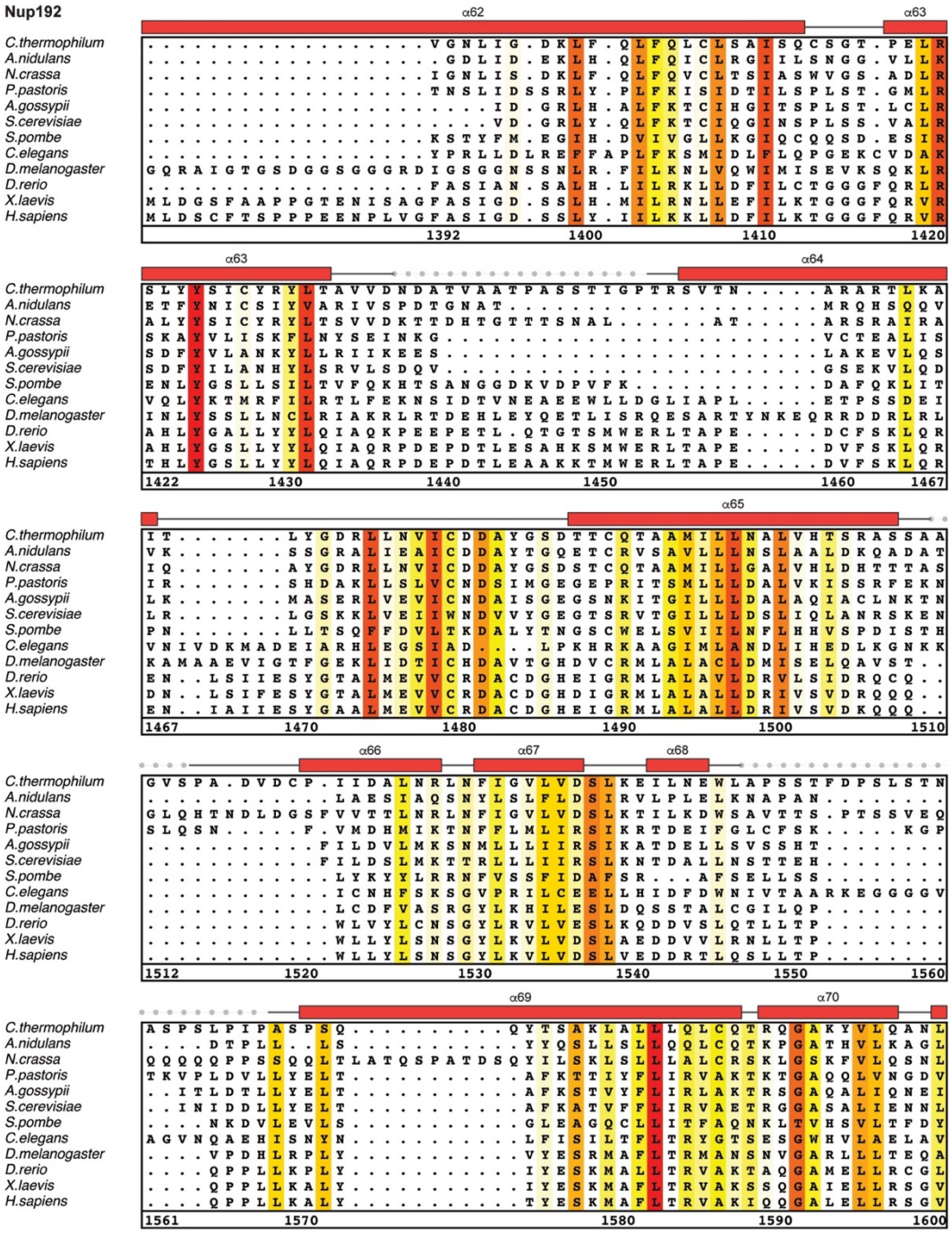


Fig. S30 continued.

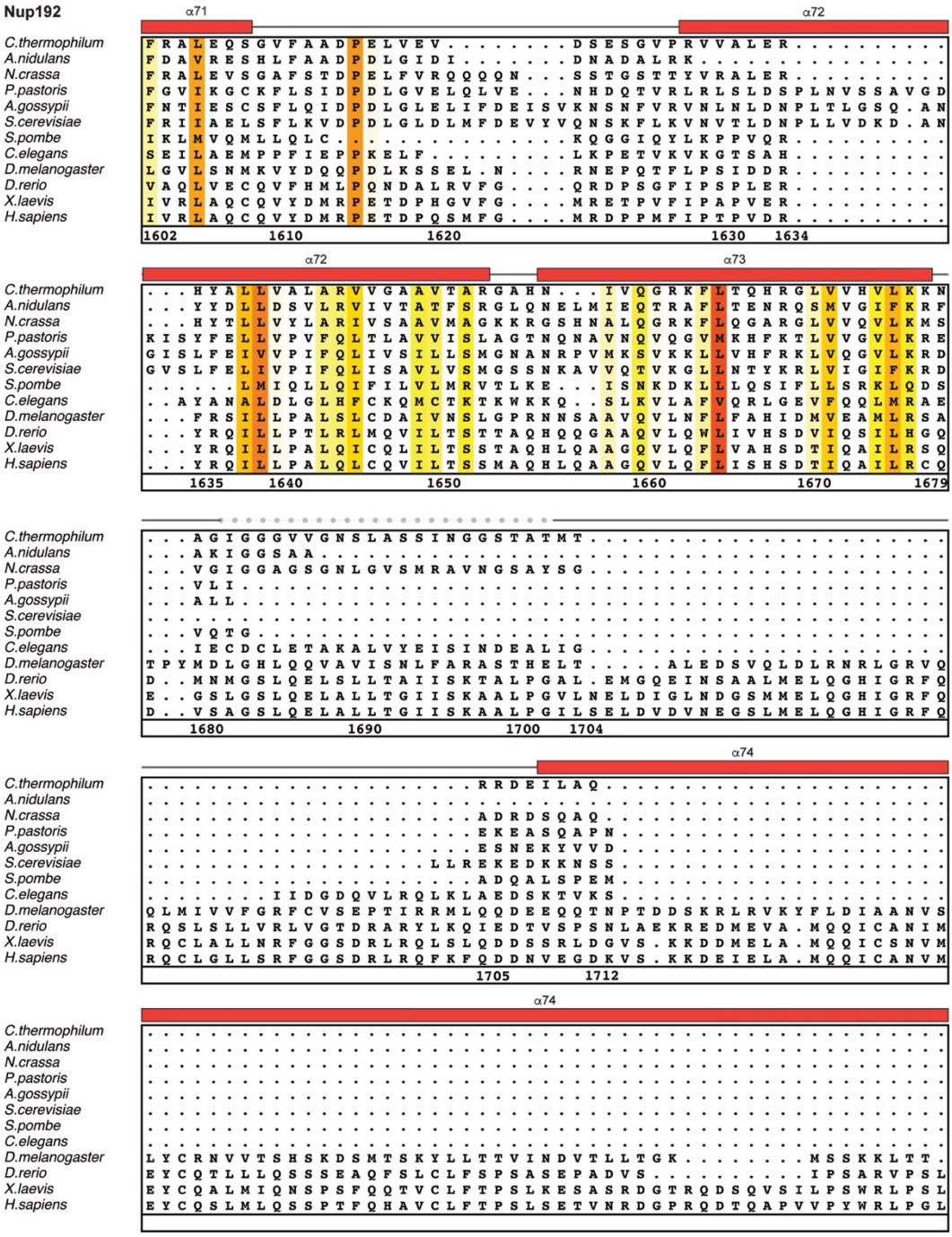


Fig. S30 continued.

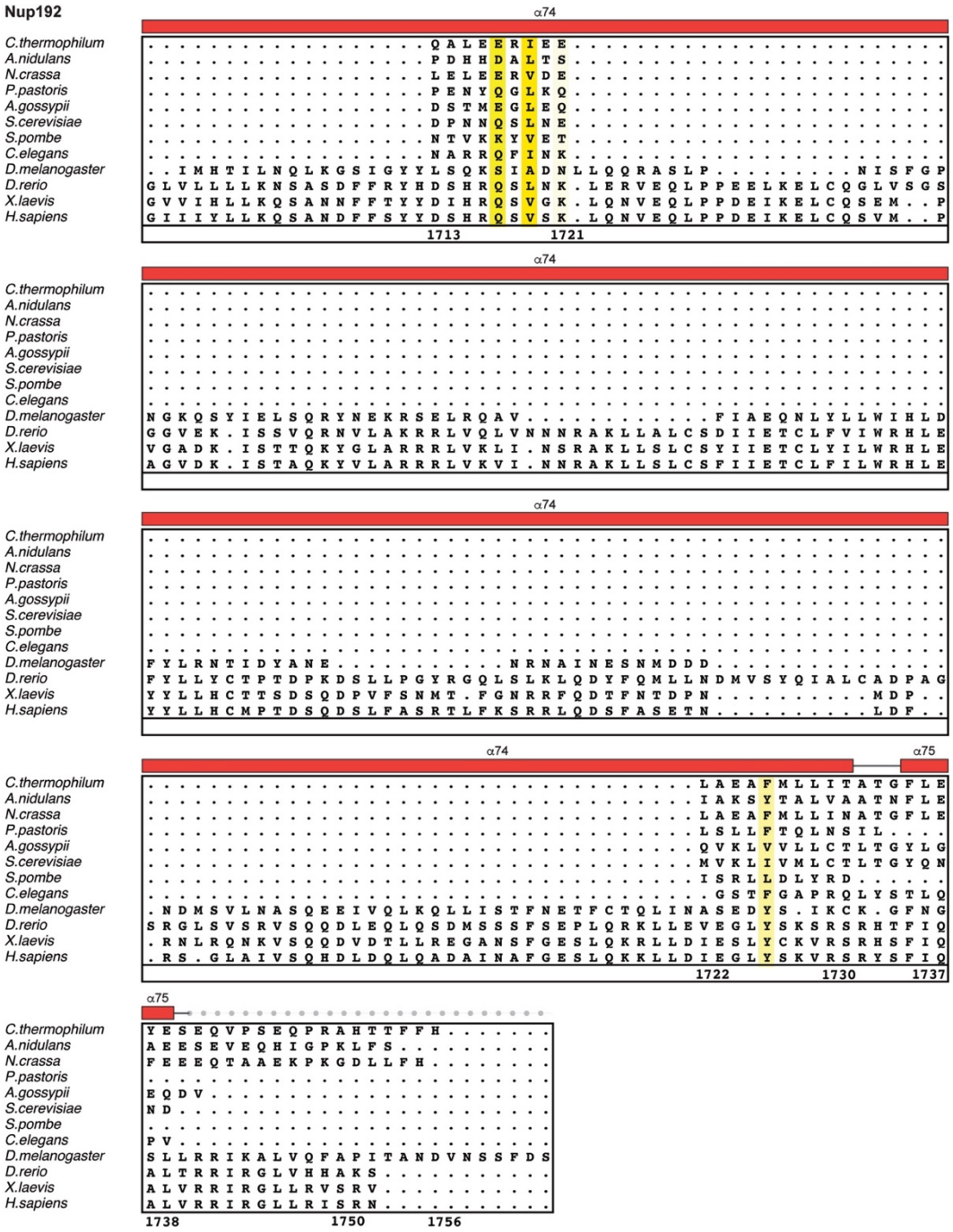


Fig. S30 continued.

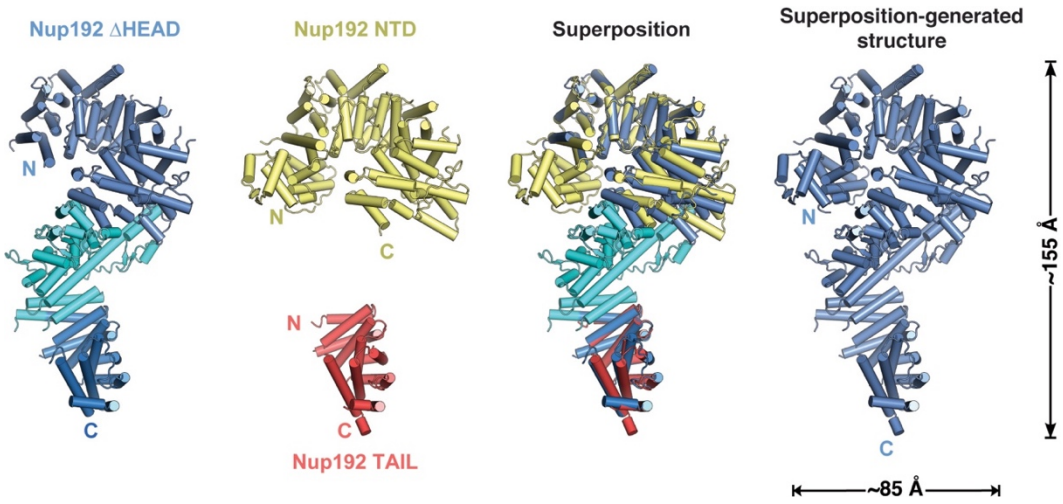


Fig. S31.

Superposition-generated structure of full-length Nup192. Cartoon representations of the crystal structures of $Nup192^{\Delta HEAD}$ (colored as in Fig. 4F), $Nup192^{NTD}$ (yellow), and $Nup192^{TAIL}$ (red) and their superposition are shown. A cartoon representation of the superposition-generated structure of full-length Nup192 (blue) generated from the superposition is shown on the right.

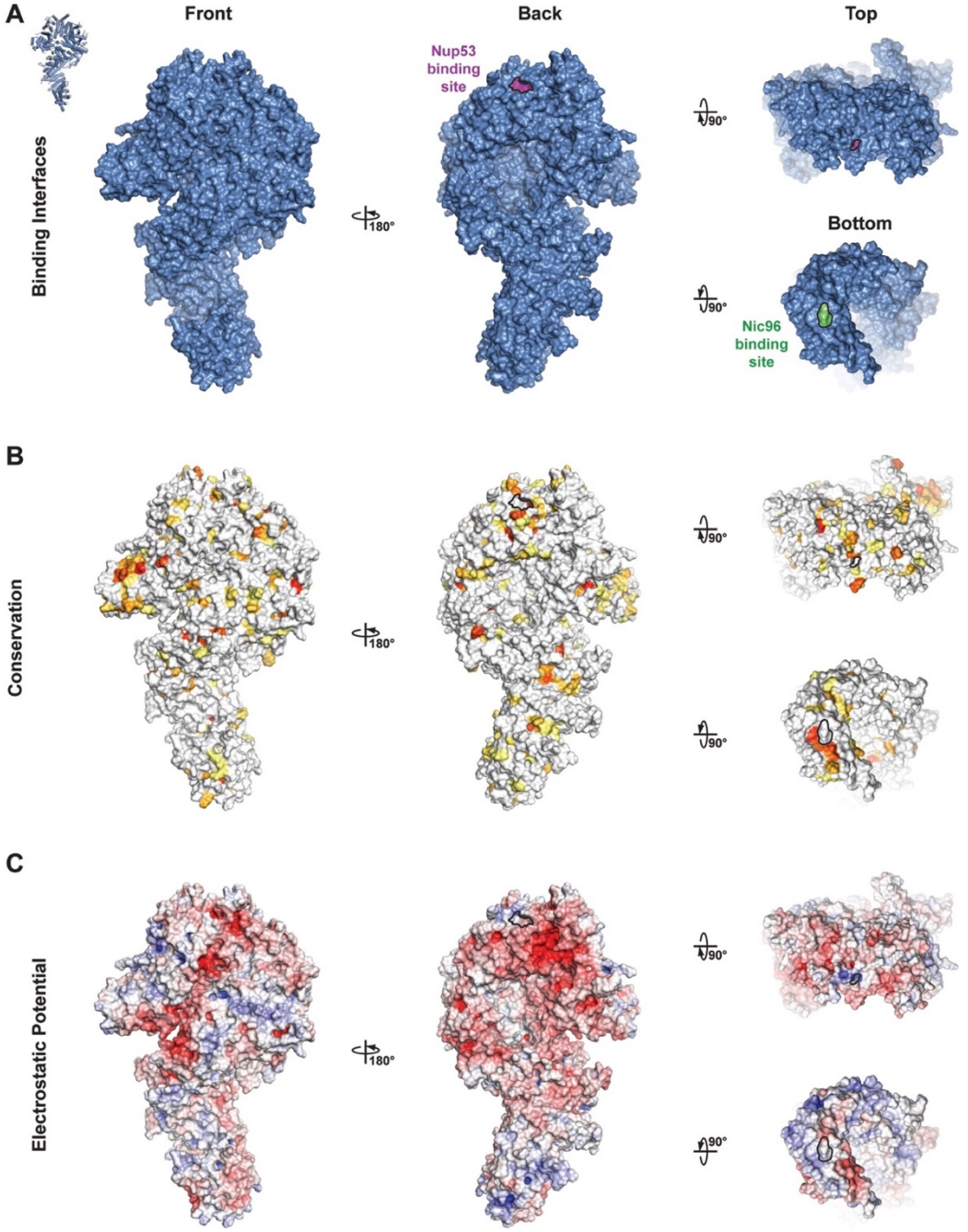


Fig. S32.
Surface properties of Nup192. Surface representations of Nup192 are shown in four different orientations. The surfaces corresponding to previously identified point mutations that disrupt interactions with Nup53 or Nic96 are outlined in black. (**A**) A surface representation of Nup192 with previously identified point mutations that disrupt interactions with Nup53 or Nic96 colored purple or green, respectively (*15, 20*). (**B**) Surface representation colored according to sequence identity based on the alignment in fig. S30. (**C**) Surface representation colored according to electrostatic potential from -10 k_BT/e (red) to +10 k_BT/e (blue).

Fig. S33.

Domain structures of symmetric core nucleoporins are conserved between *H. sapiens* and *C. thermophilum*. Predicted domain boundaries are shown for the symmetric core nucleoporins from *H. sapiens* using *H. sapiens* nomenclature. Percent sequence identity to *C. thermophilum* is shown in parenthesis.

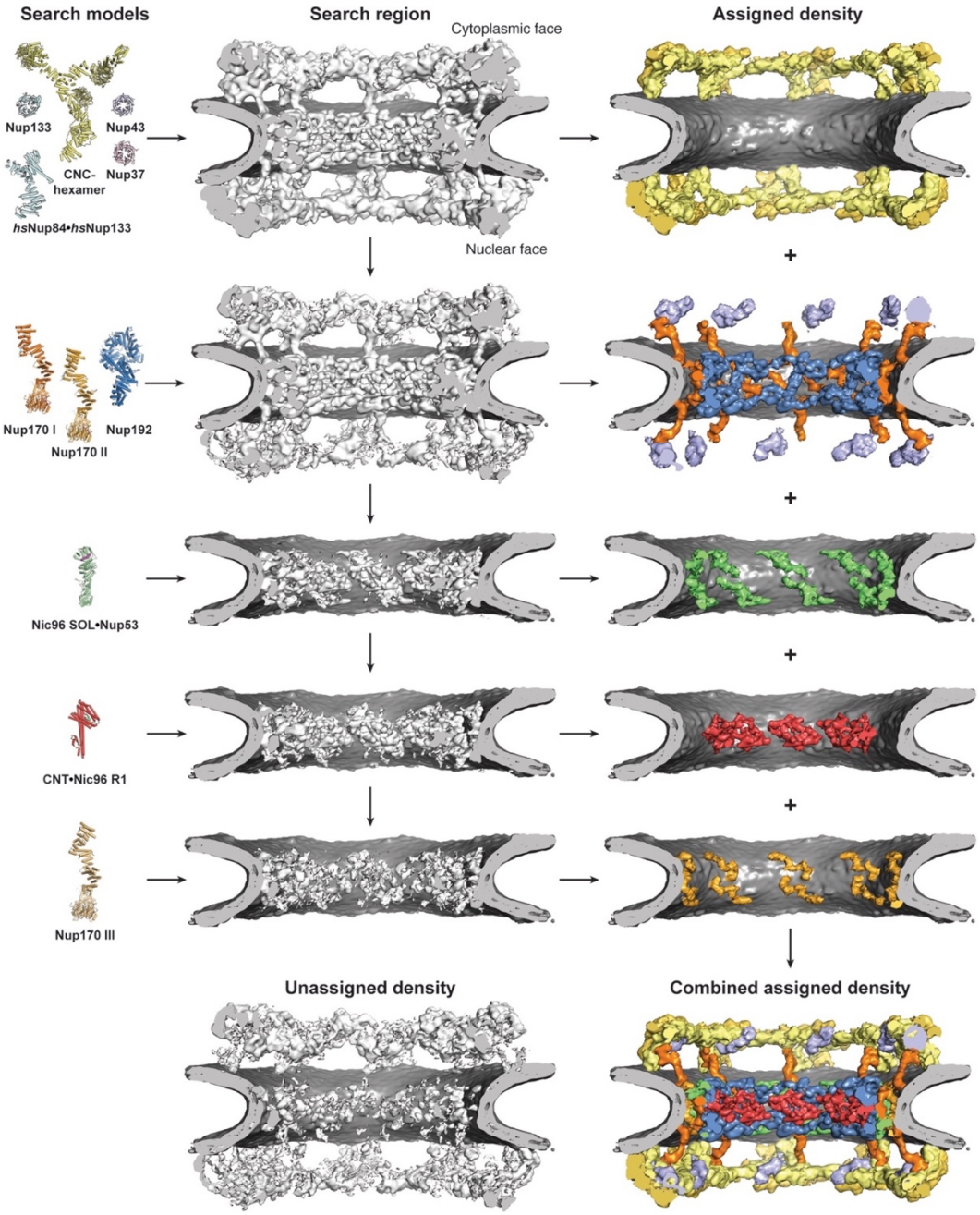


Fig. S34.
Flowchart of the incremental approach used to dock crystal structures into the cryoET reconstruction of the intact human NPC. For each docking step, the crystal structures used for global searches are shown on the left (search model) and the cryoET reconstruction the searches were performed with is shown in the middle (search region). Newly assigned density that was removed from the cryoET reconstruction in subsequent searches is shown on the right and colored as the crystal structure that was placed (assigned density). The remaining unassigned density and combined assigned density are shown at the bottom middle and right, respectively.

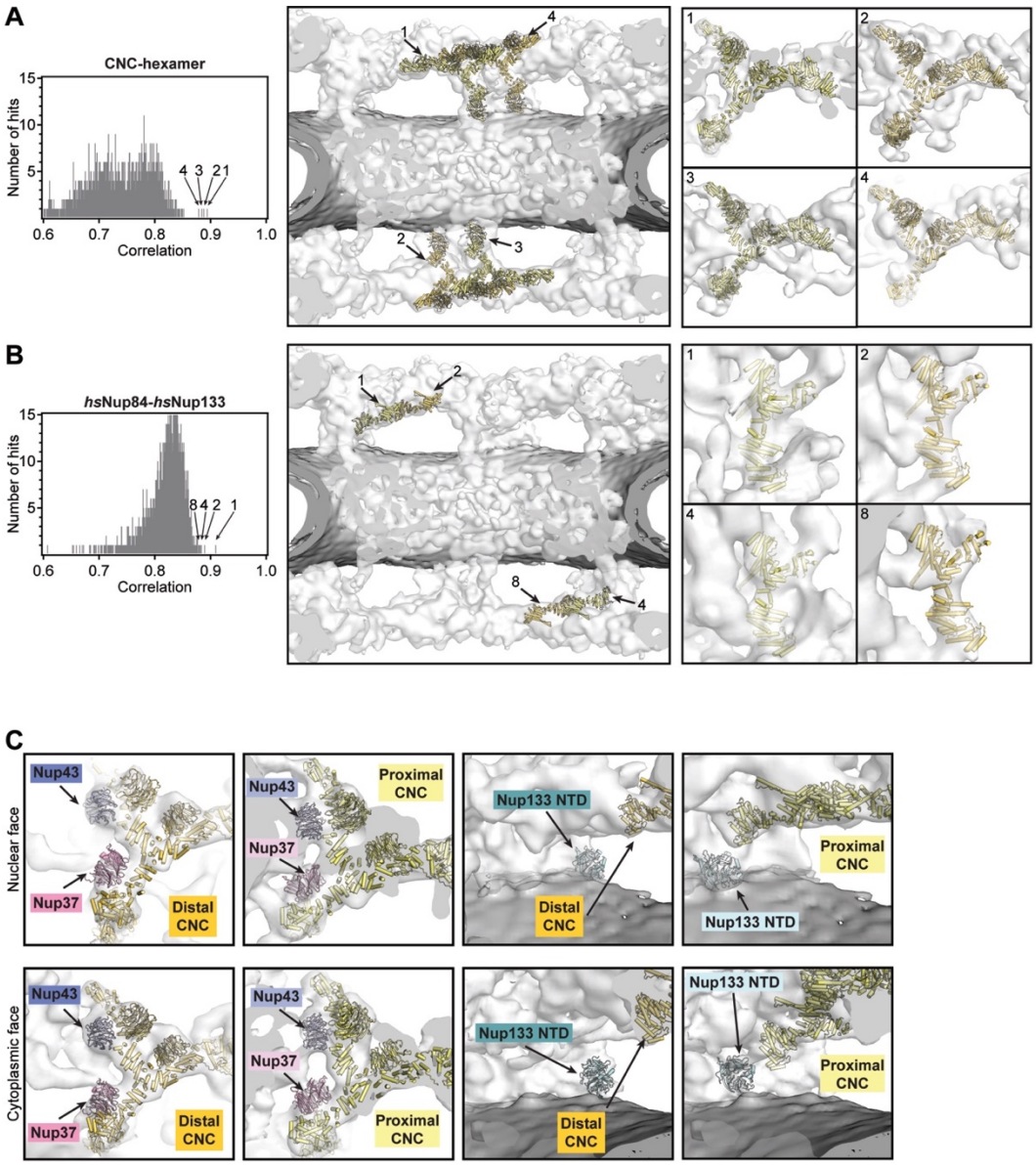


Fig. S35.

Docking of the CNC-hexamer and Nup84•Nup133 crystal structures into the cryoET reconstruction of the intact human NPC. On the left, histograms of the cross-correlation scores from a global search with 50,000 random initial placements are shown for (**A**) the yeast CNC-hexamer (PDB ID 4XMM) and (**B**) *hs*Nup84•*hs*Nup133 hetero-dimer (PDB ID 3I4R) (*14, 23*). Arrows and corresponding numbers indicate the unique solutions that were accepted and the rank of the score, respectively. The arrangement of the unique solutions in one spoke is shown in the middle. The densities corresponding to the nuclear envelope and the NPC are colored in dark gray and white, respectively. Representative views illustrating the quality of the fits are shown on the right. The numbers on the top left of the box indicate which solution is depicted. (**C**) Manual docking of Nup43 (PDB ID 4I79; periwinkle), Nup37 (PDB ID 4FHM; pink), and Nup133^{NTD} (PDB ID 1XKS; turquoise) β-propellers into cryoET density guided by previously published biochemical and structural data (*39-41, 50*).

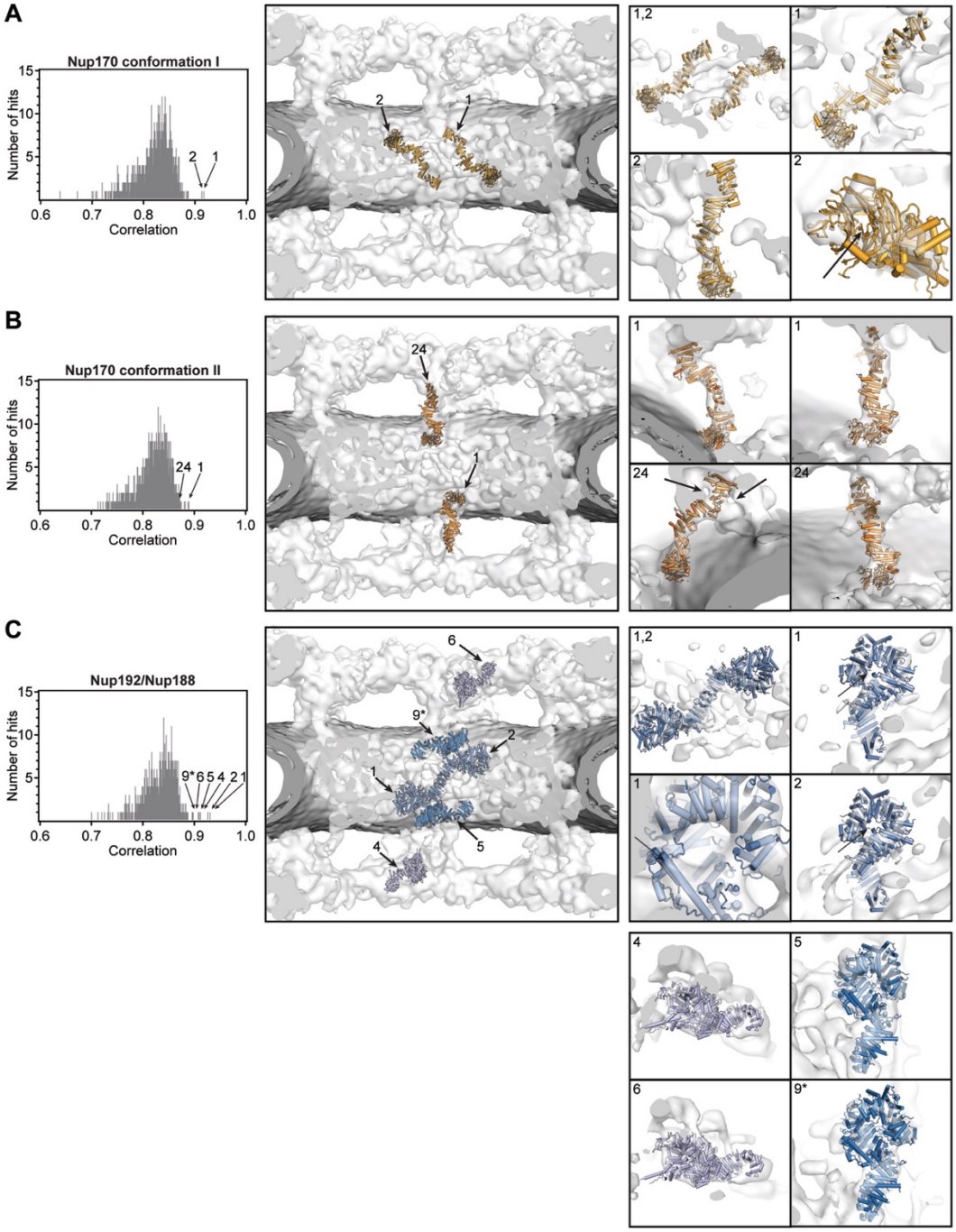


Fig. S36.
Docking of the superposition-generated full-length Nup170 and Nup192 structures into the cryoET reconstruction of the intact human NPC. On the left, histograms of the cross-correlation scores from a global search with 50,000 random initial placements are shown for (**A**) Nup170 conformation I (light orange), (**B**) Nup170 conformation II (dark orange), and (**C**) Nup192 (blue), Nup188 (purple). The Nup170 conformations I and II correspond to conformations 5 and 2, respectively, in fig. S25. Arrows and corresponding numbers indicate the unique solutions that were accepted and the rank of the score, respectively. For Nup192,

we accepted 5 placements from the global search results, two placements in the inner ring, two placements in the outer rings, and one placement on the nuclear peripheral side of the inner ring. A matching cytoplasmic peripheral placement was not found in global searches, but the manual placement on the cytoplasmic peripheral side of the inner ring generated a score which would rank as the 9^{th} highest score (asterisk). The arrangement of the unique solutions in one spoke is shown in the middle. The densities corresponding to the nuclear envelope and the NPC are colored in dark gray and white, respectively. Representative views illustrating the quality of the fits are shown on the right. The numbers on the top left of the box indicate which solution is depicted.

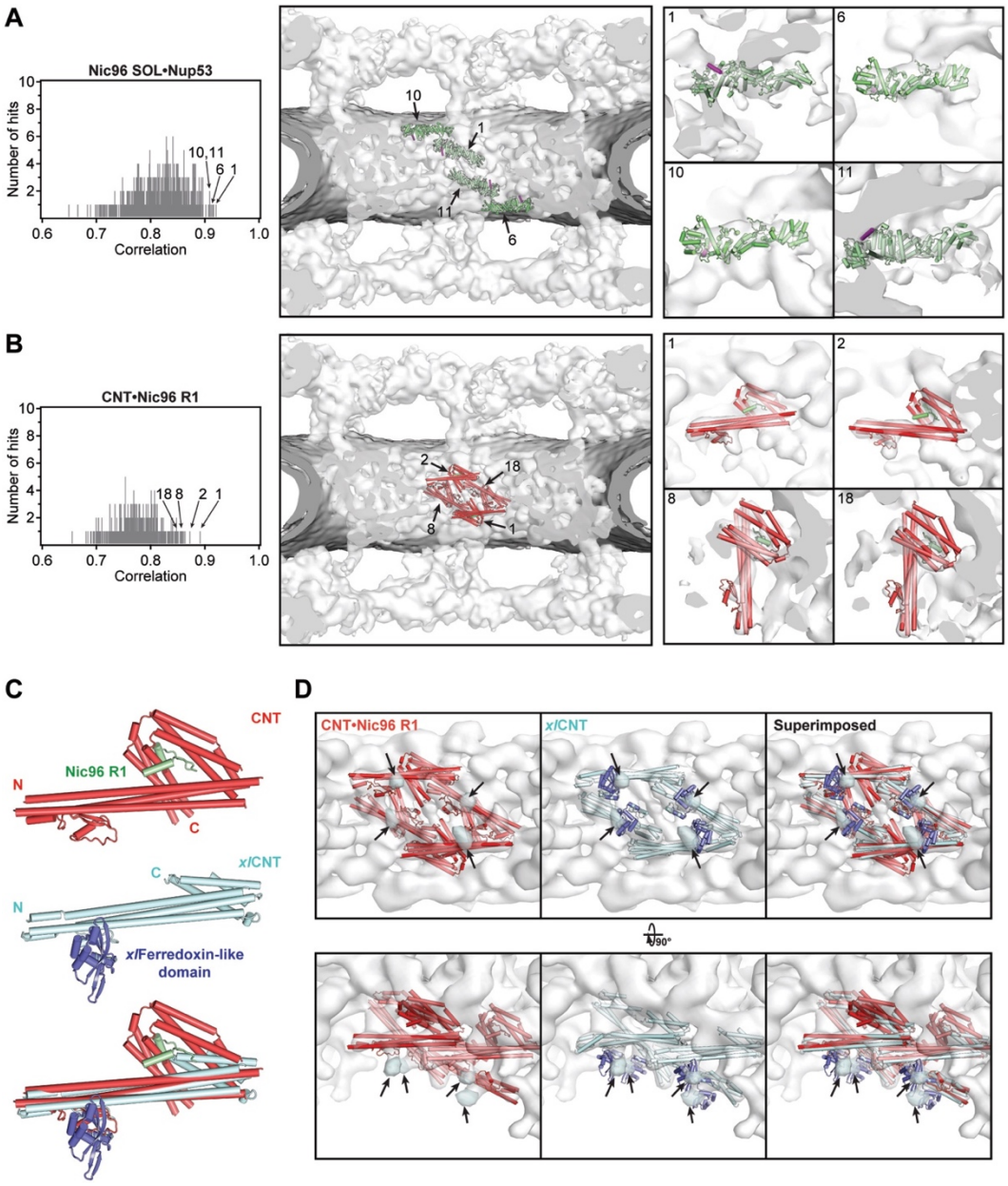


Fig. S37.

Docking of Nic96^{SOL} and CNT crystal structures into the cryoET reconstruction of the intact human NPC. On the left, histograms of the cross-correlation scores from a global search with 50,000 random initial placements are shown for (**A**) Nic96^{SOL}•Nup53^{R2} (green) and (**B**) CNT•Nic96^{R1} (PDB 5CWS; red). Arrows and corresponding numbers indicate the unique solutions that were accepted and the rank of the score, respectively. The arrangement of the unique solutions in one spoke is shown in the middle. The densities corresponding to the nuclear envelope and the NPC are colored in dark gray and white, respectively. Representative views illustrating the quality of the fits are shown on the right. The numbers on the top left of the box indicate which solution is depicted. (**C**) Crystal structures of the *C. thermophilum* CNT•Nic96^{R1} hetero-tetramer (PDB ID 5CWS; red and green; top), the *X. laevis* CNT hetero-trimer (PDB ID 5C3L; light cyan; middle) and ferredoxin-like domain

(PDB ID 5C2U; purple; middle), and their superposition (bottom) are shown in cartoon representation (*15, 16*). (**D**) The four knobs of unexplained density directly adjacent to the docked CNT•Nic96^{R1} crystal structures (colored in cyan and marked with arrows) are readily explained by the presence of the ferredoxin-like domain, a metazoan-specific insertion in Nup57. Side views from within the central transport channel of the NPC (top) and top views from the cytoplasm (bottom) are shown of the cryoET reconstruction with the docked CNT•Nic96^{R1} crystal structures (left), the model of *X. laevis* CNT with ferredoxin-like domain (middle), and their superposition (right).

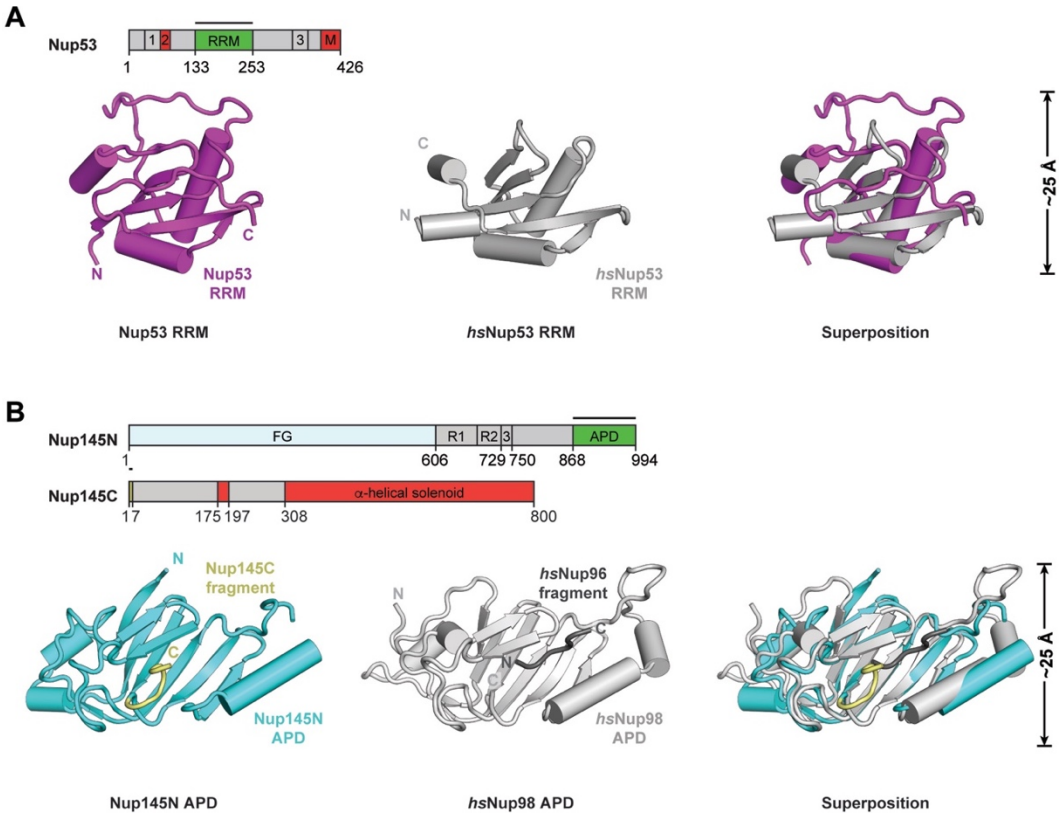


Fig. S38.

Crystal structures of Nup53^{RRM} and Nup145N^{APD}•Nup145C^{N}. (**A**) Crystal structures of *C. thermophilum* Nup53^{RRM} (purple), *H. Sapiens* Nup53^{RRM} (PDB ID 4LIR; gray), and their superposition are shown in cartoon representation. The core fold is conserved, but there are minor alterations in the size of helices and loops. (**B**) Crystal structures of the *C. thermophilum* Nup145N^{APD}•Nup145C^{N} complex (cyan), the *H. sapiens* Nup98^{APD}•Nup96^{N} complex (PDB ID 1KO6; light gray), and their superposition are shown in cartoon representation. As previously observed for the *hs*Nup98^{APD}•*hs*Nup96^{N} hetero-dimer, fusion of the Nup145C^{N} peptide to Nup145N^{APD} and introduction of the catalytically inactive T994A mutation allowed for structure determination of the Nup145N^{APD}•Nup145C^{N} complex. Nup145C^{N} occupied the same binding grove as previously observed in the *hs*Nup98•*hs*Nup96 complex (*30*). As references, the domain structures of Nup53, Nup145N, and Nup145C are shown above the crystal structures with black bars indicating the crystallized fragments.

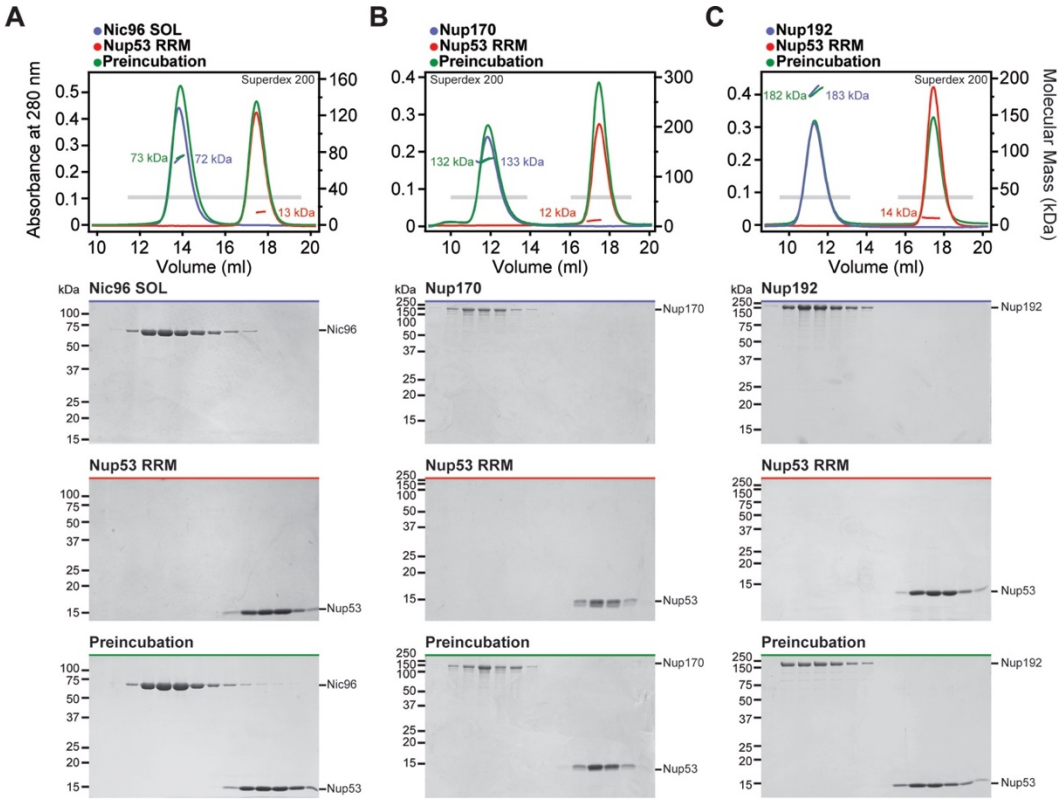


Fig. S39.
Nup53^{RRM} does not interact with inner ring scaffold nucleoporins or the CNC-hexamer. (**A**-**E**) SEC-MALS profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). SEC profiles were obtained using a Superdex 200 10/300 column or Superose 6 10/300 column as indicated. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

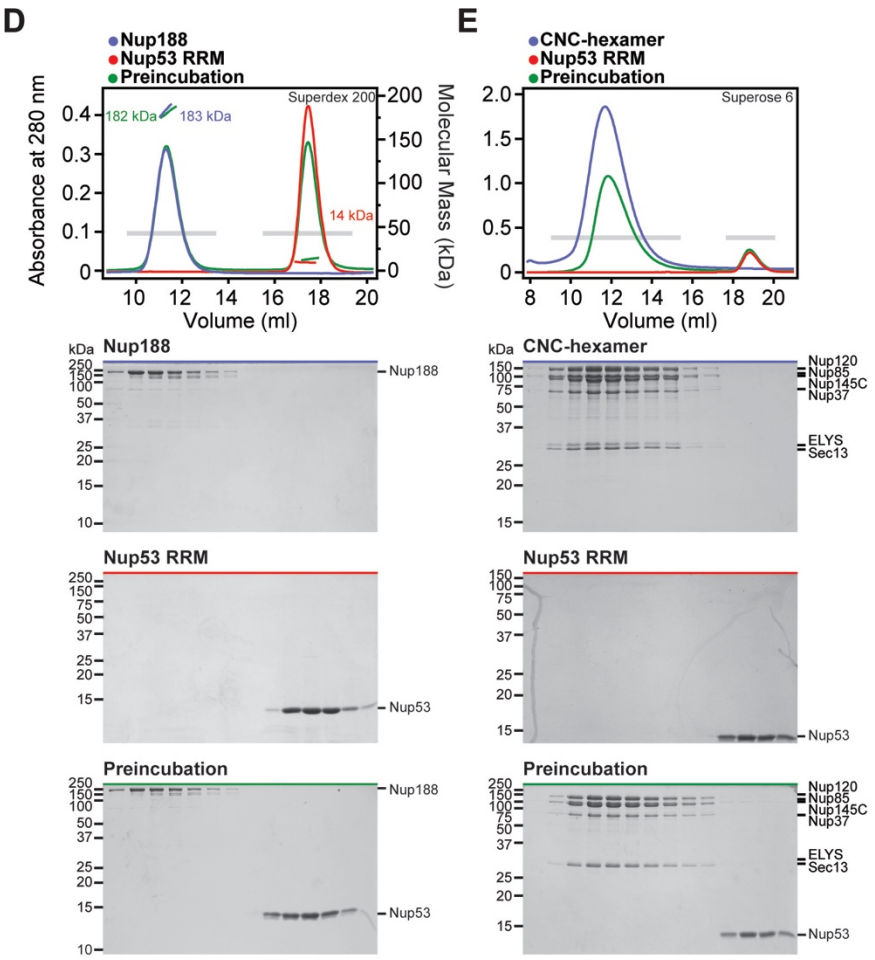


Fig. S39 continued.

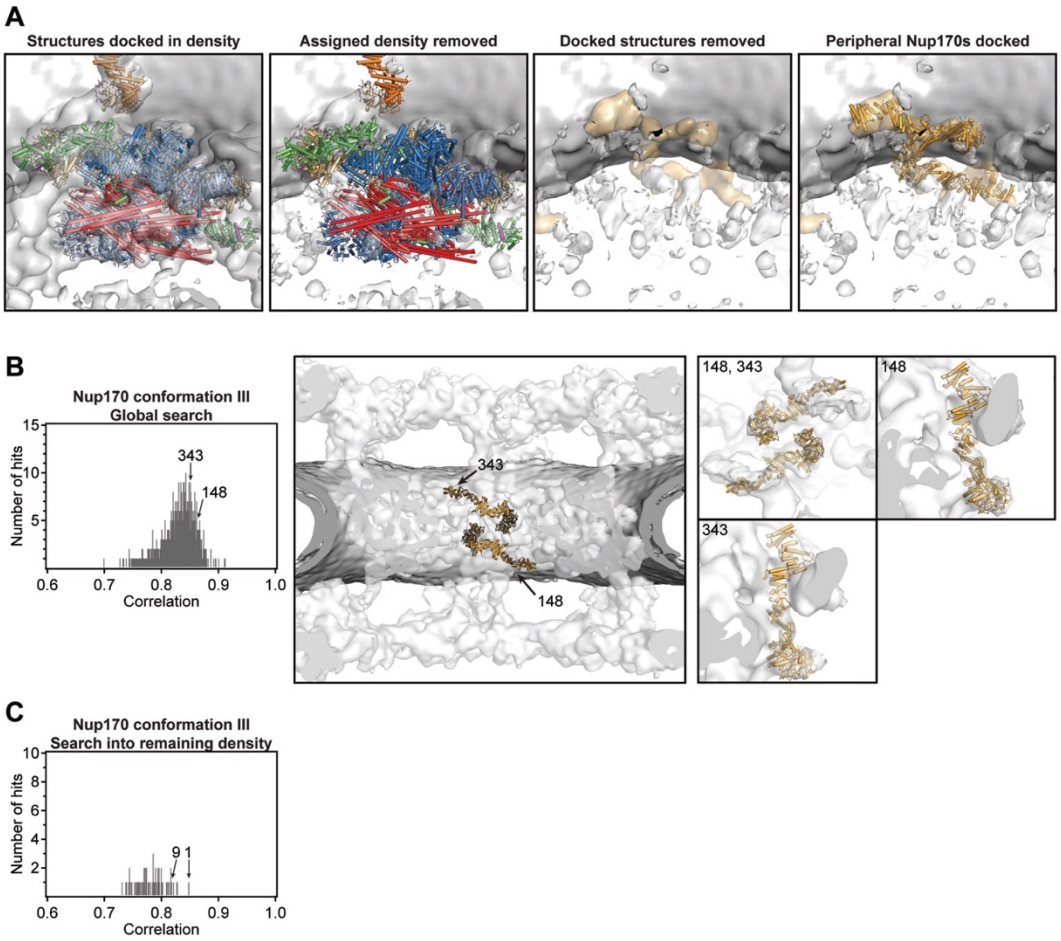


Fig. S40.

Identification of a third pair of Nup170 molecules in the inner ring of the cryoET reconstruction of the intact human NPC. (**A**) Views of the cryoET density in the inner ring illustrating all density and the docked proteins from the unbiased search (far left), the cryoET density remaining after removal of the assigned density and the docked proteins (middle left), the remaining cryoET density without the docked proteins from the unbiased search revealing density that is shaped like a Nup170 molecule colored in orange (middle right), and the same density with a third conformation of Nup170 docked (far right). (**B**) Search results for two additional Nup170 molecules fitted into a cryoET map from which the density for the CNCs had been removed. The histogram on the left shows the cross-correlation scores for a pair of Nup170 conformation III (light orange) from a global search with 50,000 random initial placements. Nup170 conformation III corresponds to conformation 8 in fig. S25. Arrows and corresponding numbers indicate the unique solutions that were accepted and the rank of the score, respectively. The arrangement of the unique solutions in one spoke is shown in the middle. The densities corresponding to the nuclear envelope and the NPC are colored in dark gray and white, respectively. Representative views illustrating the quality of the fits are shown on the right. The numbers on the top left of the box indicate which solution is depicted. (**C**) A histogram illustrating the search results for fitting Nup170 conformation III into a cryoET map of the inner ring from which the density corresponding to all other components had been removed.

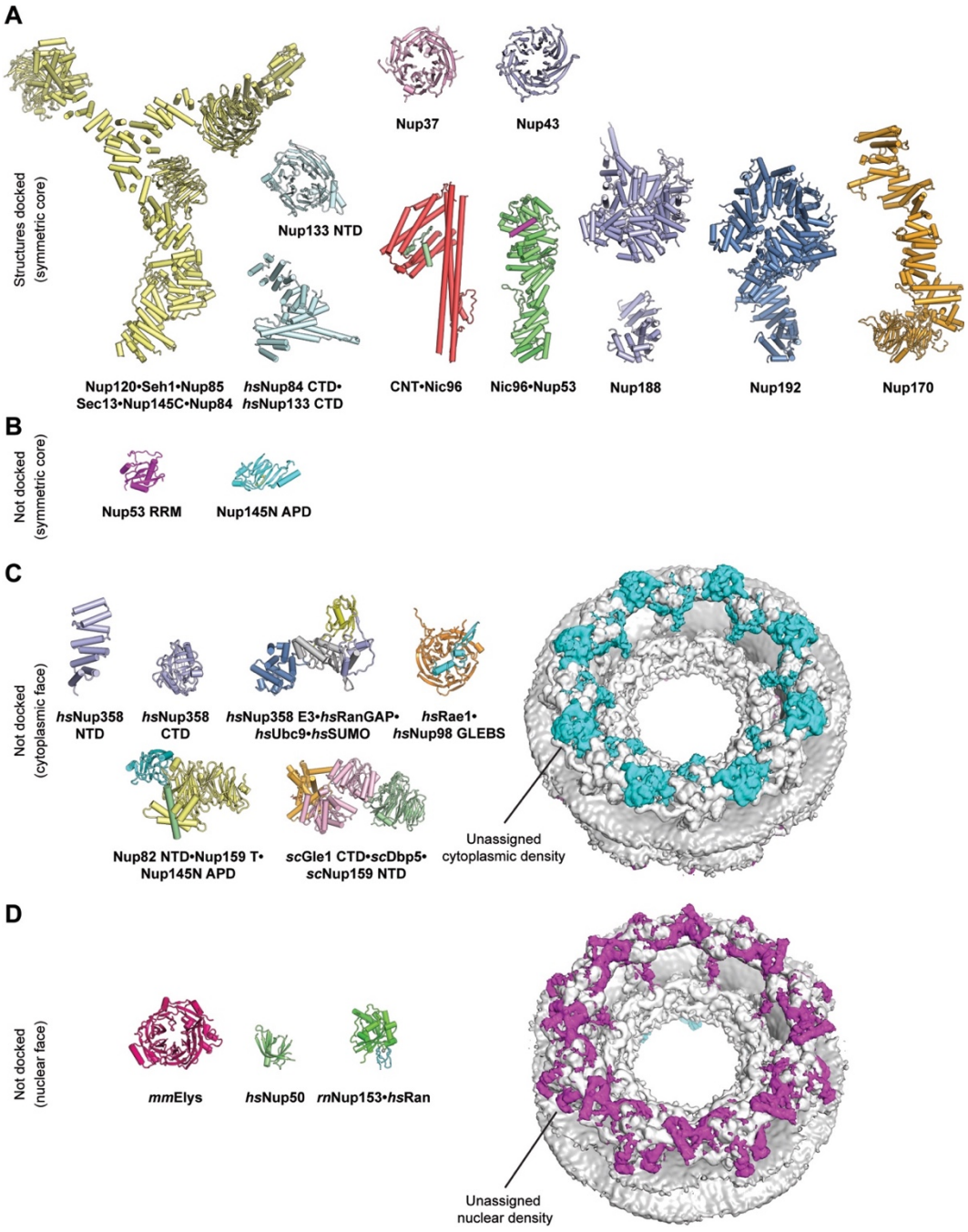


Fig. S41.

Summary of crystal structure docking into the cryoET reconstruction of the intact human NPC. (**A**) Crystal structures of the symmetric core nucleoporins that could be successfully docked into a cryoET reconstruction of the intact human NPC. (**B-D**) Crystal structures of symmetric core nucleoporin domains, the cytoplasmic filament nucleoporins, and nuclear basket nucleoporins, which were not docked due to their small size, indistinctive shape, or lack of biochemical restraints. Cytoplasmic and nuclear views of the cryoET reconstruction of the intact human NPC are shown on the right. The assigned density of the symmetric NPC core is colored in gray and the unassigned densities on the cytoplasmic and

nuclear faces are colored in cyan and purple, respectively. The volume of the unassigned density appears to be sufficient to accommodate the majority of the remaining structured protein mass of the asymmetric cytoplasmic filament and nuclear basket nucleoporins. All crystal structures are shown to scale in cartoon representation.

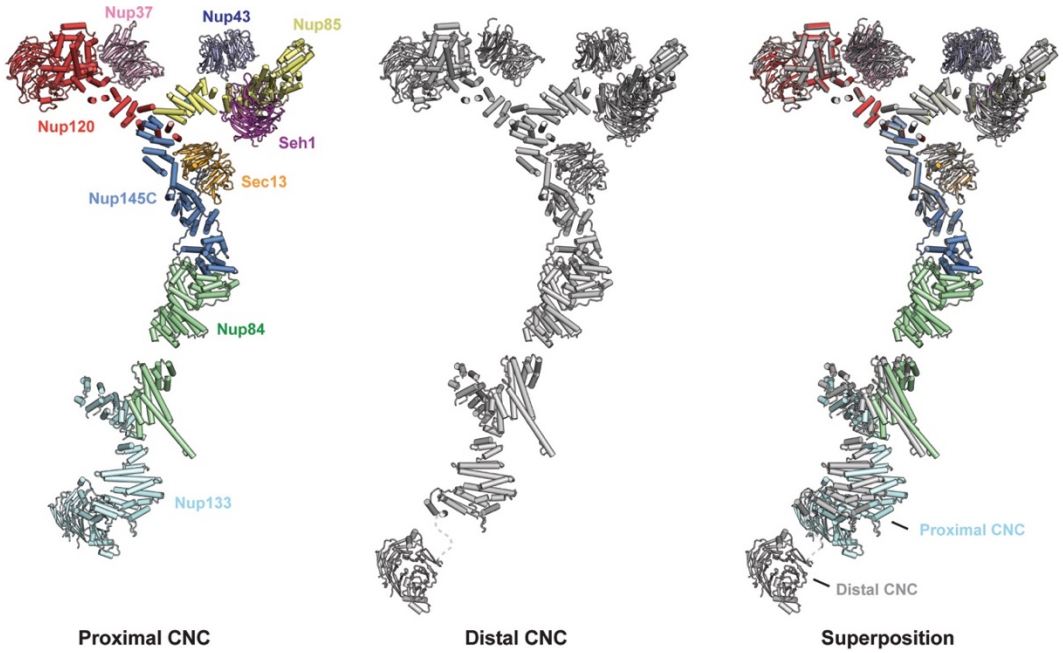


Fig. S42.

Comparison of the distal and proximal CNC docking solutions. Individual distal (multicolored; left) and proximal (gray; middle) CNCs are shown in cartoon representation. Their superposition reveals a dramatically different orientation for Nup133^{NTD} and slightly different orientations of the Nup133 and Nup84 solenoids (right).

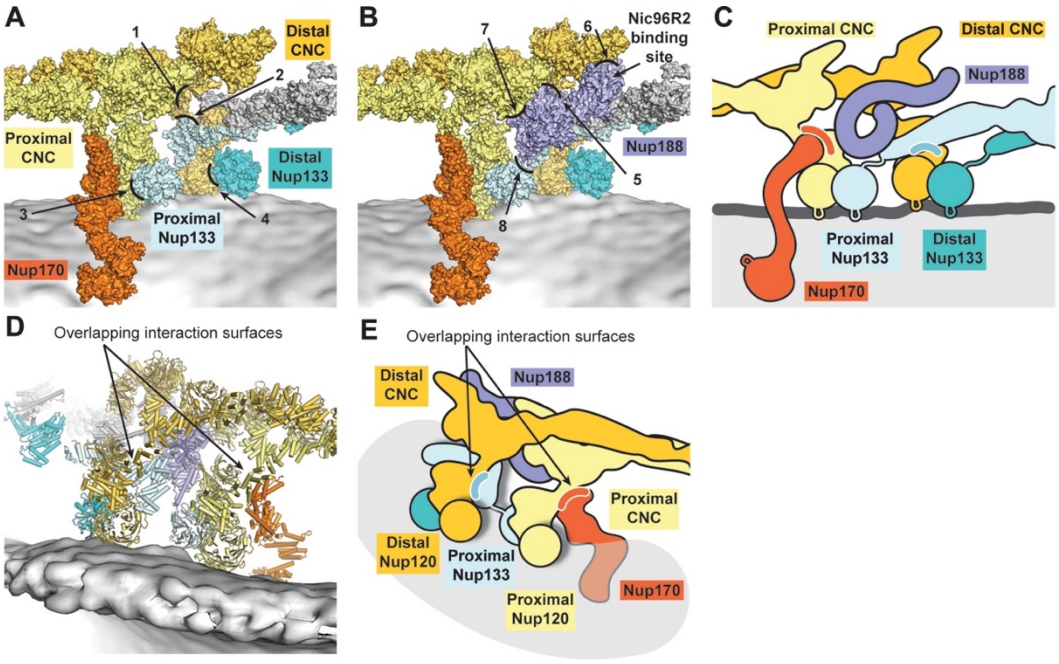


Fig. S43.

Interactions mediating inter-spoke assembly. (**A**) The outer ring inter-spoke interface is shown in surface representation, but with Nup188 removed for clarity. Contacts observed between CNCs from different spokes are indicated by arrows and numbered from 1 to 4, corresponding to the contacts described in the text. (**B**) The same view as in panel (A), but with Nup188 also shown in a surface representation. Contacts observed between Nup188 and the CNCs at the inter-spoke interface are indicated by arrows and numbered from 5 to 8, corresponding to the contacts described in the text. An arrow highlights the Nic96^{R2} binding site on Nup188. (**C**) Schematic of the inter-spoke interface, from the same view as in panel (A). (**D**) The outer ring inter-spoke interface, shown in cartoon representation, ~180° rotated from the view in panel (A). Nup170 and Nup133 contact overlapping interfaces on proximal and distal Nup120 molecules. (**E**) Schematic of the outer ring inter-spoke interface drawn in the same view as in panel (D).

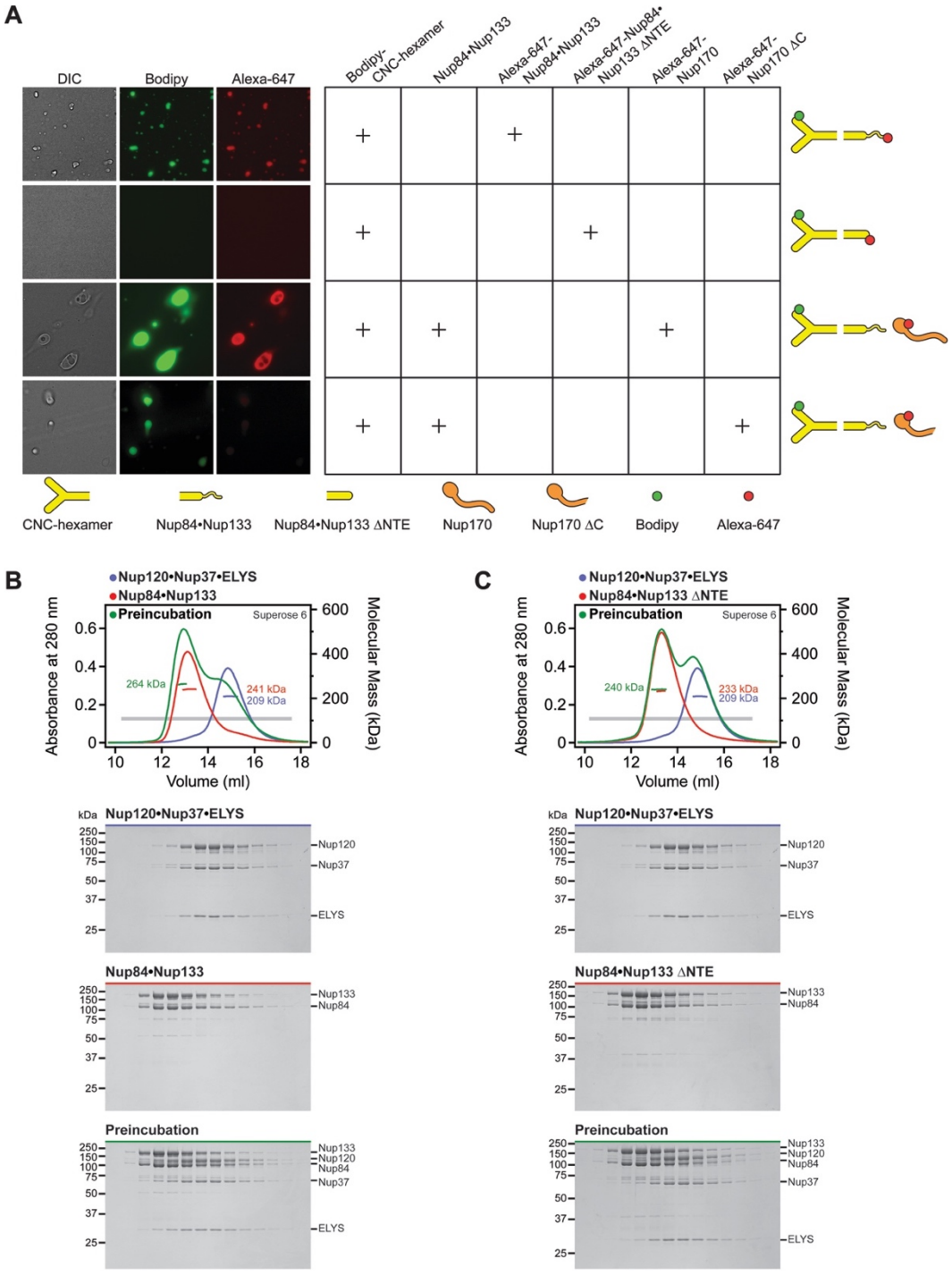


Fig. S44.

Oligomerization of the CNC. (**A**) The CNC forms oil droplets in the presence of the unstructured N-terminal extension (NTE) of Nup133. The CNC-hexamer was labeled with Bodipy and various other nucleoporins were labeled with Alexa-647 and tested for their ability to form oil droplets. Nup170 incorporates CNC oil droplets, but this incorporation is ablated by deletion of C-terminal helices. Tested protein combinations are indicated in the table and shown schematically on the right. All experiments were repeated at least three

times. (**B**, **C**) SEC-MALS profiles of nucleoporin complexes are shown individually (blue or red) and after their preincubation (green). There is a weak interaction between Nup120•Nup37•ELYS and Nup84•Nup133 only in the presence of the NTE of Nup133. SEC profiles were obtained using a Superose 6 10/300 GL column. Measured molecular masses are indicated for the peak fractions. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

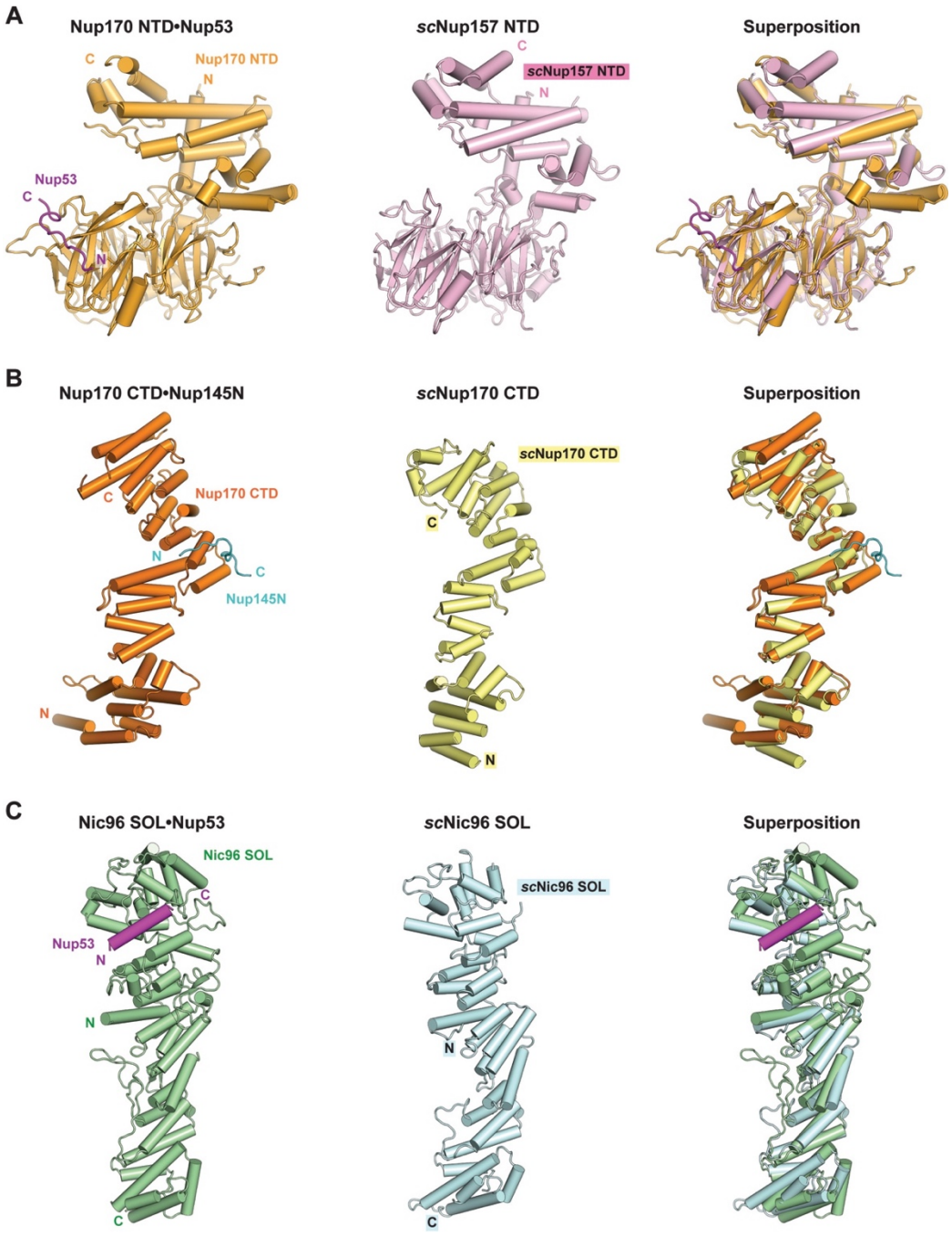


Fig. S45.

Overall nucleoporin folds are evolutionarily conserved. Structural comparison of *C. thermophilum* Nup170 and Nic96 complexes reported here and previously determined *apo S. cerevisiae* crystal structures, revealing extensive evolutionary conservation of the overall folds. (**A**) $Nup170^{NTD}$•$Nup53^{R3}$ (left), *sc*$Nup157^{NTD}$ (PDB ID 4MHC; middle), and their superposition (right) are shown in cartoon representation (*21*). (**B**) $Nup170^{CTD}Nup145N^{R3}$ (left), *sc*$Nup170^{CTD}$ (PDB ID 3I5P; middle), and their superposition (right) are shown in cartoon representation (*23*). (**C**) $Nic96^{SOL}$•$Nup53^{R2}$ (left), *sc*$Nic96^{SOL}$ (PDB ID 2RFO; middle), and their superposition (right) are shown in cartoon representation (*24*).

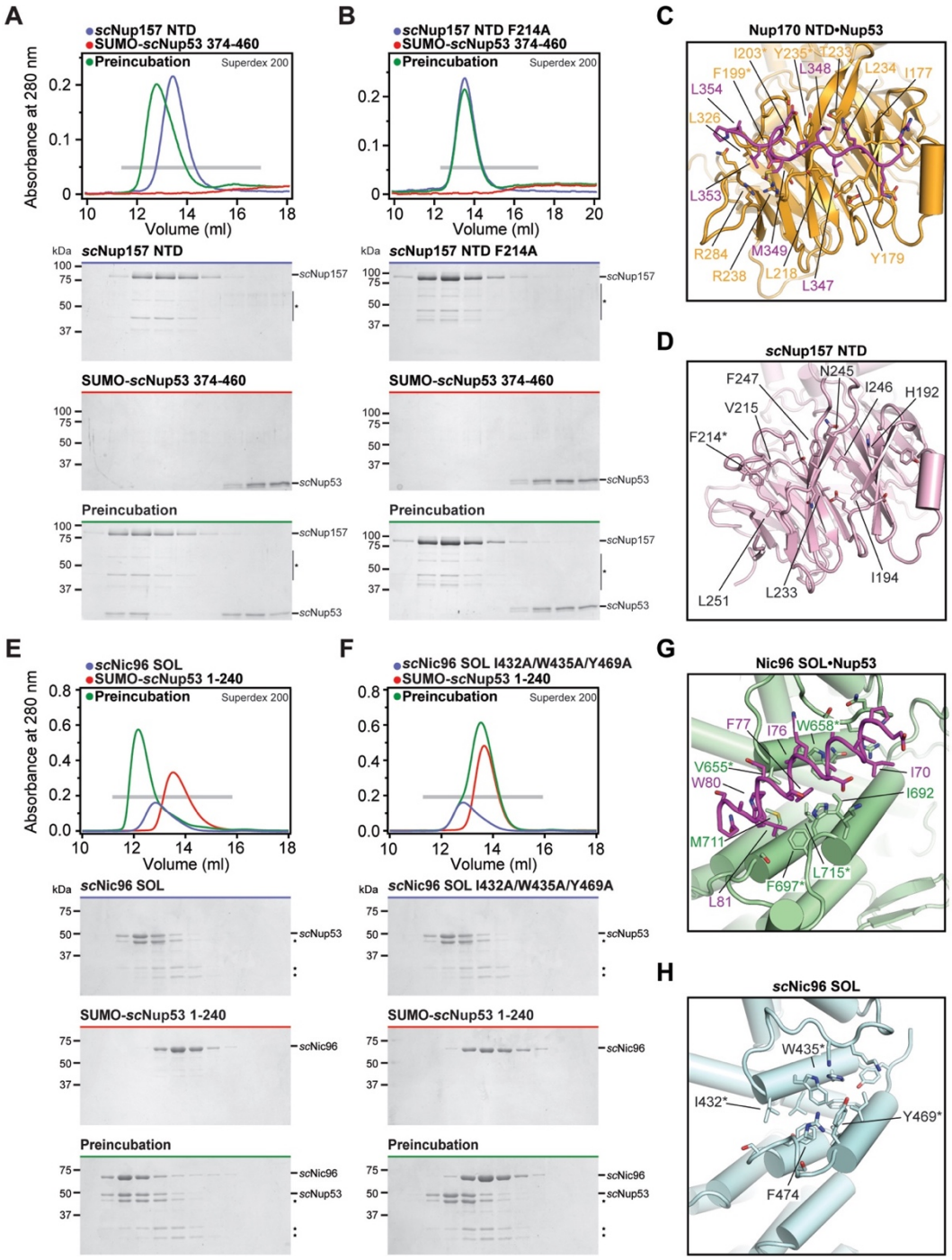


Fig. S46.

Binding pockets on *sc*Nup157 and *sc*Nic96 for *sc*Nup53 are evolutionarily conserved. (**A**, **B**) SEC interaction experiments were performed with constructs of *S. cerevisiae* Nup157^{NTD} and Nup53 homologous to the crystallized Nup170^{NTD}•Nup53^{R3} complex and with a mutant Nup157^{NTD} construct in which a conserved F214 residue in the Nup53 binding pocket was mutated to alanine. SEC profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). (**C**) A close-up view of the

Nup170^{NTD}•Nup53^{R3} complex, same as Fig. 3C. (**D**) A close-up view of the corresponding surface of *sc*Nup157 (PDB ID 4MHC) reveals that the hydrophobic binding pocket for Nup53^{R3} is intact (*21*). An asterisk highlights the *sc*Nup157^{NTD} mutant that abolished the interaction with *sc*Nup53. (**E**, **F**) Interaction experiments were performed with fragments of *S. cerevisiae* Nic96^{SOL} and Nup53 homologous to the crystallized Nic96^{SOL}•Nup53^{R2} complex and with mutant *sc*Nic96^{SOL} in which conserved hydrophobic residues within the Nup53 binding pocket were mutated to alanine. SEC profiles of nucleoporins are shown individually (blue and red) and after their preincubation (green). (**G**) A close-up view of the Nic96^{SOL}•Nup53^{R2} complex, same as Fig. 4C. (**H**) A close-up view of the corresponding surface of *sc*Nic96^{SOL} (PDB ID 2RFO) reveals that the hydrophobic groove that recognizes Nup53^{R2} is intact (*24*). Asterisks highlight the *sc*Nic96^{SOL} mutants that abolished the interaction with *sc*Nup53. All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

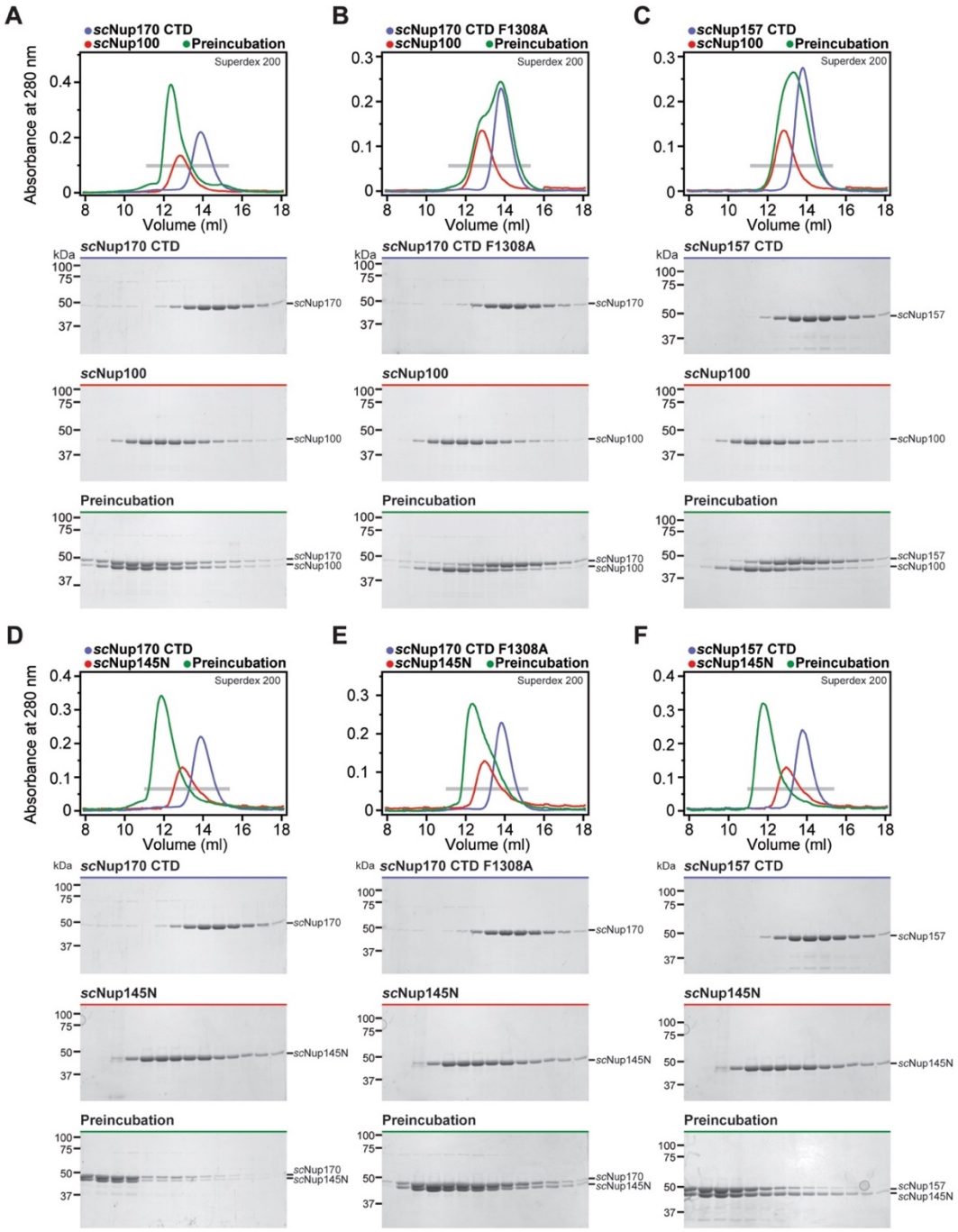


Fig. S47.

Binding pockets on *sc*Nup157 and *sc*Nup170 for the Nup145N paralogs *sc*Nup100 and *sc*Nup145N are conserved. (**A-F**) SEC profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

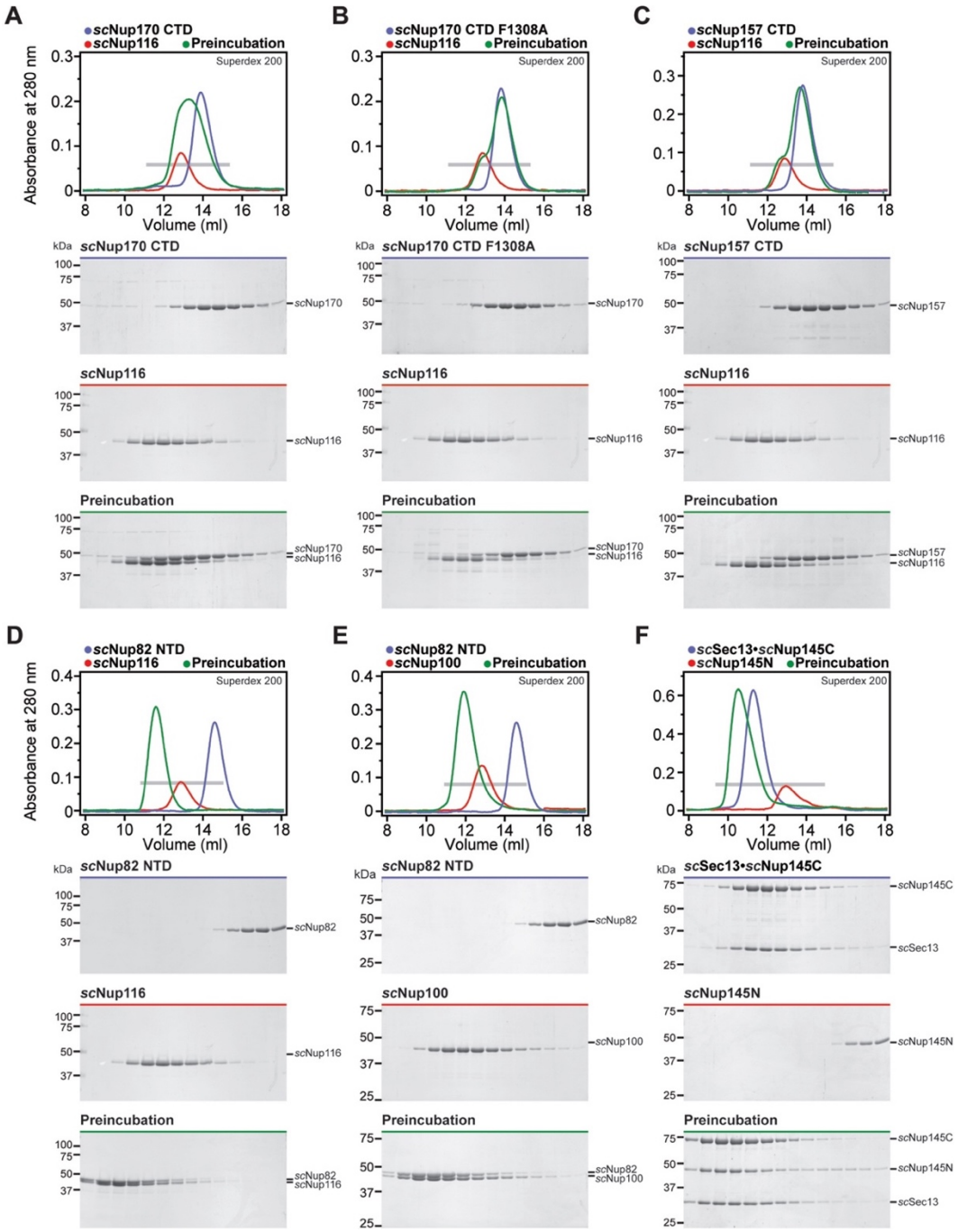


Fig. S48.

Other Nup145N interactions are conserved in the three *S. cerevisiae* paralogs *sc*Nup100, *sc*Nup145N, and *sc*Nup116, but *sc*Nup116 only binds very weakly to *sc*Nup170^{CTD} and *sc*Nup157^{CTD}. (**A-F**) SEC profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

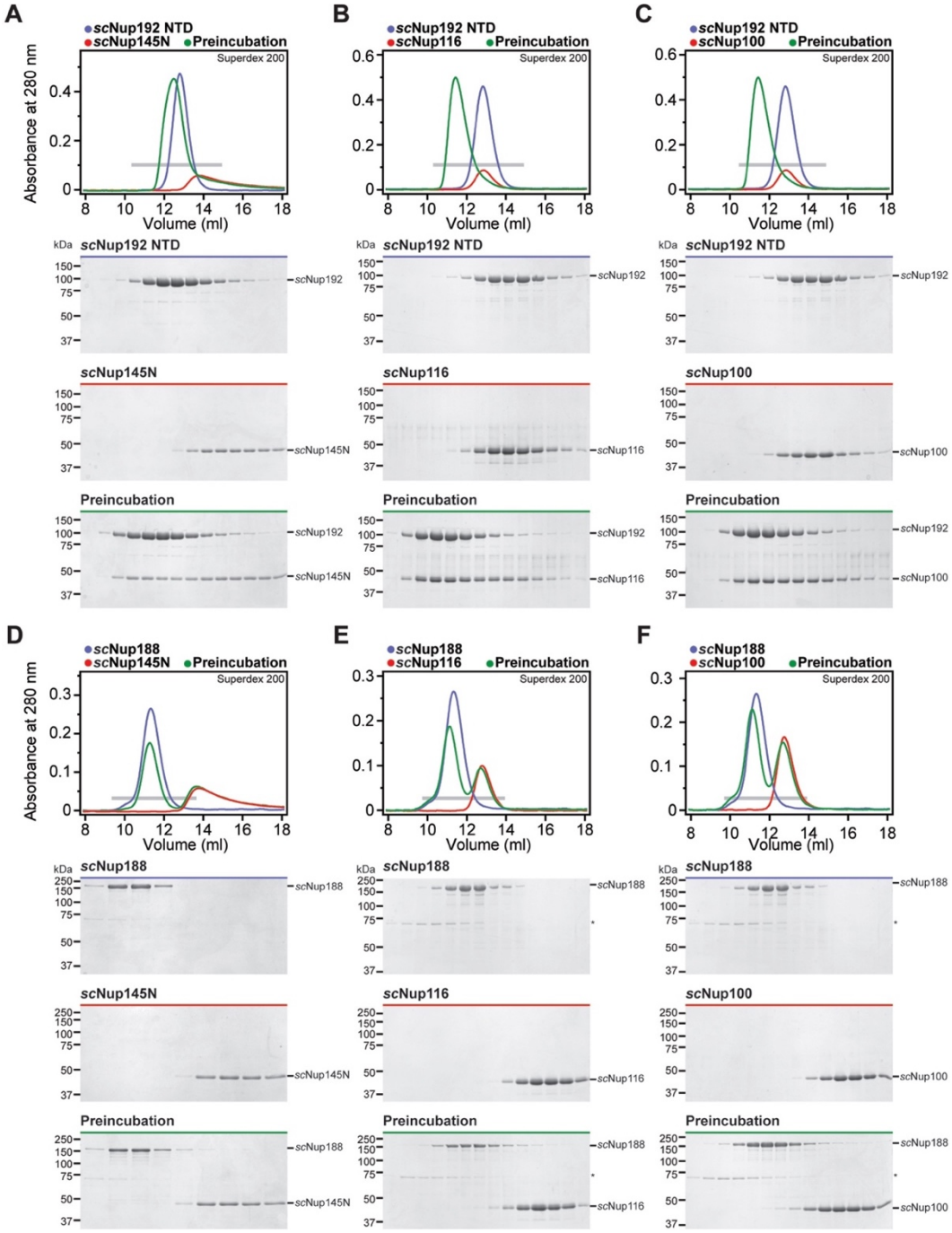


Fig. S49.

All three *S. cerevisiae* Nup145N paralogs bind to *sc*Nup192^{NTD}, but none bind to *sc*Nup188. (**A-F**) SEC profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining.

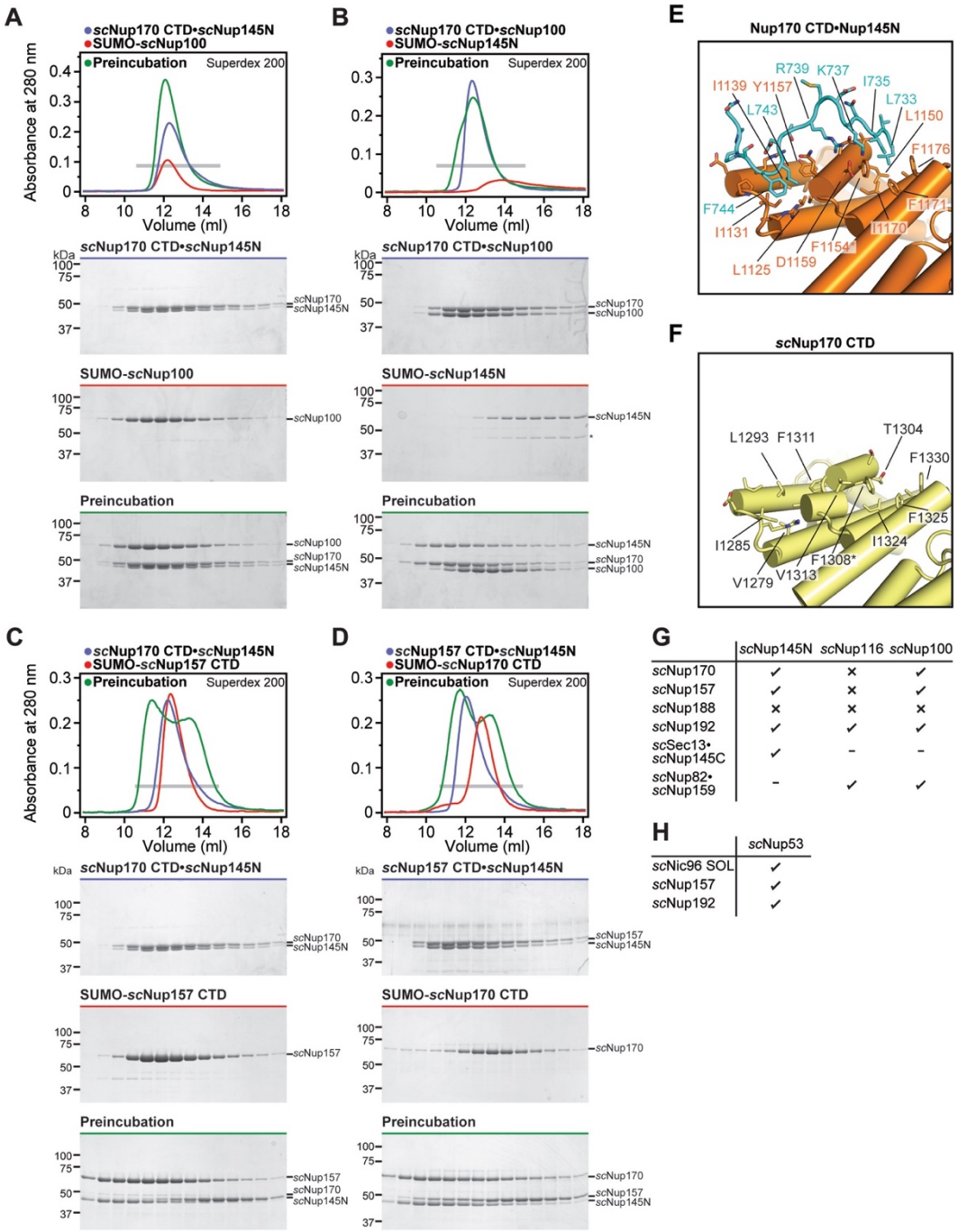


Fig. S50.

***sc*Nup145N and *sc*Nup100 bind to *sc*Nup170^{CTD} and *sc*Nup157^{CTD} in a mutually exclusive fashion.** (**A**, **B**) *sc*Nup145N and *sc*Nup100 bind to *sc*Nup170^{CTD} in a mutually exclusive fashion and *sc*Nup145N outcompetes *sc*Nup100. (**C**, **D**) *sc*Nup157^{CTD} and *sc*Nup170^{CTD} bind to *sc*Nup145N in a mutually exclusive fashion. SEC profiles of nucleoporins or nucleoporin complexes are shown individually (blue and red) and after their preincubation (green). All SEC profiles were obtained using a Superdex 200 10/300 GL column. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized

by Coomassie staining. (**E**) Close-up view of the Nup170^{CTD}•Nup145N^{R3} complex. (**F**) Close-up view of the corresponding surface in *sc*Nup170^{CTD} (PDB ID 3I5P) revealed that both hydrophobic binding pockets for Nup145N^{R3} are conserved (*23*). The conserved residue, F1308, which was mutated in the interaction experiments in figs. S47 and S48 is indicated by an asterisk. (**G**) Summary of the results of interaction experiments performed with *S. cerevisiae* Nup145N homologs. Check marks indicate complexes that can form in SEC experiments, crosses indicate complexes that do not form, and dashes indicate complexes that were not tested. (**H**) Summary of the results of interaction experiments performed with *S. cerevisiae* Nup53. Check marks indicate complexes that can form in SEC experiments, crosses indicate complexes that do not form, and dashes indicate complexes that were not tested.

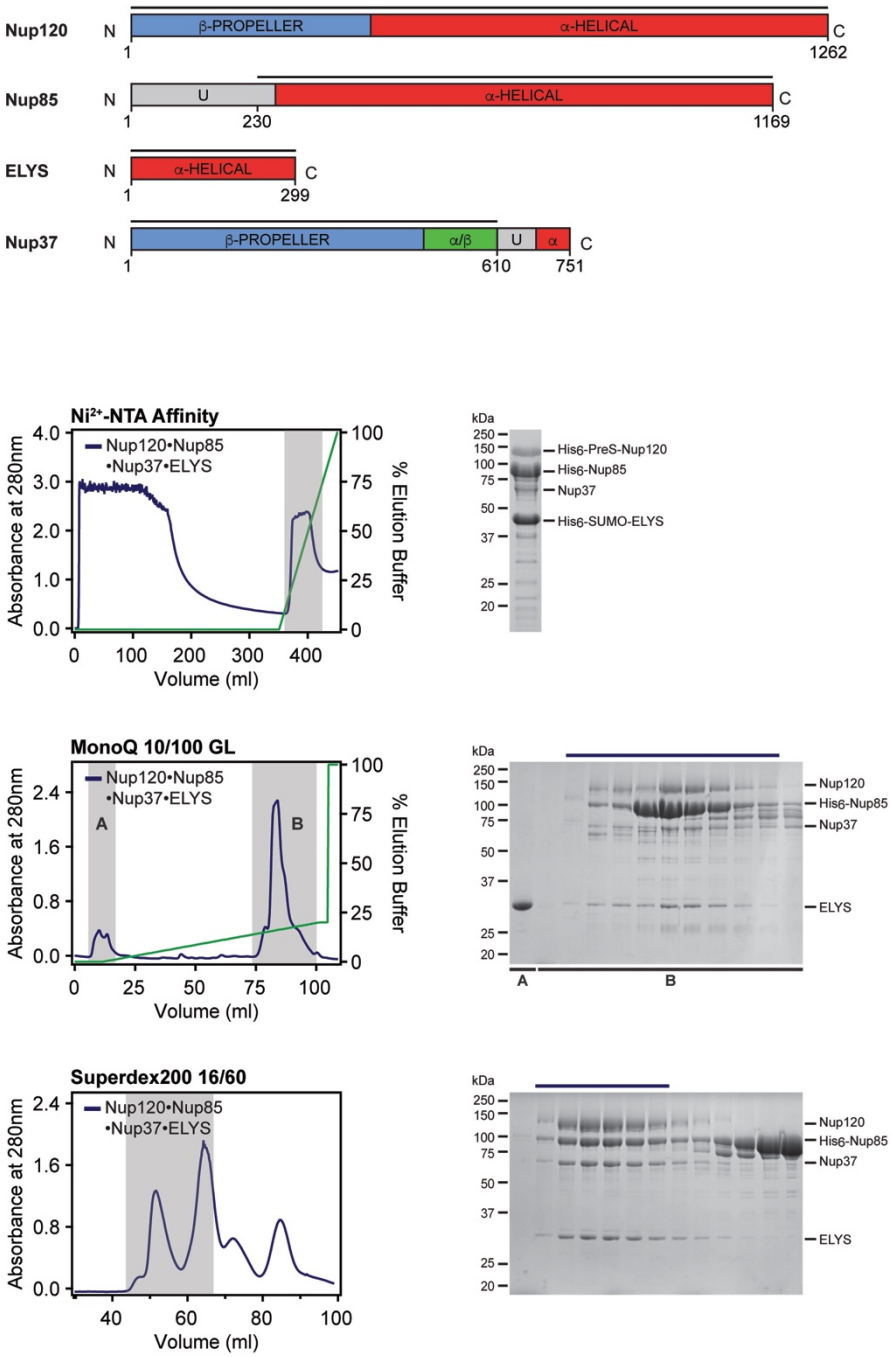


Fig. S51.

Purification protocol for Nup120•Nup37•ELYS•Nup85 hetero-tetramer. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

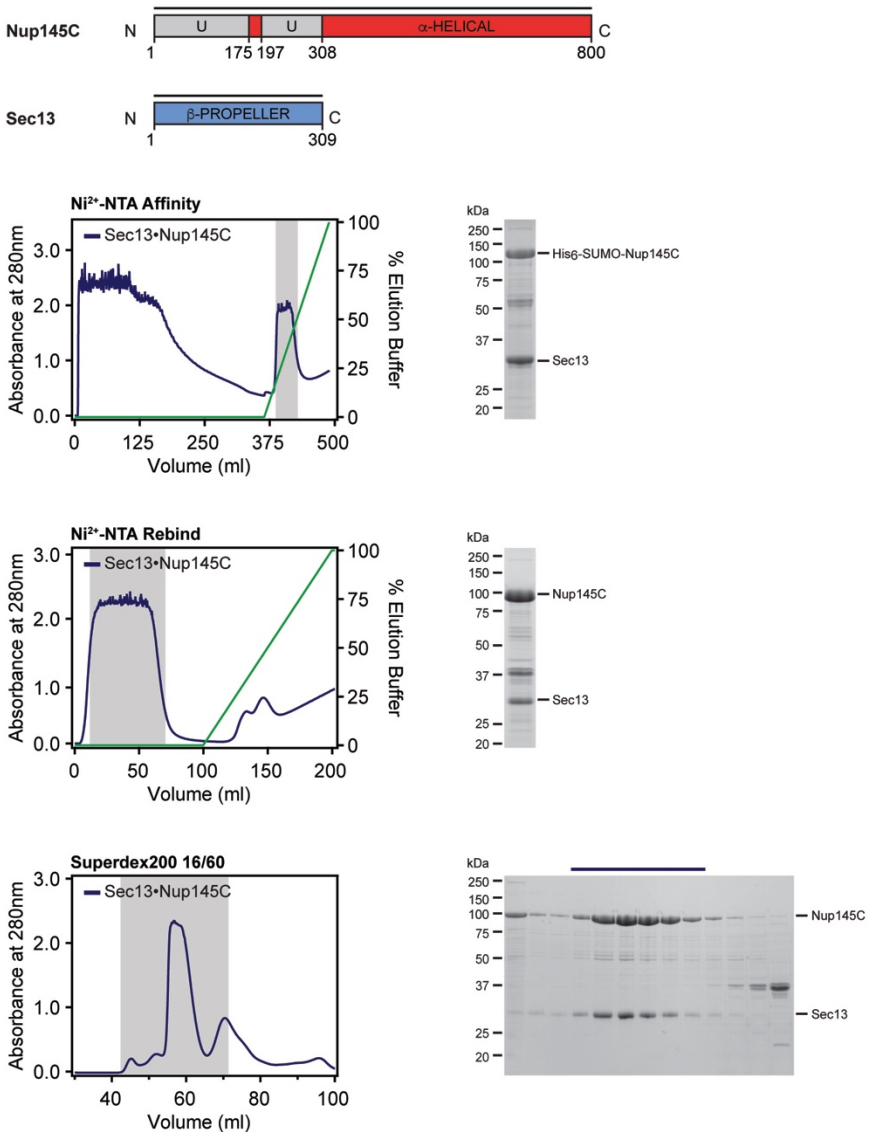


Fig. S52.

Purification protocol for the Sec13•Nup145C hetero-dimer. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

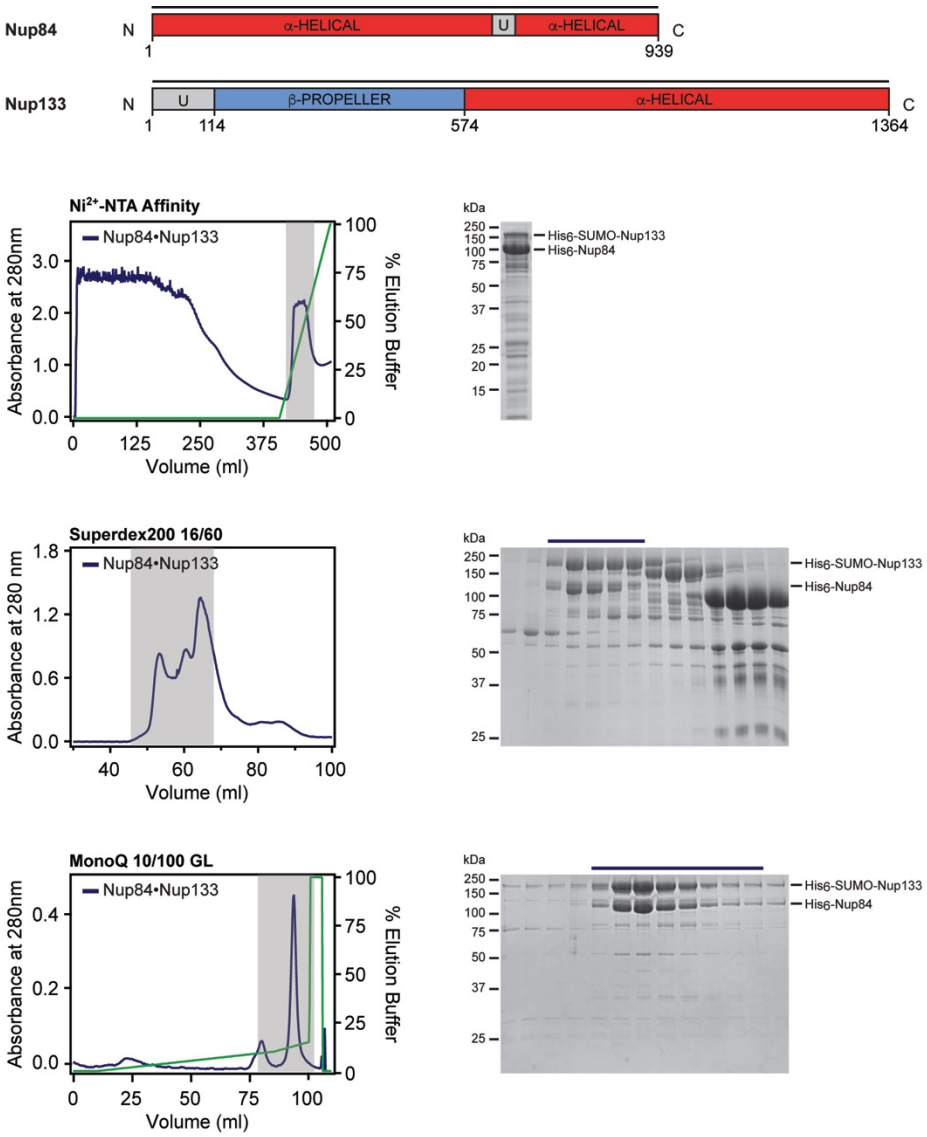


Fig. S53.

Purification protocol for the Nup84•Nup133 hetero-dimer. Domain boundaries for the purified proteins are shown above with black lines indicating the construct boundaries. Sequential steps of purification via chromatography are shown from top to bottom, with the fractions pooled for the subsequent step highlighted in grey. SDS-PAGE gels for each step of the purification are shown next to each chromatogram. Buffer conditions can be found in Table S2.

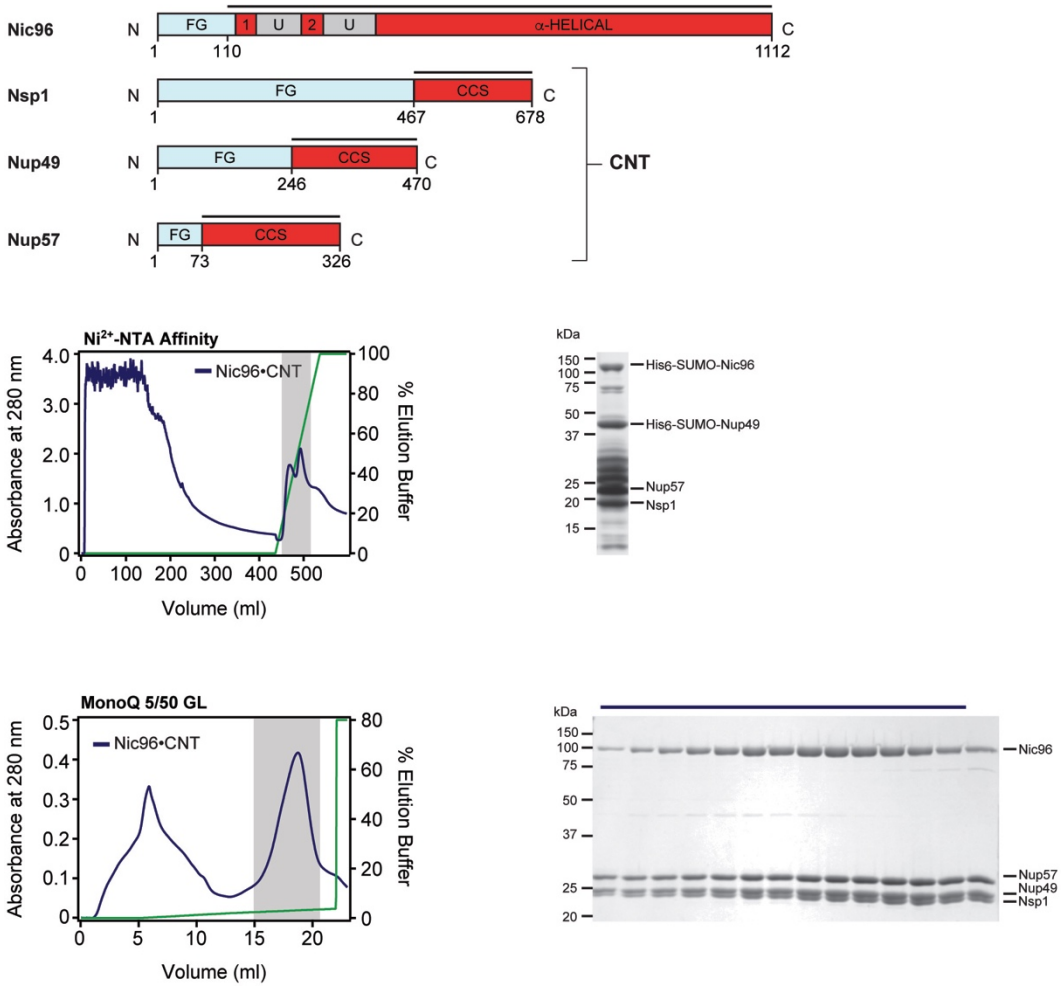


Fig. S54.

Purification protocol for the CNT•Nic96 hetero-tetramer. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

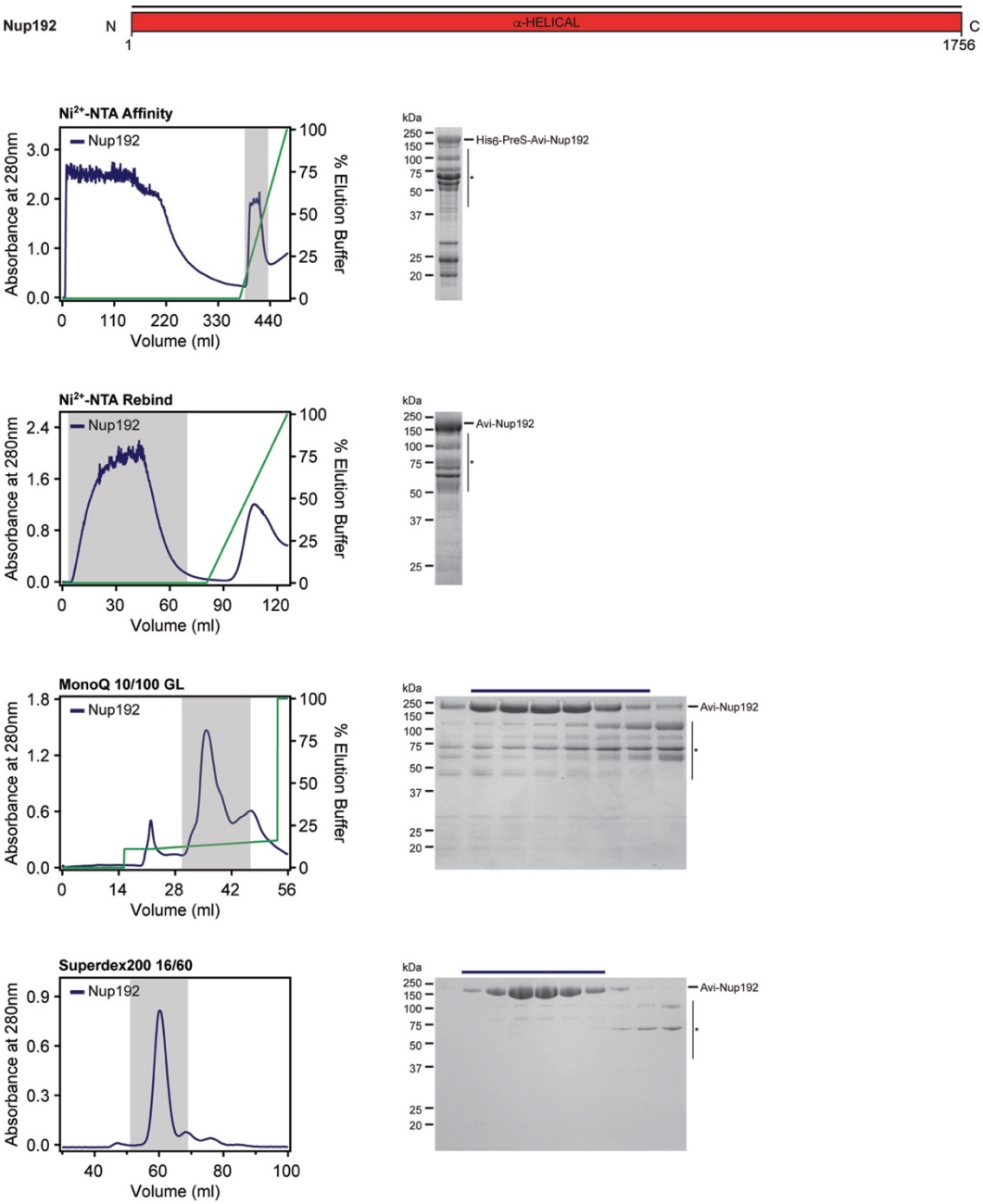


Fig. S55.

Purification protocol for Nup192. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

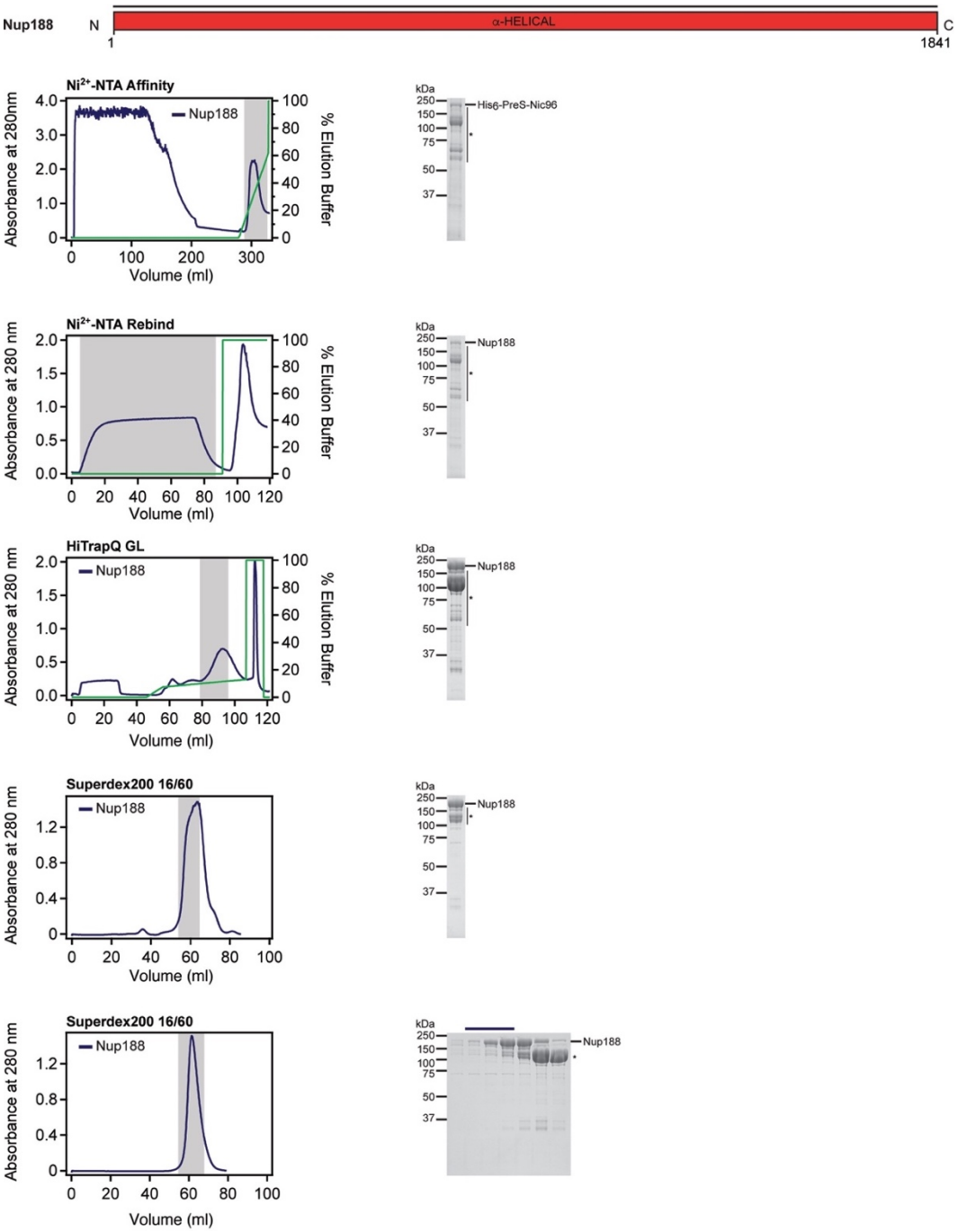


Fig. S56.

Purification protocol for Nup188. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

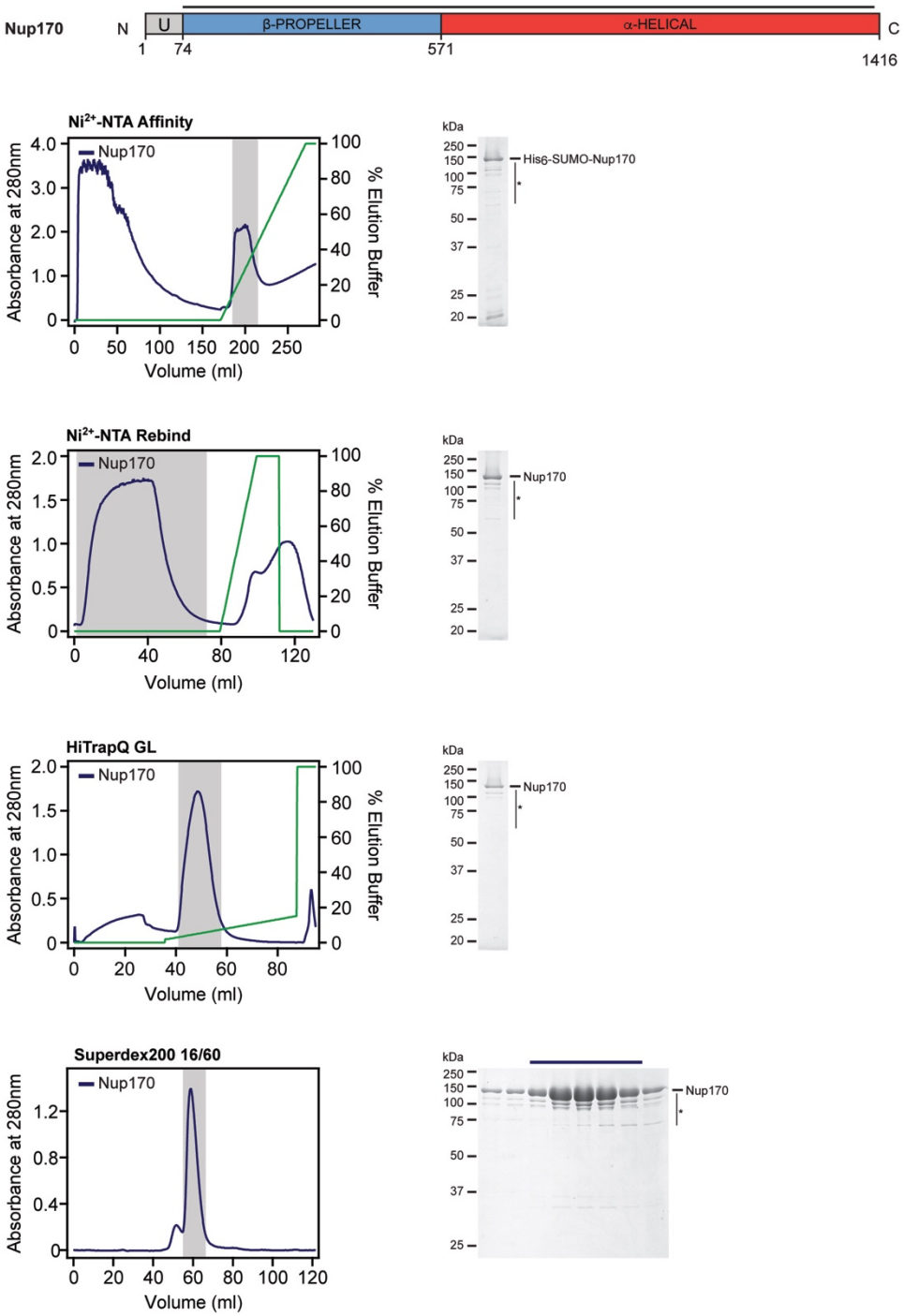


Fig. S57.

Purification protocol for Nup170. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

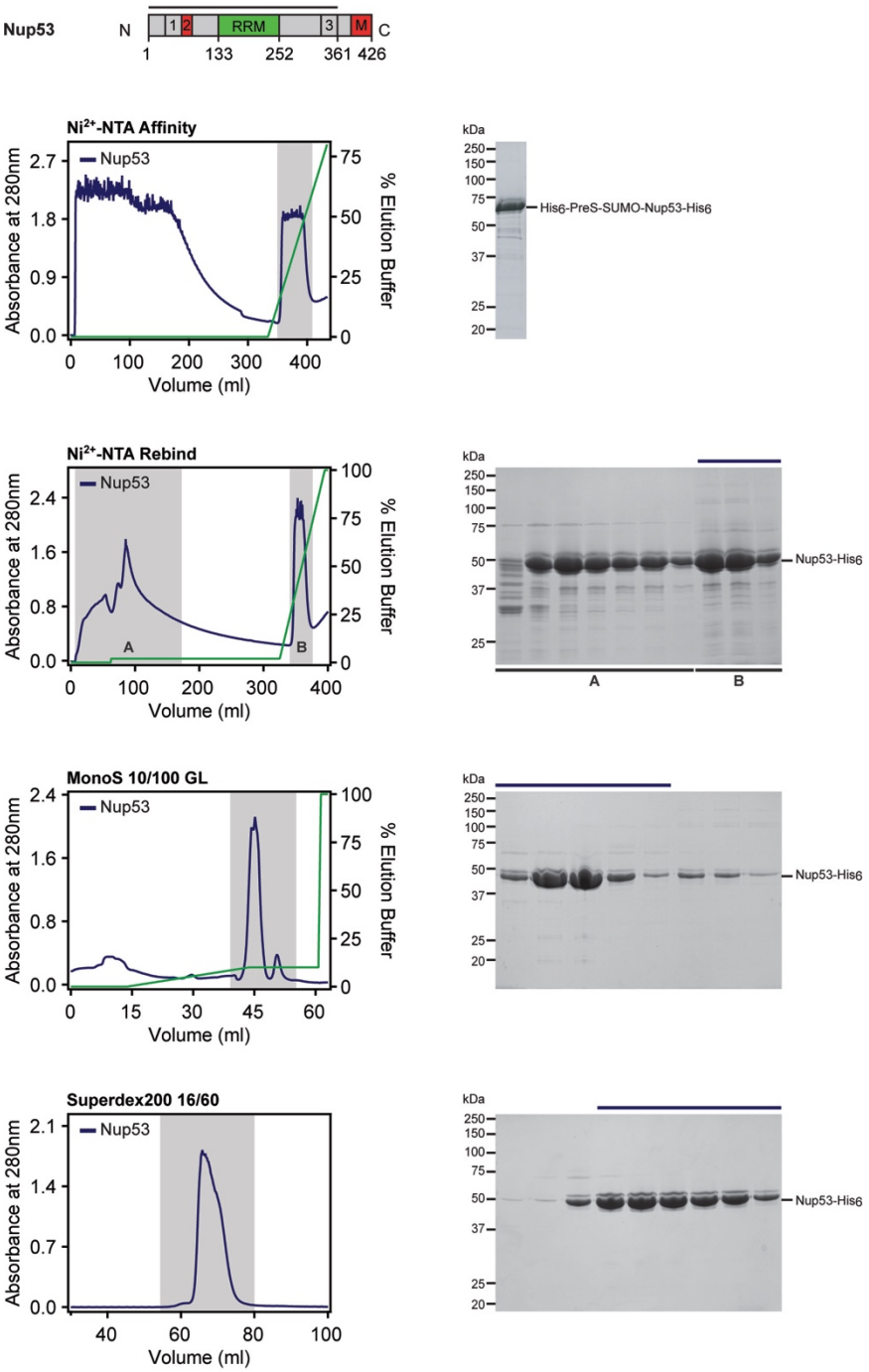


Fig. S58.

Purification protocol for Nup53. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

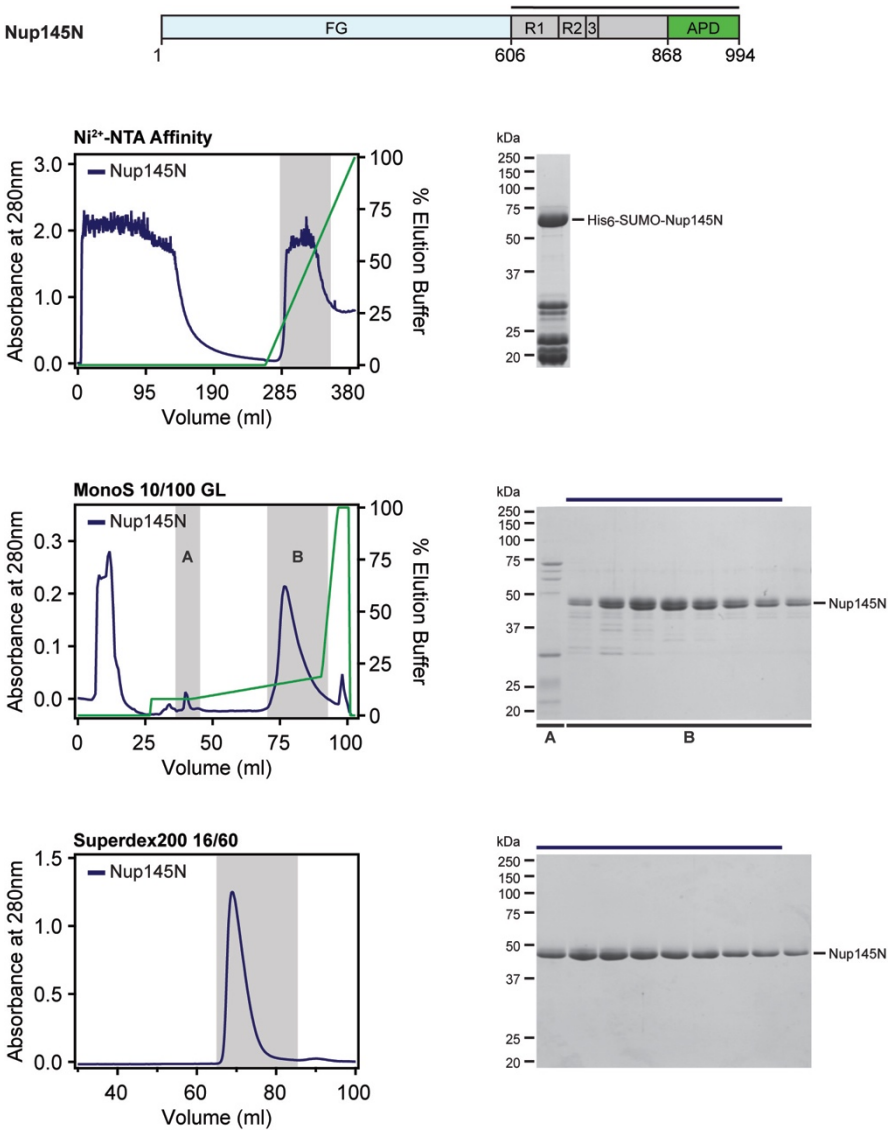


Fig. S59.

Purification protocol for Nup145N. Domain boundaries of the purified nucleoporins are shown with black lines indicating the construct boundaries. Sequential chromatography purification steps are shown from top to bottom with the employed columns indicated. Gray bars indicate fractions that were resolved on SDS-PAGE gels and visualized by Coomassie staining. Pooled fractions are indicated with a black bar above the SDS-PAGE gels. For details of the buffer conditions see Table S2.

SUPPLEMENTAL REFERENCES

1. A. Hoelz, E. W. Debler, G. Blobel, The structure of the nuclear pore complex. *Annu. Rev. Biochem.* **80**, 613-643 (2011).
2. A. Ibarra, M. W. Hetzer, Nuclear pore proteins and the control of genome functions. *Genes Dev.* **29**, 337-349 (2015).
3. V. Le Sage, A. J. Mouland, Viral subversion of the nuclear pore complex. *Viruses* **5**, 2019-2042 (2013).
4. A. Kohler, E. Hurt, Gene regulation by nucleoporins and links to cancer. *Mol. Cell* **38**, 6-15 (2010).
5. K. Zhang *et al.*, The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61 (2015).
6. A. C. Woerner *et al.*, Cytoplasmic protein aggregates interfere with nucleo-cytoplasmic transport of protein and RNA. *Science* **351**, 173-176 (2015).
7. H. M. Kaneb *et al.*, Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **24**, 1363-1373 (2015).
8. B. D. Freibaum *et al.*, GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129-133 (2015).
9. A. Jovicic *et al.*, Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* **18**, 1226-1229 (2015).
10. A. Cook, F. Bono, M. Jinek, E. Conti, Structural biology of nucleocytoplasmic transport. *Annu. Rev. Biochem.* **76**, 647-671 (2007).

11. K. Thierbach *et al.*, Protein interfaces of the conserved Nup84 complex from Chaetomium thermophilum shown by crosslinking mass spectrometry and electron microscopy. *Structure* **21**, 1672-1682 (2013).

12. K. Kelley, K. E. Knockenhauer, G. Kabachinski, T. U. Schwartz, Atomic structure of the Y complex of the nuclear pore. *Nat. Struct. Mol. Biol.* **22**, 425-431 (2015).

13. K. H. Bui, A. von Appen *et al.*, Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* **155**, 1233-1243 (2013).

14. T. Stuwe, A. R. Correia *et al.*, Architecture of the nuclear pore complex coat. *Science* **347**, 1148-1152 (2015).

15. T. Stuwe, C. J. Bley, K. Thierbach, S. Petrovic *et al.*, Architecture of the fungal nuclear pore inner ring complex. *Science* **350**, 56-64 (2015).

16. H. Chug, S. Trakhanov, B. B. Hulsmann, T. Pleiner, D. Gorlich, Crystal structure of the metazoan Nup62*Nup58*Nup54 nucleoporin complex. *Science* **350**, 106-110 (2015).

17. J. Fischer, R. Teimer, S. Amlacher, R. Kunze, E. Hurt, Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. *Nat. Struct. Mol. Biol.* **22**, 774-781 (2015).

18. N. Eisenhardt, J. Redolfi, W. Antonin, Interaction of Nup53 with Ndc1 and Nup155 is required for nuclear pore complex assembly. *J. Cell Sci.* **127**, 908-921 (2014).

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

20. T. Stuwe, D. H. Lin, L. N. Collins, E. Hurt, A. Hoelz, Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. *Proc. Natl. Acad. Sci. USA* **111**, 2530-2535 (2014).

21. H. S. Seo, B. J. Blus, N. Z. Jankovic, G. Blobel, Structure and nucleic acid binding activity of the nucleoporin Nup157. *Proc. Natl. Acad. Sci. USA* **110**, 16450-16455 (2013).

22. K. R. Andersen *et al.*, Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* **2**, e00745 (2013).

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

24. N. Schrader *et al.*, Structural basis of the nic96 subcomplex organization in the nuclear pore channel. *Mol. Cell* **29**, 46-55 (2008).

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

26. P. Sampathkumar *et al.*, Structure, dynamics, evolution, and function of a major scaffold component in the nuclear pore complex. *Structure* **21**, 560-571 (2013).

27. S. Amlacher *et al.*, Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell* **146**, 277-289 (2011).

28. A. von Appen *et al.*, In situ structural analysis of the human nuclear pore complex. *Nature*, (2015).

29. M. T. Teixeira *et al.*, Two functionally distinct domains generated by in vivo cleavage of Nup145p: a novel biogenesis pathway for nucleoporins. *EMBO J.* **16**, 5086-5097 (1997).

30. A. E. Hodel *et al.*, The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. *Mol. Cell* **10**, 347-358 (2002).

31. J. A. Killian, G. von Heijne, How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429-434 (2000).

32. G. Drin *et al.*, A general amphipathic alpha-helical motif for sensing membrane curvature. *Nat. Struct. Mol. Biol.* **14**, 138-146 (2007).

33. S. J. Kim *et al.*, Integrative structure-function mapping of the nucleoporin Nup133 suggests a conserved mechanism for membrane anchoring of the nuclear pore complex. *Mol. Cell. Proteomics* **13**, 2911-2926 (2014).

34. D. I. Fernandez, T. H. Lee, M. A. Sani, M. I. Aguilar, F. Separovic, Proline facilitates membrane insertion of the antimicrobial peptide maculatin 1.1 via surface indentation and subsequent lipid disordering. *Biophys J.* **104**, 1495-1507 (2013).

35. E. Laurell *et al.*, Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. *Cell* **144**, 539-550 (2011).

36. D. Flemming *et al.*, Analysis of the yeast nucleoporin Nup188 reveals a conserved S-like structure with similarity to karyopherins. *J. Struct. Biol.* **177**, 99-105 (2012).

37. M. Eibauer *et al.*, Structure and gating of the nuclear pore complex. *Nat. Commun.* **6**, 7532 (2015).

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

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

40. H. S. Seo *et al.*, Structural and functional analysis of Nup120 suggests ring formation of the Nup84 complex. *Proc. Natl. Acad. Sci. USA* **106**, 14281-14286 (2009).

41. C. Xu *et al.*, Crystal structure of human nuclear pore complex component NUP43. *FEBS Lett.* **589**, 3247-3253 (2015).

42. D. I. Kim *et al.*, Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci. USA* **111**, E2453-2461 (2014).

43. A. Ori *et al.*, Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol. Syst. Biol.* **9**, 648 (2013).

44. D. Devos *et al.*, Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol.* **2**, e380 (2004).

45. S. S. Katta, C. J. Smoyer, S. L. Jaspersen, Destination: inner nuclear membrane. *Trends Cell Biol.* **24**, 221-229 (2014).

46. R. Ungricht, U. Kutay, Establishment of NE asymmetry-targeting of membrane proteins to the inner nuclear membrane. *Curr. Opin. Cell Biol.* **34**, 135-141 (2015).

47. P. Li *et al.*, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012).

48. C. P. Lusk *et al.*, Nup53p is a target of two mitotic kinases, Cdk1p and Hrr25p. *Traffic* **8**, 647-660 (2007).

49. F. Alber *et al.*, The molecular architecture of the nuclear pore complex. *Nature* **450**, 695-701 (2007).

50. I. C. Berke, T. Boehmer, G. Blobel, T. U. Schwartz, Structural and functional analysis of Nup133 domains reveals modular building blocks of the nuclear pore complex. *J. Cell Biol.* **167**, 591-597 (2004).

CHAPTER 4

STRUCTURAL AND FUNCTIONAL ANALYSIS OF THE C-TERMINAL DOMAIN OF NUP358/RANBP2

This chapter was adapted from:

Daniel H. Lin, Stephan Zimmermann, Tobias Stuwe, Evelyn Stuwe, André Hoelz (2013). Structural and functional analysis of the C-terminal domain of Nup358/RanBP2, ***J. Mol. Biol.***, 425(8):1318-1329.

ABSTRACT

The nuclear pore complex (NPC) is the sole mediator of bi-directional transport between the nucleus and cytoplasm. Nup358 is a metazoan-specific nucleoporin that localizes to the cytoplasmic filaments and provides several binding sites for the mobile nucleocytoplasmic transport machinery. Here we present the crystal structure of the C-terminal domain (CTD) of Nup358 at 1.75 Å resolution. The structure reveals that the CTD adopts a cyclophilin-like fold with a non-canonical active site configuration. We determined biochemically that the CTD possesses weak peptidyl-prolyl isomerase activity and show that the active site cavity mediates a weak association with the HIV-1 capsid protein, supporting its role in viral infection. Overall, the surface is evolutionarily conserved, suggesting that the CTD serves as a protein-protein interaction platform. However, we demonstrate that the CTD is dispensable for nuclear envelope localization of Nup358, suggesting that the CTD does not interact with other nucleoporins.

INTRODUCTION

Nuclear pore complexes (NPCs) are large, proteinaceous transport organelles that perforate the nuclear envelope and are the sole mediator of bi-directional nucleocytoplasmic transport.[1] The human NPC is composed of 34 distinct proteins – 31 soluble nucleoporins (nups) and 3 integral membrane proteins of the pore membrane domain (poms). Each NPC protein occurs in multiple copies to generate a highly symmetrical structure with about 1,000 individual polypeptide chains and an estimated molecular mass of approximately 120 MDa.[1] Electron microscopy structures of the NPC have shown a doughnut-shaped symmetric core that is embedded in ~100 nm pores perforating the nuclear envelope. The symmetric NPC core is decorated with different nucleoporin subsets that form distinct filamentous structures protruding from the nucleoplasmic and cytoplasmic faces.[1] On the nucleoplasmic face, the filaments bundle into a basket-like structure and provide binding sites for various components of other macromolecular machineries that carry out diverse cellular processes at the nuclear envelope, including chromatin remodeling, transcription, and DNA repair.[1] On the cytoplasmic face, the filaments are extended and reach far into the cytoplasm, providing binding sites for a plethora of transport complexes and the mRNA export machinery.[1]

In humans, the cytoplasmic filaments are primarily composed of three proteins: Nup88, Nup214/CAN, and Nup358/RANBP2 (referred to as Nup358 in the following text).[1] Nup358 is a unique component of the metazoan NPC and is the largest known nucleoporin, composed of 3,224 residues in humans.[2,3] Nup358 can be divided into several domains: an N-terminal TPR domain, an α-helical region, four Ran-binding domains, eight tandem zinc fingers, a SUMO E3 ligase domain, and a C-terminal domain that displays sequence homology to cyclophilins (Fig. 1a).[2,3] Thus far, only the SUMO E3 ligase domain, two Ran-binding domains and the N-terminal TPR domain have been structurally characterized.[4-7] Nup358 provides key interaction sites for various components of the nucleocytoplasmic

transport machinery including Ran, SUMOylated RanGAP1, the NXF1•p15 mRNA export heterodimer, and the export karyopherin CRM1.[1,7] Aditionally, we have recently shown that the N-terminal domain binds to single-stranded RNA.[8]

Apart from its essential function in nucleocytoplasmic transport, Nup358 is involved in many other important cellular processes. During mitosis, Nup358 localizes to kinetochores and facilitates spindle formation as well as chromosome congression and segregation.[9,10] A striking link exists between the cytoplasmic filament nups, their binding partners, and neoplastic disease. The expression of Nup358 is upregulated in plasma cells from patients with multiple myeloma, and mutations and fusions of the Nup358 gene have been associated with inflammatory myofibroblastic tumors.[11,12] Moreover, mutations in Nup358 have been linked to heightened susceptibility of otherwise healthy children to acute necrotizing encephalopathy following common viral infections such as influenza. [13] Recently, Nup358 has also been implicated in the delivery and integration of the genomic material of HIV-1 to the genome of terminally differentiated non-dividing cells.[14-17]

In order to gain insight into structure and function of Nup358, we have determined the crystal structure of the C-terminal domain (CTD) at 1.75 Å resolution. Nup358^{CTD} adopts a canonical cyclophilin fold, but displays substantial alterations of the active site cavity. Surprisingly, we find biochemically that the observed structural changes fail to render the enzyme inactive and that the active site cavity retains weak peptidyl-prolyl isomerase activity. Moreover, we demonstrate that the active site cavity also facilitates binding to the HIV-1 capsid protein, as previously shown for cyclophilin A. In addition to the active sitecavity, Nup358^{CTD} features three extensive evolutionarily conserved surface patches, suggesting that the domain also functions as a protein-protein interaction platform. However, we determined by immunofluorescence microscopy that the CTD is dispensable for nuclear envelope localization of Nup358 *in vivo,* suggesting that it does not interact with other

nucleoporins. Together, these data suggest that Nup358^{CTD} is exposed at the tips of the cytoplasmic filaments of the NPC to facilitate binding to physiological substrates and could potentially be hijacked by the pre-integration complex of HIV-1 to facilitate nuclear import and genome delivery.

RESULTS

Structure determination

Based on a sequence alignment between human Nup358 and cyclophilin A (CypA), we generated an expression construct encompassing residues 3062 to 3224 that we termed the Nup358 C-terminal domain (CTD). $Nup358^{CTD}$ is monomeric in solution as determined by multiangle light scattering coupled to size exclusion chromatography with a measured molecular mass of 17.9 kDa (theoretical 18.4 kDa) (Fig. S1). $Nup358^{CTD}$ crystallized in the monoclinic space group $P2_1$ with six molecules in the asymmetric unit. The structure of $Nup358^{CTD}$ was solved by molecular replacement and the final model was refined to 1.75 Å resolution with R_{work} and R_{free} values of 12.2 % and 15.8 %, respectively. For details of the data collection and refinement statistics see Table 1.

Structural overview

Overall, $Nup358^{CTD}$ is a globular domain with a diameter of approximately 30 Å. $Nup358^{CTD}$ displays the canonical cyclophilin fold of an 8-stranded β-barrel (β1 to β8) that is flanked on three sides by an α-helix (αA to αC), as first observed in CypA (Fig. 1b,c).[18] The core of the domain is extensively decorated by several long and well-ordered loops that generate a large surface area. An ~20 Å long and ~10 Å wide surface groove is formed at the intersection between β-strands β3, β4, β5, and β6, the α-helix αB, and the loop linking β5 and β6. This groove is utilized for substrate binding and catalysis in CypA and other cyclophilins.[18]

In order to identify surfaces of $Nup358^{CTD}$ that might be functionally important, we aligned sequences of $Nup358^{CTD}$ from a diverse set of species (Fig. S2), and mapped the sequence conservation onto the surface of $Nup358^{CTD}$. A comparison of the conservation and electrostatic potential surfaces revealed that $Nup358^{CTD}$ possesses three evolutionarily

conserved surface areas: (1) A groove in the surface that forms the active site in other cyclophilins, (2) an extensive hydrophobic patch on the back of the domain that is generated by hydrophobic substitutions which are not present in other cyclophilins, and (3) an acidic patch on the top of the domain (Fig. 2).

The active site of Nup358^{CTD}

The 17 cyclophilin-like domains present in human proteins are very well conserved in both sequence and structure, especially at the active site, where there are only a handful of substitutions at residues that are catalytically important.[19] In CypA, the model enzyme for studying cyclophilin peptidyl-prolyl isomerization activity, the active site cavity is generated by approximately fifteen residues, which in turn accommodate four to six residues of a linear peptide.[18] The mechanism of peptidyl-prolyl isomerization by cyclophilins remains unclear, but there are several features that are conserved across family members.

Hydrophobic residues form the bottom of the cavity that accommodates a substrate *trans*- proline and several residues in the cavity form polar contacts with the peptide backbone of the substrate peptide.[19] Specifically, a critical arginine (Arg55 in CypA) that is invariant across cyclophilins forms key hydrogen bonds with the isomerized amide bond and the C- terminal backbone carbonyl.[19] Nup358^{CTD} shares ~60 % sequence identity with human CypA and the crystal structures of CypA and Nup358^{CTD} superimpose with a root-mean-square deviation of ~0.4 Å over 130 Cα atoms (Fig. 3).[20] Of the fifteen residues that contribute to the putative active site in Nup358^{CTD}, nine are identical with their corresponding positions in CypA (Fig. 3). In particular, the residues that form the majority of the contacts between CypA and its substrates, His3186, Leu3182, Gln3171, and Phe3120, as well as the critical arginine, Arg3115, are conserved. However, other critical residues in the CypA active site that make key contacts with the substrate (Trp121, Phe113, Met61, and Ala103) are altered

in the putative Nup358^{CTD} active site (His3181, Val3173, Val3121, and Gln3163). Specifically, the imidazole ring of His3181 in Nup358^{CTD} is still positioned to act as a hydrogen bond donor for the substrate backbone, analogous to Trp121 in CypA, but would be unable to form the additional hydrophobic contacts. Moreover, the substitutions of the bulky hydrophobic side chains Phe113 and Met61 for the smaller side chains Val3173 and Val3121 substantially reduce the extent to which Nup358^{CTD} can form hydrophobic contacts with substrates, and the substitution of Gln3163 for Ala103 introduces a larger side chain where the substrate normally packs closely to the β-carbon of Ala103 (Fig. 3).

Our conservation analysis reveals that most of the putative Nup358^{CTD} active site residues are nearly invariant across various vertebrates. However, whereas Nup358^{CTD} possesses a cyclophilin-like fold, the putative active site of Nup358^{CTD} displays marked differences from canonical cyclophilins, raising the possibility that Nup358^{CTD} either has an altered peptidyl-prolyl isomerase substrate specificity or is catalytically inactive.

Catalytic activity

In order to determine whether Nup358^{CTD} is an active peptidyl-prolyl isomerase, we used a chymotrypsin-coupled spectrophotometric assay on the isomerization of Suc-Ala-Ala-Pro-Phe-2,4-difluoroanilide, a synthetic peptide that mimics CypA substrates.[21] We found that Nup358^{CTD} indeed possessed peptidyl-prolyl isomerase activity, but with a much lower efficiency than CypA. The Michaelis-Menten constants K_m and k_{cat} of Nup358^{CTD} were determined to be 304.6 μM and 72 s^{-}, compared to 309.9 μM and 119,913 s^{-1} for CypA, respectively (Fig. 4a,b).

To confirm that the weak enzymatic activity of Nup358^{CTD} was in fact mediated by the putative active site, we generated the V3173W mutant that introduces the bulky aromatic

residue tryptophan directly into the active site cavity. The behavior of the Nup358^{CTD} V1373W on a gel filtration column was indistinguishable from that of the wild-type protein, indicating that the mutation did not affect proper protein folding. We found that the Nup358^{CTD} V3173W mutant failed to accelerate *cis*/*trans* isomerization over the background rate (Fig. 4c), consistent with previous results that the analogous F113W mutation in CypA dramatically reduces enzymatic activity.[22] Our data establish that Nup358^{CTD} is indeed an active peptidyl-prolyl isomerase, but with substantially reduced enzymatic activity, supporting the results of our structural analysis of the active site cavity.

HIV-1 capsid protein interaction

Interactions between cyclophilins and the HIV-1 capsid protein are a critical part of the viral life cycle. CypA binds to the HIV-1 capsid protein (HIV-1^{CA}) and is specifically incorporated into the viral capsid, preventing Ref-1 mediated restriction of HIV-1 infection, and thus is essential for HIV-1 infectivity.[23-25] The crystal structure of the CypA•HIV-1^{CA} heterodimer revealed that CypA recognizes and binds a proline-rich loop of HIV-1^{CA} with its active site (Fig. 5).[26] Recent studies have proposed that Nup358 plays a role in the delivery and genomic integration of the genetic material of HIV-1 to the nucleus by facilitating an interaction to HIV-1^{CA} via its C-terminal domain.[15-17]

To characterize the interaction between HIV-1^{CA} and Nup358^{CTD}, we employed a size exclusion chromatography binding assay. Whereas HIV-1^{CA} and CypA formed a stable stoichiometric 1:1 complex (Fig. 6a), we found the interaction between HIV-1^{CA} and Nup358^{CTD} to be weaker (Fig. 6b). To determine whether the interaction was mediated by the Nup358^{CTD} active site, we tested whether the V3173W mutant is capable for forming a complex with HIV-1^{CA}. Strikingly, the weak interaction observed between Nup358^{CTD} and HIV-1^{CA} was completely abolished in the Nup358^{CTD} V3173W mutant (Fig. 6c). To confirm

the weak interaction detected by size exclusion chromatography, we additionally employed isothermal titration calorimetry. Although the interaction strength is at the limit for reliable quantification of thermodynamic parameters (K_d > 200), we were able to observe a significant interaction between Nup358^{CTD} and HIV-1^{CA}. In agreement with our size exclusion chromatography assay, we did not observe any interaction between Nup358^{CTD} V3173W and HIV-1^{CA} (Fig. S3).

The results of the isothermal titration calorimetry experiments are consistent with the relatively minor peak shift observed for the interaction between HIV-1^{CA} and Nup358^{CTD} compared to the shift observed for the complex formed by HIV-1^{CA} and CypA. Thus, Nup358^{CTD} binds the HIV-1 capsid protein with its active site, but it does so with a reduced binding affinity, consistent with the observed structural alterations of the active site.

In vivo localization assay

Nup358 localizes to kinetochores after the breakdown of the nuclear envelope, but it is a stable component of the cytoplasmic filaments during interphase.[2,3,9] To determine whether Nup358^{CTD} is involved in the attachment of Nup358 to NPCs, we performed a series of immunofluorescence localization experiments in HEK293 cells. Both the HA-tagged full-length Nup358 and a C-terminal truncation that lacks the CTD (Nup358$^{\Delta CTD}$) yielded nuclear envelope rim staining that coincides with staining by mAb414, an antibody that recognizes a group of phenylalanine-glycine (FG)-repeat-containing nucleoporins at the nuclear rim.[27] In contrast, the staining of a construct containing only the HA-tagged Nup358^{CTD} failed to enrich at the nuclear rim (Fig. 7). These data establish that the CTD is not required for nuclear envelope localization and dispensable for Nup358 anchoring to the cytoplasmic face of the NPC.

CONCLUSIONS

We determined the crystal structure of the C-terminal domain of Nup358, a unique component of the metazoan NPC, and established biochemically that the domain possesses peptidyl-prolyl isomerase activity in addition to its previously described E3 ligase activity.[28] We show biochemically that Nup358^{CTD} interacts with the HIV-1 capsid protein. Additionally, analysis of the evolutionary conservation of the Nup358^{CTD} surface suggests that the domain also functions as a protein-protein interaction platform. However, it is unlikely that the physiological binding partners are other nucleoporins as Nup358^{CTD} does not localize to the nuclear envelope and is dispensable for nuclear envelope localization of Nup358 *in vivo*.

Despite many years of research on cyclophilins, the precise catalytic mechanism by which these enzymes catalyze proline isomerization remains poorly understood. Deeper mechanistic insights remain elusive partly because of sequence variation within the active site and a lack of physiological substrates. The active site cavity of Nup358^{CTD} displays similarity to those found in the cyclophilin-like domains of cyclophilin G and NK-tumor recognition protein, distinct from "classical" cyclophilin active sites.[29] Further experiments with this subset of cyclophilins are required to see whether their distinct active site configuration recognizes a unique set of substrates.

What role could a peptidyl-prolyl isomerase have at the periphery of the NPC? Besides a role in protein folding, peptidyl-prolyl isomerases also participate in a wide range of signaling pathways and Nup358^{CTD} may act in a similar way on an unknown substrate. Alternatively, it is possible that Nup358^{CTD} may also function as a protein-protein interaction module at the tips of the cytoplasmic filaments, perhaps serving as a surveillance domain to attract cargo molecules far away from the transport channel of the NPC. We speculate that this functionality may also serve to specifically recognize a thus unknown binding partner,

leading to increased concentration at the cytoplasmic face of the NPC, similar to the Ran-binding domains of Nup358. Given the increased local concentration of Nup358 that is generated by its localization to the NPC, even relatively weak interactions may play physiogically important roles. Viral proteins such as the HIV-1 capsid protein may hijack such functionality to facilitate their nuclear import potentially by promoting uncoating or interactions with transport machinery as suggested previously.[17] Our studies provide a structural and biochemical platform to dissect these interactions.

METHODS

Protein expression and purification. DNA fragments encoding residues 3062-3224 from human Nup358, residues 1-165 of human Cyclophilin A, and residues 1-146 of the HIV-1 capsid protein were cloned into a modified pET28a vector, containing an N-terminal hexahistidine-tag followed by a PreScission protease cleavage site using NdeI and NotI restriction sites.[30] $Nup358^{CTD}$ V3173W was generated by site directed mutagenesis and confirmed by DNA sequencing. All proteins were expressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) in LB media. Expression was induced at an OD_{600} of approximately 0.6 with 0.5 mM IPTG at 18 °C for 16 hours. Cells were harvested by centrifugation and resuspended in a buffer containing 25 mM TRIS, pH 8.0, 500 mM NaCl, 20 mM imidazole, 5 mM β- mercaptoethanol (β-ME) and complete EDTA-free protease inhibitor cocktail (Roche).

For purification, the cells were lysed with a cell disruptor (Avestin) and DNAse I (Roche) was added to the lysate before centrifugation at 30,000 x g for 1 hour. The supernatant was filtered with a 0.45 μm filter (Millipore) and loaded onto a Ni-NTA column (GE Healthcare) equilibrated in buffer A (20 mM TRIS, pH 8.0, 500 mM NaCl, 20 mM imidazole, and 5 mM β-ME). Protein was eluted with a linear gradient of buffer B (20 mM TRIS, pH 8.0, 500 mM NaCl, 500 mM imidazole, and 5 mM β-ME). Fractions were pooled and loaded onto a HiPrep Desalting column (GE Healthcare) equilibrated in 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM β-ME. PreScission protease was added to the eluate and the protein was cleaved at 4 °C for 16 hours. Digested protein was loaded onto a Ni-NTA column and collected in the flowthrough, concentrated in a centrifugal filter (Millipore), and loaded on a Superdex 75 16/60 column (GE Healthcare) equilibrated in buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 10 mM DTT. Protein-containing fractions were

pooled and concentrated to 65 mg/ml for crystallization studies and 30 mg/ml for biochemistry studies.

Protein crystallization and data collection. Protein crystallization was carried out at 21 °C in hanging drops consisting of 1.0 µl protein solution and 1.0 µl reservoir solution. Crystals were grown in 0.1 M TRIS-HCl, pH 8.3, 0.2 M NaCl, and 0.9 M K/Na tartrate with a protein concentration of 65 mg/ml. Crystals were cryoprotected in the reservoir solution supplemented with 20 % glycerol and flash frozen in liquid nitrogen. X-ray diffraction data was collected at 100 K at Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL).

Structure determination and refinement. X-ray diffraction data was processed with the HKL-2000 denzo/scalepack package. A model of human Cyclophilin A, PDB code 19MC was used as a search model for molecular replacement with Phaser.[20,31] Subsequent model building and refinement using anisotropic B-factors were performed with Coot and PHENIX, respectively.[32,33] The final model contains residues 3,062 to 3,224 and was refined to 1.75 Å resolution with R_{work} and R_{free} values of 12.2 % and 15.8 %, respectively. The stereochemical quality of the model was assessed with PROCHECK and MolProbity.[34,35] For details of the data collection and refinement statistics see Table 1.

Kinetic analysis. Enzyme kinetics were analyzed following a previously established chymotrypsin-coupled spectrophotometric assay.[21] Peptidyl-prolyl isomerase activity was measured with a spectrophotometer (Shimadzu) at 9 °C using the cleavage of Suc-Ala-Ala-Pro-Phe-2,4- difluoroanilide (Bachem), which led to an increase of absorbance at 290 nm.

CypA and Nup358^{CTD} were assayed at final concentrations of 20 nM and 2 μM, respectively. The final concentration of α-chymotrypsin (Sigma) was 150 μg/ml. Suc-Ala-Ala-Pro-Phe-2,4- difluoroanilide (Bachem) was dissolved in a solution of Tetrafluoroethylene (TFE) and 500 mM LiCl and the reaction was started by the addition of Suc-Ala-Ala-Pro-Phe-2,4-difluoroanilide to the cuvette for final concentrations that ranged from 50 μM to 500 μM. The reaction was carried out in 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM β-ME.

Analytical size-exclusion chromatography. HIV-1^{CA} was mixed in approximately equimolar amounts with Nup358^{CTD}, Nup358^{CTD}, or CypA and incubated on ice for approximately 30 minutes before injection onto a Superdex 75 10/300 GL gel-filtration column (GE Healthcare) equilibrated in 20 mM TRIS, pH 8.0, 100 mM NaCl, and 5 mM β-ME. Complex formation was monitored by the shift of the respective protein peaks in the chromatogram and by SDS-PAGE of the protein-containing fractions.

ITC measurements. ITC measurements were performed at 23 °C using an ITC200 calorimeter (GE Healthcare). Protein samples were stored in a buffer containing 20 mM TRIS, pH 8.0, 100 mM NaCl, and 4 mM β-ME. To perform the titration, 2.5 μL of 3.7 mM HIV-1^{CA} was injected into 200 μL of 350 μM Nup358^{CTD} or 310 μM Nup358^{CTD} V3173W every 240 s. The heat generated from dilution was subtracted for baseline correction. Baseline corrected data were analyzed with Origin 7.0 software with MicroCal add-ons. All experiments were performed at least three times.

Immunofluorescence microscopy. For immunofluorescence, Nup358 and Nup358 fragments (residues 1-3224, 1-145, and 146-3224) were inserted into the pCMV-HA vector (Clontech) using SalI and NotI restriction sites. The resulting Nup358 fragments contain an

N-terminal HA-tag. HEK293T cells were grown on poly-L-lysine coated slides (Sigma) until ~50 % confluency.

Transfection was performed with TransIT-LT1 transfection reagent (Mirus) according to the manufacturer's instructions. After 48 hours, the medium was removed and cells were washed in phosphate buffered saline (PBS) and fixed in PBS, supplemented with 2 % (w/v) formaldehyde for 5 min at room temperature. After two washes with PBS, the cells were permeabilized with PBS containing 0.1 % (v/v) TritonX-100 (Sigma-Aldrich) for 10 minutes at room temperature. The cells were then washed in PBS and blocked in PBS, supplemented with 10 % (v/v) Fetal Calf Serum (FCS) for 20 min at room temperature. For nuclear envelope staining, the cells were incubated with a 1:500 dilution of the monoclonal antibody mAb414 (Abcam) in PBS buffer, supplemented with 0.1 % (w/v) saponin and 10 % (v/v) FCS for 16 hours at 4 °C. Secondary antibody incubation was performed with a 1:3000 dilution of Alexa Fluor 568 goat anti mouse (Invitrogen) in PBS, supplemented with 0.1 % (w/v) saponin and 10 % (v/v) FCS for 1 hour at room temperature, followed by three washes with PBS. For detection of the HA-tagged proteins, the cells were incubated with a monoclonal antibody anti-HA Fluor488 conjugate antibody (Invitrogen) for 1 hour at room temperature, washed three times with PBS, and mounted onto coverslips with ProLong Gold Antifade reagent with DAPI (Invitrogen). Slides were examined with an inverted fluorescence microscopy on a Carl Zeiss AxioImagerZ.1 equipped with a Hamamatzu camera.

Illustration and figures. The sequence alignment of Nup358^{CTD} was generated using ClustalX and colored with Alscript.[36,37] The scientific illustrations were generated using PyMOL (Schrödinger, LLC; www.pymol.org), Igor (WaveMetrics), and Prism (GraphPad).

The electrostatic potential was calculated with Adaptive Poisson-Boltzmann Solver (APBS).[38]

ACKNOWLEDGEMENTS

We thank Alina Patke and the members of the Hoelz laboratory for critical reading of the manuscript, Kuang Shen for helpful discussions regarding kinetics analysis, and David Baltimore and Christopher Hill for providing cDNAs for the HIV-1 gag-pol and human Cyclophilin A, respectively. We thank Jens Kaiser and the scientific staff of SSRL beamline 12-2 for their support with X-ray diffraction measurements. We acknowledge the Gordon and Betty Moore Foundation for their support of the Molecular Observatory at the California Institute of Technology. The operations at the SSRL are supported by the Department of Energy and by the National Institutes of Health. DHL was supported by a National Research Service Award (T32GM07616) from the National Institute of General Medical Sciences. ES was supported by a Boehringer Ingelheim Fonds predoctoral fellowship. AH was supported by the Albert Wyrick V Scholar Award of the V Foundation for Cancer Research, the 54th Mallinckrodt Scholar Award of the Edward Mallinckrodt, Jr. Foundation, and a Kimmel Scholar Award of the Sidney Kimmel Foundation for Cancer Research.

REFERENCES

1. Hoelz A, Debler EW, Blobel G. The structure of the nuclear pore complex. Annu. Rev. Biochem. 2011; 80:613–643.

2. Wu J, Matunis MJ, Kraemer D, Blobel G, Coutavas E. Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. J. Biol. Chem. 1995; 270:14209–14213.

3. Yokoyama N, et al. A giant nucleopore protein that binds Ran/TC4. Nature. 1995; 376:184–188.

4. Vetter IR, Nowak C, Nishimoto T, Kuhlmann J, Wittinghofer A. Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. Nature. 1999; 398:39–46.

5. Reverter D, Lima CD. Insights into E3 ligase activity revealed by a SUMORanGAP1-Ubc9-Nup358 complex. Nature. 2005; 435:687–692.

6. Geyer JP, et al. Solution structure of the Ran-binding domain 2 of RanBP2 and its interaction with the C terminus of Ran. J. Mol. Biol. 2005; 348:711–725.

7. Kassube SA, et al. Crystal Structure of the N-Terminal Domain of Nup358/RanBP2. J. Mol. Biol. 2012; 423:752–765.

8. Cook A, Bono F, Jinek M, Conti E. Structural biology of nucleocytoplasmic transport. Annu. Rev. Biochem. 2007; 76:647–671.

9. Salina D, Enarson P, Rattner JB, Burke B. Nup358 integrates nuclear envelope breakdown with kinetochore assembly. J. Cell Biol. 2003; 162:991–1001.

10. Joseph J, Tan S-H, Karpova TS, McNally JG, Dasso M. SUMO-1 targets

RanGAP1 to kinetochores and mitotic spindles. J. Cell. Biol. 2002; 156:595–602.

11. Felix RS, et al. SAGE analysis highlights the importance of p53csv, ddx5, mapkapk2 and ranbp2 to multiple myeloma tumorigenesis. Cell. 2009; 278:41–48.

12. Ma Z, et al. Fusion of ALK to the Ran-binding protein 2 (RANBP2) gene in inflammatory myofibroblastic tumor. Genes Chromosomes Cancer. 2003; 37:98–105.

13. Neilson DE, et al. Infection-triggered familial or recurrent cases of acute necrotizing encephalopathy caused by mutations in a component of the nuclear pore, RANBP2. Am. J. Hum. Genet. 2009; 84:44–51.

14. Copeland AM, Newcomb WW, Brown JC. Herpes simplex virus replication: roles of viral proteins and nucleoporins in capsid-nucleus attachment. J. Virol. 2009; 83:1660–1668.

15. Ocwieja KE, et al. HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. PLoS Pathog. 2011; 7:e1001313.

16. Zhang R, Mehla R, Chauhan A. Perturbation of host nuclear membrane component RanBP2 impairs the nuclear import of human immunodeficiency virus -1 preintegration complex (DNA). PLoS ONE. 2010; 5:e15620.

17. Schaller T, et al. HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. PLoS Pathog. 2011; 7:e1002439.

18. Ke HM, Zydowsky LD, Liu J, Walsh CT. Crystal structure of recombinant human T-cell cyclophilin A at 2.5 A resolution. Proc. Natl. Acad. Sci. U.S.A. 1991; 88:9483–9487.

19. Davis TL, et al. Structural and Biochemical Characterization of the Human Cyclophilin Family of Peptidyl-Prolyl Isomerases. PLoS Biol. 2010; 8:e1000439.

20. Howard BR, Vajdos FF, Li S, Sundquist WI, Hill CP. Structural insights into the catalytic mechanism of cyclophilin A. Nat. Struct. Mol. Biol. 2003; 10:475–481.

21. Kofron JL, Kuzmic P, Kishore V, Colon-Bonilla E, Rich DH. Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. Biochemistry. 1991; 30:6127–6134.

22. Bosco DA, et al. Dissecting the microscopic steps of the cyclophilin A enzymatic cycle on the biological HIV-1 capsid substrate by NMR. J. Mol. Biol. 2010; 403:723–738.

23. Thali M, et al. Functional association of cyclophilin A with HIV-1 virions. Nature. 1994; 372:363– 365.

24. Franke EK, Yuan HE, Luban J. Specific incorporation of cyclophilin A into HIV-1 virions. Nature. 1994; 372:359–362.

25. Towers GJ, et al. Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. Nat. Med. 2003; 9:1138–1143.

26. Gamble TR, et al. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell. 1996; 87:1285–1294.

27. Davis LI, Blobel G. Identification and characterization of a nuclear pore complex protein. Cell. 1986; 45:699–709.

28. Pichler A, Gast A, Seeler JS, Dejean A, Melchior F. The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity. Cell. 2002; 108:109–120.

29. Galat A. Variations of sequences and amino acid compositions of proteins that

sustain their biological functions: An analysis of the cyclophilin family of proteins. Arch. Biochem. Biophys. 1999; 371:149–162.

30. Hoelz A, Nairn AC, Kuriyan J. Crystal Structure of a Tetradecameric Assembly of the Association Domain of Ca2+/Calmodulin-Dependent Kinase II. Mol. Cell. 2003; 11:1241–1251.

31. McCoy AJ, et al. Phaser crystallographic software. J. Appl. Crystallogr. 2007; 40:658–674.

32. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:486–501.

33. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:213–221.

34. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK - a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 1993; 26:283–291.

35. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 2007; 35:W375–83.

36. Larkin MA, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948.

37. Barton GJ. ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. 1993; 6:37– 40.

38. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U.S.A. 2001; 98:10037–10041

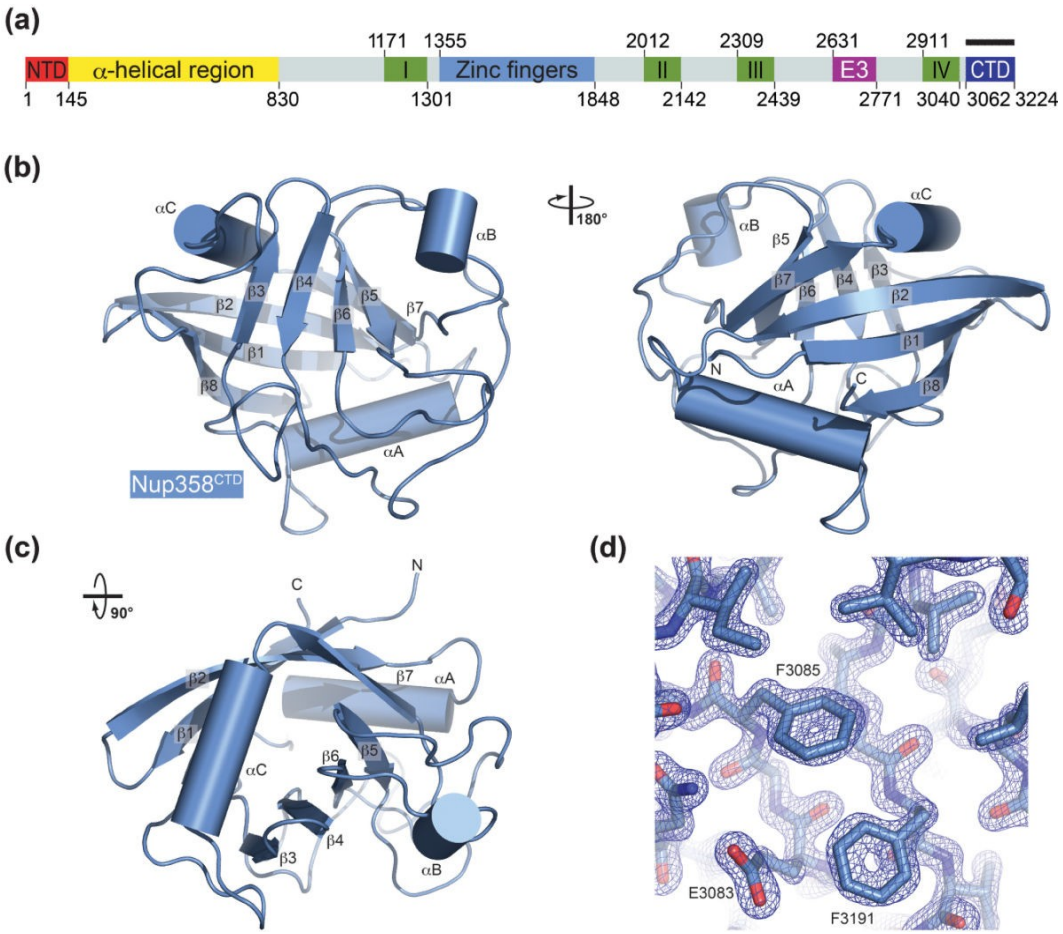


Fig. 1.

Domain organization and structure of Nup358^{CTD}. (a) Domain organization of human Nup358. Domain boundaries are indicated by residue numbers. The bar above the domain structure denotes the crystallized fragment. I, II, III, and IV, Ran binding domains; NTD, N-terminal domain; CTD, C-terminal domain; E3, E3 ligase domain. (b) Cartoon representation of Nup358^{CTD} with a view rotated 180° along the vertical axis shown on the right. (c) Cartoon representation of Nup358^{CTD} rotated 90° along the horizontal axis from above. (d) Representative $2|F_O|-|F_C|$ electron density map contoured at 1.5 σ.

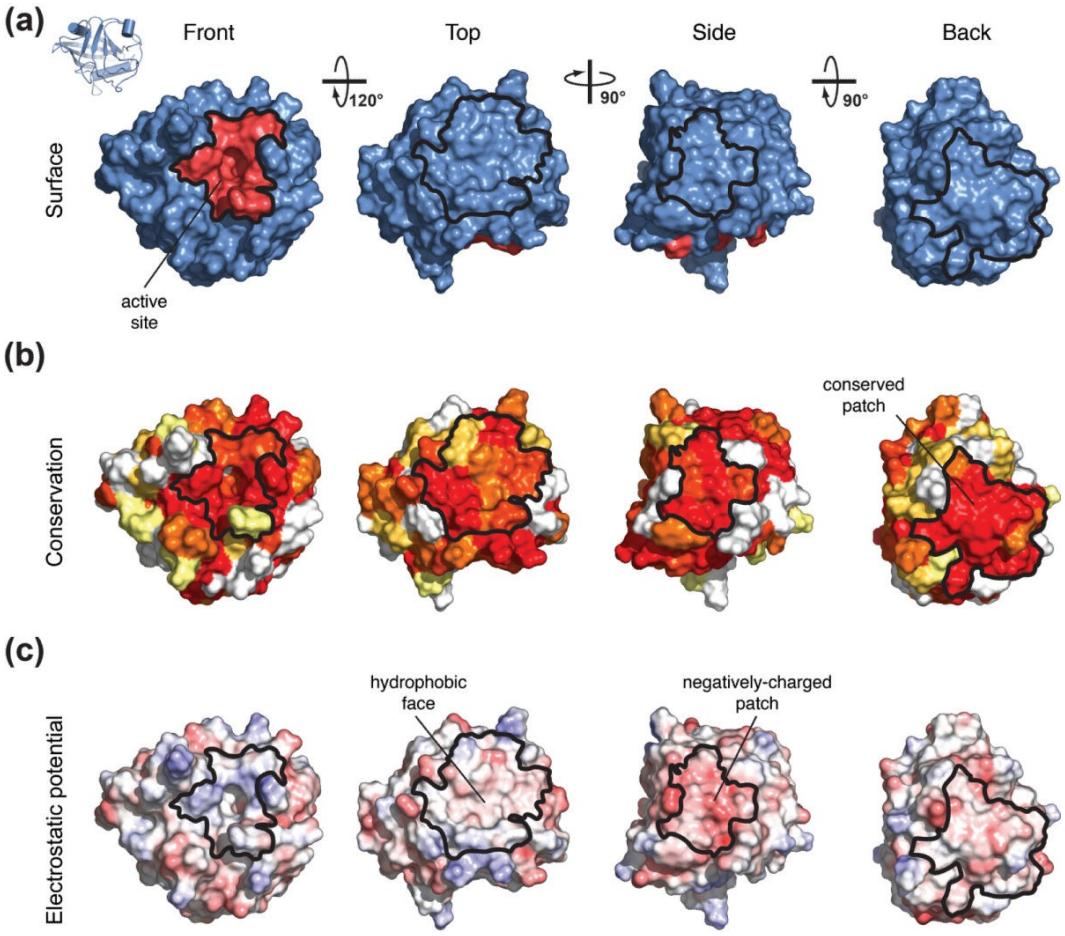


Fig. 2.
Surface properties of Nup358^{CTD}. (a) Surface representation of Nup358^{CTD}, with the active site colored in red. (b) Surface representation colored according to sequence identity based on a multi-species sequence alignment (Fig. 2). The identity at each position is mapped onto the surface and is shaded in a color gradient from white (60 % less than 60 % identity) to red (100 % identity). (c) Surface representation colored according to electrostatic potential from -10 k_BT/e (red) to +10 k_BT/e (blue).

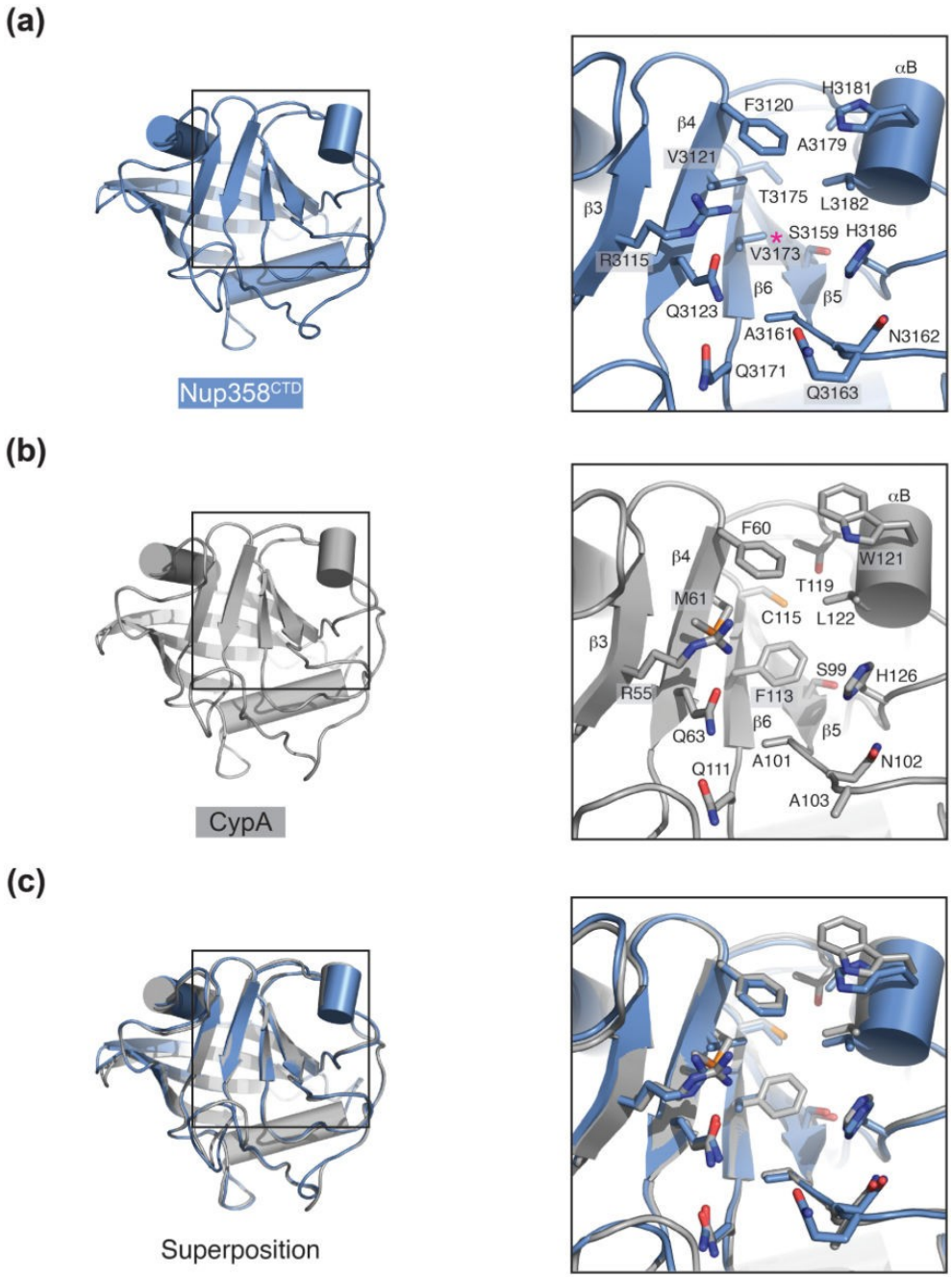


Fig. 3.

Comparison of Nup358^{CTD} and Cyclophilin A active sites. (a) Detailed view of the Nup358^{CTD} active site. (b) Detailed view of the Cyclophilin A active site (PDB Code 1M9C). (c) Overlay of the active sites from Nup358^{CTD} and Cyclophilin A. Critical active site residues are shown in stick representation, and the Cα-traces are shown in coil representation, according to the coloring scheme in A. The orientation of all active site residues is identical.

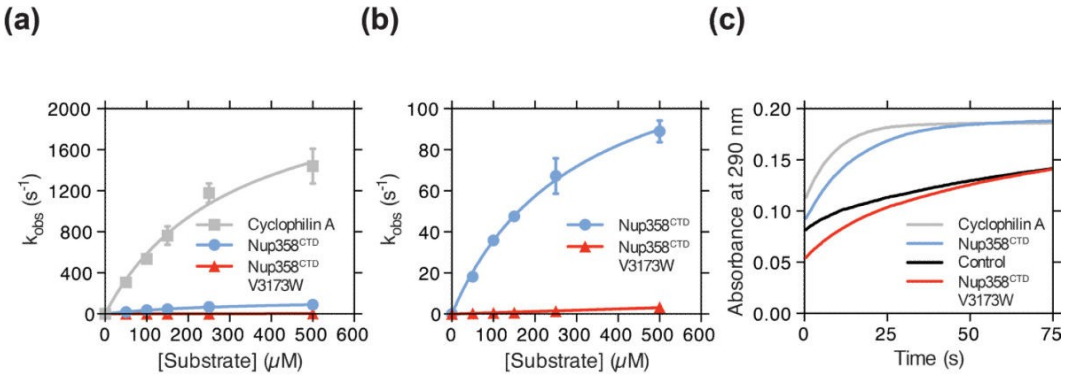


Fig. 4.

Nup358^{CTD} possesses peptidyl-prolyl isomerase activity. (a) Michaelis-Menten plot of the peptidyl-prolyl isomerization of Suc-Ala-Ala-Pro-Phe-2,4-difluoroanilide by Cyclophilin A, Nup358^{CTD}, and Nup358^{CTD} V3173W. (b) Michaelis-Menten plot of the peptidyl-prolyl isomerization of Suc-Ala-Ala-Pro-Phe-2,4-difluoroanilide by Nup358^{CTD} and Nup358^{CTD} V3173W. Note the different scale of the y-axis from panel (a). (c) Representative time-course traces of the peptidyl-prolyl isomerization of Suc-Ala-Ala-Pro-Phe-2,4-difluoroanilide by Cyclophilin A, Nup358^{CTD}, Nup358^{CTD} V3173W, and in the absence of an enzyme.

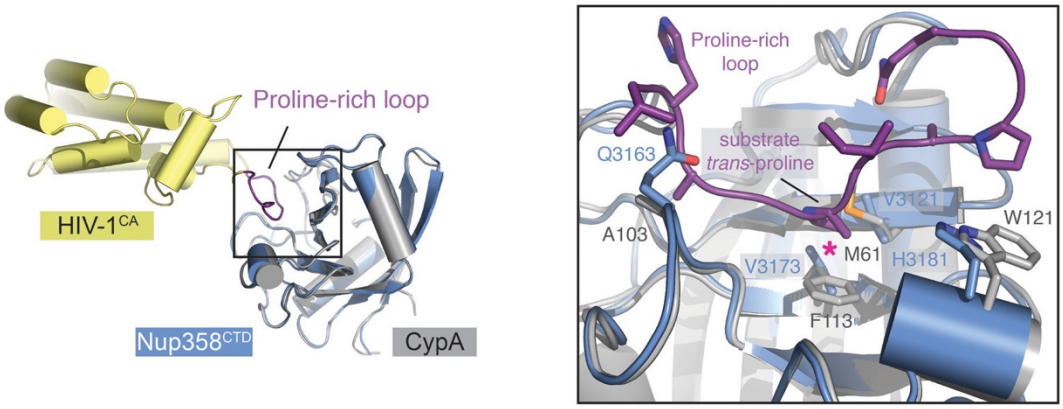


Fig. 5.
Structural comparison of Nup358CTD to the Cyclophilin A•HIV-1^{CA} complex. The structure of Nup358^{CTD} overlaid on the structure of the Cyclophilin A•HIV-1^{CA} complex (PDB code 1M9C). The right panel is a close-up view of the interaction with the HIV-1^{CA} loop rotated 90° along the vertical axis from the left panel. Note the clash between the Nup358^{CTD} Q3163 and the HIV-1^{CA} proline-rich loop.

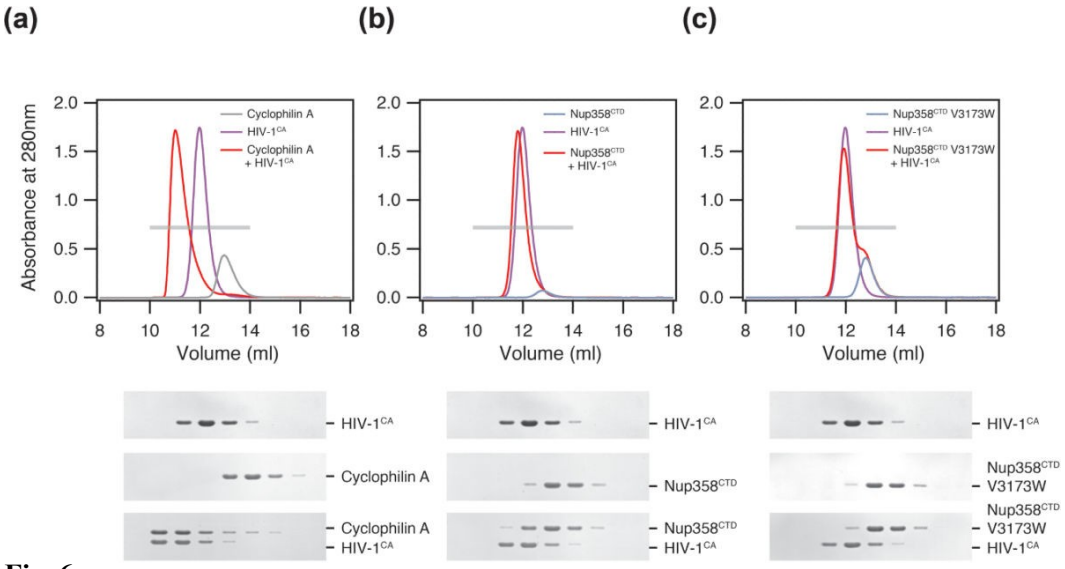


Fig. 6.
Nup358^{CTD} binds weakly to the HIV-1 capsid protein. (a-c) Size exclusion chromatography interaction analysis of HIV-1^{CA} with (a) Cyclophilin A, (b) Nup358^{CTD}, and (c) Nup358^{CTD} V3173W. The analyzed fractions are indicated in gel-filtration profile by a grey bar. Notably, experiments were carried out with identical protein concentrations, the different peak height of the wild-type Nup358^{CTD} is a result of a lower absorbance coefficient.

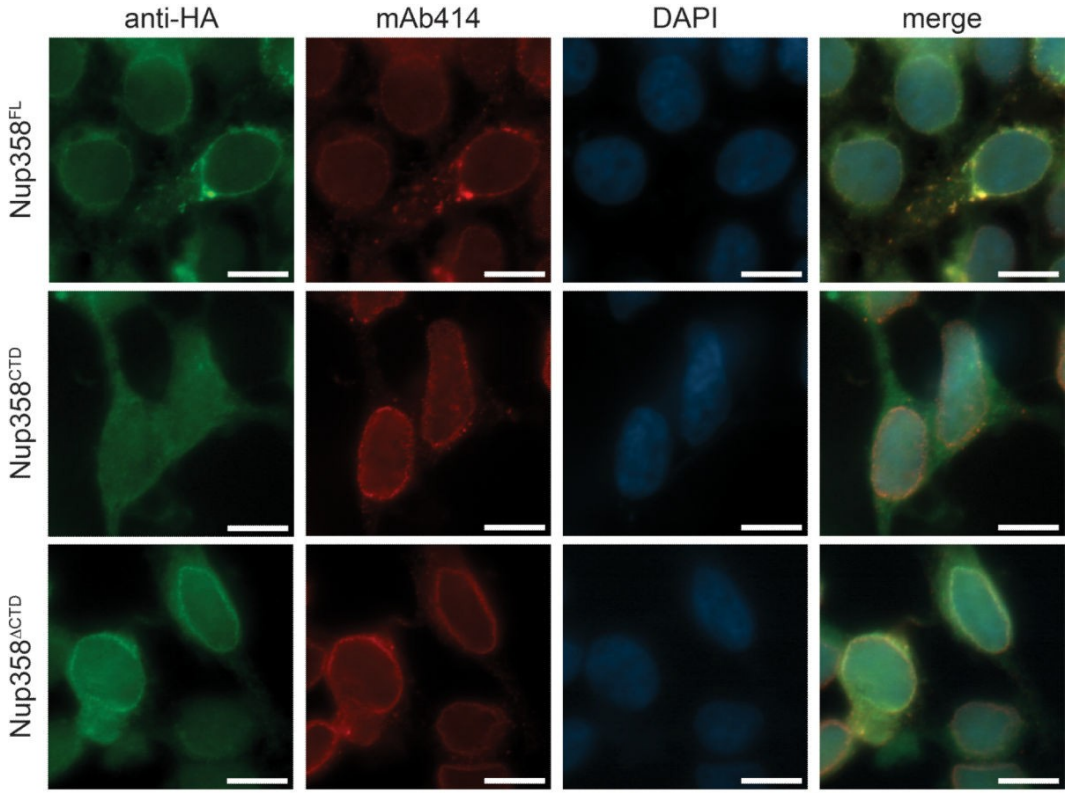


Fig. 7.

Nup358^{CTD} is dispensable for nuclear envelope localization. Nup358^{CTD} and Nup358 fragments carrying a N-terminal HA-tag were transiently transfected into HEK293T cells and analyzed by fluorescence microscopy. HA-tagged Nup358 protein localization was detected with an α-HA antibody (green). The monoclonal α-mAb414 antibody (red) and DAPI (blue) were used as a reference for nuclear envelope and nucleus staining, respectively. The right panel shows the merged images.

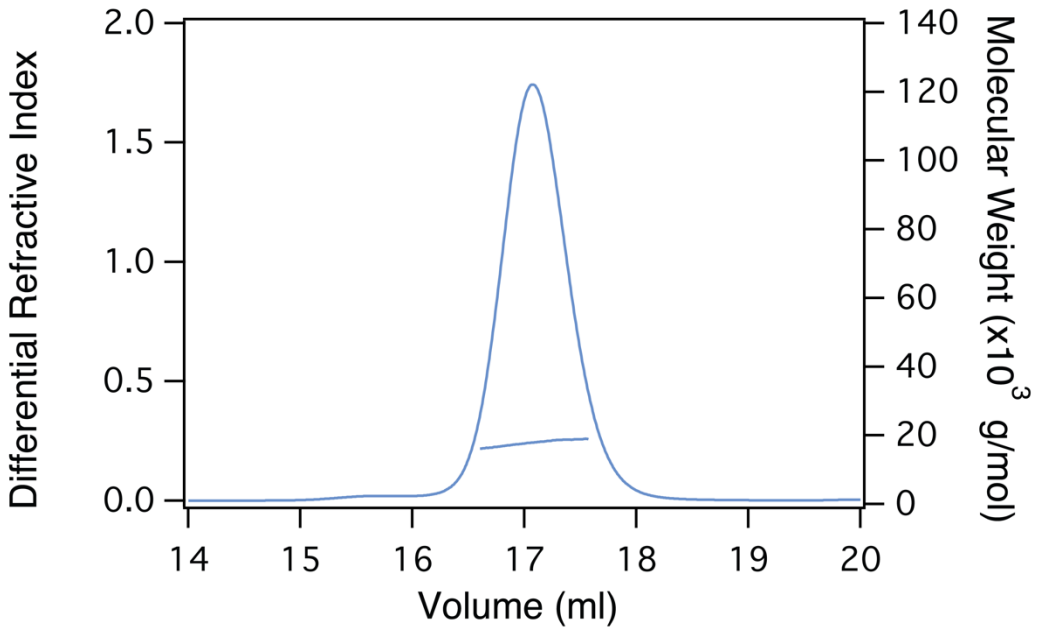


Fig. S1.
Multi-angle light scattering analysis of Nup358^{CTD}. The differential refractive index and determined molecular weight are plotted against elution volume off of a Superdex 75 10/300 GL gel-filtration column.

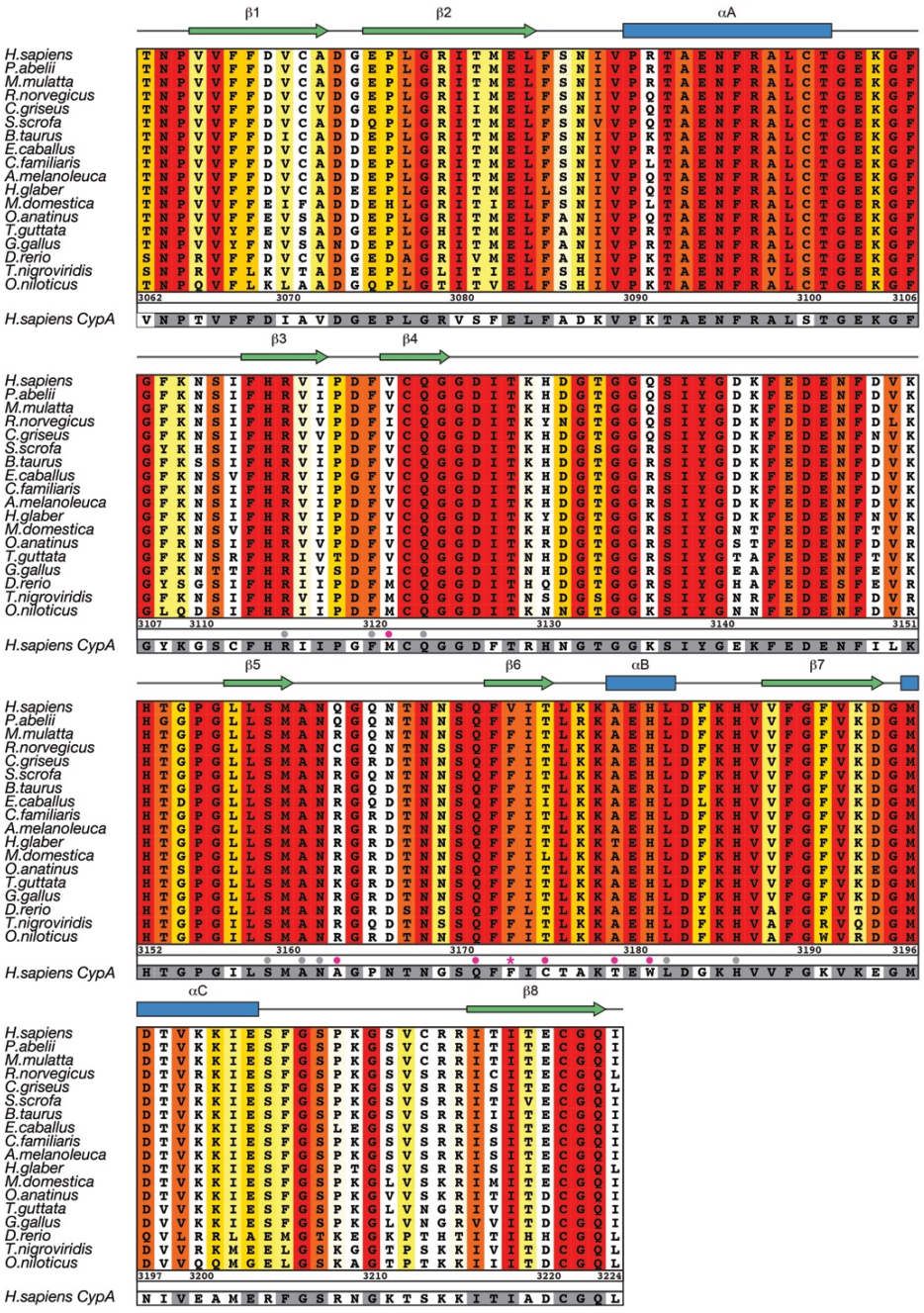


Fig. S2.

Multi-species sequence alignment of Nup358^{CTD}. The overall sequence identity is shaded from white (less than 60 % identity) to red (100 % identity). The secondary structure as observed in the Nup358^{CTD} structure is shown above the alignment, with blue bars representing α helices and green arrows representing β strands. The sequence of human cyclophilin A is displayed at the bottom of the alignment with residues identical to human Nup358^{CTD} shaded in grey. Gray and magenta dots indicate identical and different active site residues between human Nup358^{CTD} and human cyclophilin A, respectively. The numbering below the sequence is relative to human Nup358.

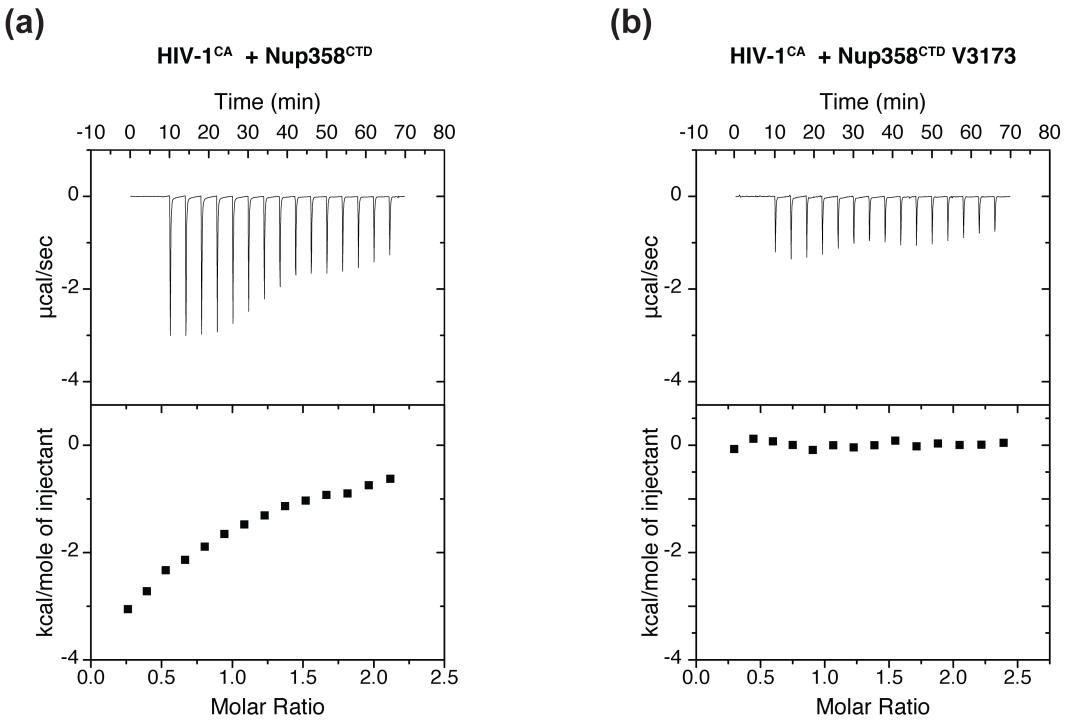


Fig. S3.

Isothermal titration calorimetry (ITC) analysis. Upper parts of each box show raw data and lower parts show integrated heat changes corrected for heat from dilution for interactions between (a) $Nup358^{CTD}$ and HIV-1^{CA} and (b) $Nup358^{CTD}$ V3173W and HIV-1^{CA}.

Table 1.

Crystallographic Analysis

Data collection	
Protein	$Nup358^{CTD}$
Synchrotron	SSRL
Beamline	BL12-2
Space group	$P2_1$
Cell parameters	
a, *b*, *c* (Å)	a=104.1, b=97.1, c=108.2
α, β, γ (°)	α=90.0 β=117.4 γ=90.0
Wavelength (Å)	0.9795
Resolution (Å)	50.0 – 1.75
R_{sym} (%)	10.5 (98.5)
</> / <σ/>	12.8 (2.3)
Completeness (%)	97.7 (92.8)
No. observations	1,742,028
No. unique reflections	188,791 (17,824)
Redundancy	9.2 (7.7)
Refinement	
Resolution (Å)	50.0 – 1.75
No. reflections total	188,714
No. reflections test set	9,500 (5.03%)
R_{work}/ R_{free}	12.2/ 15.8
No. atoms	17,190
Protein	15,408
Ligand	244
Water	1,538
B-factors	
Protein	22.6
Water	38.7
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.2
Ramachandran plot	
Favored (%)	97.6
Additionally allowed (%)	2.4
Outliers (%)	0.0
MolProbity score	1.19

[a]SSRL, Stanford Synchrotron Radiation Lightsource

[b]Highest-resolution shell is shown in parentheses

[c]As determined by MolProbity

CHAPTER 5

STRUCTURAL AND FUNCTIONAL ANALYSIS OF mRNA EXPORT REGULATION BY THE NUCLEAR PORE COMPLEX

This chapter was written by Daniel. H Lin, Ana R. Correia, Sarah W. Cai, and André Hoelz.

SUMMARY

The nuclear pore complex (NPC) controls the passage of macromolecules between the nucleus and cytoplasm, but despite progress towards determining its molecular architecture, the mechanisms by which it directly participates in macromolecular transport remain poorly understood. We performed a structural and functional analysis of how the NPC regulates the final step of mRNA export, in which the DEAD-box helicase DDX19 (Dbp5 in yeast) is activated by the nucleoporins Gle1, Nup214, and Nup42 to remove the mRNA export factor Nxf1•Nxt1 from mRNAs. X-ray crystal structures of Gle1•Nup42 from *S. cerevisiae*, *H. sapiens*, and *C. thermophilum* reveal a conserved mode of binding and a critical role for Nup42 in Gle1 thermostability. Analysis of mutations linked to neurodegenerative disease implicates Gle1 stability as an important determinant in human disease. Reconstitution of ATPase activity establishes that human DDX19 activation is IP_6 independent, in contrast to the fungal system. Crystal structures of human Gle1•Nup42•DDX19 in an early activation intermediate reveal the structural basis of IP_6-independent activation of DDX19 by Gle1. These results provide a foundation for detailed mechanistic analysis of mRNA export in humans.

INTRODUCTION

In eukaryotes, genomic DNA is segregated into the nucleus, serving both to ensure genomic integrity and to spatially separate transcription from translation in the cytoplasm. For the flow of genetic information from the nucleus to the cytoplasm to occur, mRNAs must be exported out of the nucleus through nuclear pore complexes (NPCs). NPCs are massive macromolecular machines perforating the nuclear envelope, each composed of ~1000 protein subunits (called nucleoporins) totaling to a molecular mass of ~120 MDa (Hoelz et al., 2011). By fusing the inner and outer nuclear membranes, NPCs create pores through the nuclear envelope, but simultaneously generate a passive diffusion barrier composed of disordered protein sequences enriched in phenylalanine-glycine repeats (FG repeats). Architecturally, NPCs are composed of a ~ 50-60 MDa symmetric core that is decorated by different proteins on the nuclear and cytoplasmic faces which are referred to as the nuclear basket and cytoplasmic filaments, respectively. Whereas the overall architecture of the symmetric core of the human NPC has recently been elucidated (Kosinski et al., 2016; Lin et al., 2016), the organization of the cytoplasmic filaments and nuclear basket remain poorly understood.

Preparation of mRNAs for nuclear export is a highly-coordinated process that begins co-transcriptionally and proceeds through the addition and removal of mRNA-binding proteins over the course of transcription, splicing, and other nuclear processing events until an export competent messenger ribonucleoprotein (mRNP) is formed (Stewart, 2010). While some mRNAs may be exported through specialized pathways, the bulk of mRNA export is mediated by the conserved, heterodimeric transport factor Nxf1•Nxt1. Nxf1•Nxt1 binds RNA without strong sequence preference and because it can also bind to the FG repeats in the NPC, it is able to shepherd mRNAs through the diffusion barrier. Upon reaching the cytoplasmic face of the NPC, mRNPs containing Nxf1•Nxt1 encounter the cytoplasmic face-specific nucleoporins Gle1, Nup42, and Nup214. These nucleoporins specifically activate

the DEAD-box helicase DDX19, which removes Nxf1•Nxt1 from the mRNP. This activity ensures the directionality of mRNA export, because once Nxf1•Nxt1is removed, the mRNP cannot cross the diffusion barrier to reenter the nucleus.

DDX19 is a member of the DEAD-box helicases, a large family of RNA-dependent ATPases composed of two RecA domains (which we will refer to as the $Ddx19^{NTD}$ and $Ddx19^{CTD}$). Because complete stimulation of human DDX19 activity has not yet been reconstituted, many insights into the regulation of DDX19 have come from the study of the fungal homologs, as DDX19, Nup214, Gle1, and Nup42 are highly conserved. Genetic, biochemical, and structural studies of the yeast proteins have revealed that Gle1 and DDX19 bind to each other via their C-terminal domains (CTDs) in an interaction bridged by the small molecule inositol hexaphosphate (IP_6) (Alcazar-Roman et al., 2006; Dossani et al., 2009; Montpetit et al., 2011; Weirich et al., 2006). Gle1, IP_6, and RNA cooperate to stimulate DDX19 ATPase activity, although the precise mechanism by which stimulation is effected remains controversial (Folkmann et al., 2011; Montpetit et al., 2011). The precise functional roles for Nup42 and Nup214 also remain unclear. The interaction between Nup214 and DDX19 is required for steady-state localization of DDX19 to the NPC, but occurs in a mutually exclusive manner with RNA and inhibits DDX19 activity (Hodge et al., 1999; Napetschnig et al., 2009; Schmitt et al., 1999; von Moeller et al., 2009). Nup42 binds to $Gle1^{CTD}$ (Kendirgi et al., 2005; Saavedra et al., 1997), but its effects on DDX19 activity are unknown.

In addition to their roles in mRNA export, DDX19 and Gle1 have also been implicated in other cellular functions including transcription regulation, DNA damage response, translation initiation, and RNA processing (Aditi et al., 2015; Alcazar-Roman et al., 2010; Bolger and Wente, 2011; Hodroj et al., 2017; Mikhailova et al., 2017; Neumann et al., 2016). Importantly, Gle1 has been linked to several human diseases. Specific point

mutations of Gle1 are associated with lethal contracture congenital syndrome 1 (LCCS1), lethal arthrogryposis with anterior horn cell disease (LAAHD), and amyotrophic laterals sclerosis (ALS) (Kaneb et al., 2015; Nousiainen et al., 2008). Gle1 mislocalization and impairment of nucleocytoplasmic transport in general are also implicated in Huntington's disease and familial ALS linked to the C9orf72 expansion (Freibaum et al., 2015; Gasset-Rosa et al., 2017; Grima et al., 2017; Jovicic et al., 2015; Shang et al., 2017; Shi et al., 2017; Zhang et al., 2015).

To gain further insight into how DDX19 function is regulated by the NPC, we sought to characterize the molecular architecture of the nucleoporins that regulate it. This led us to identify a mechanism of mutually exclusive binding between Nup155, Gle1, and Nup98. Crystal structures of the Gle1•Nup42 complex from *S. cerevisiae*, *C. thermophilum*, and *H. sapiens* revealed the highly conserved structural basis of their interaction. Nup42 has a profound effect on the thermostability of Gle1, which enabled us to analyze the human DDX19 ATPase cycle in the context of Gle1 for the first time. We found that human DDX19 activation is IP_6 independent and analysis of Gle1 and DDX19 sequences suggests that this may be true for all metazoans. Crystal structures of the human Gle1•Nup42•DDX19 complex reveal the adaptations that facilitate IP_6 independent activation in humans and the specific conformational changes that Gle1 induces in DDX19 to enable ATPase activation.

RESULTS

Gle1 and Nup98 recognize the same surface on Nup155^{CTD}

To gain a better understanding of the molecular architecture of the nucleoporins that regulate mRNA export, we set out to reconstitute the interactions *in vitro* with purified, recombinant proteins. In humans Gle1 exists as two isoforms, Gle1A and Gle1B, which are primarily localized to the cytoplasm or nuclear rim, respectively (Kendirgi et al., 2003). As we were interested in the role of Gle1 in the context of the NPC, we focused our analysis on Gle1B and refer to it throughout the text as Gle1. Human Gle1 can be divided into three structural domains: an unstructured N-terminal region (Gle1^{N}, residues 1-123), a coiled-coil region (Gle1^{CC}, residues 124-355), and a highly-conserved C-terminal domain (Gle1^{CTD}, residues 379-698) (Figure 1B). Although Gle1^{CTD} is the domain that binds and stimulates DDX19, previous studies have suggested that features in all three regions are required for NPC localization (Folkmann et al., 2013; Kendirgi et al., 2005; Rayala et al., 2004).

We began our analysis by focusing on the interaction between Gle1^{N} and Nup155 (Nup170 in fungi). Although the first 28 residues of Gle1 were previously shown to be sufficient for an interaction, we used a construct that also included several charged residues at the C-terminus to enhance protein solubility (residues 2-33, Gle1^{N}) (Rayala et al., 2004). We observed formation of a stoichiometric complex between Nup155^{CTD} and Gle1^{N} in size exclusion chromatography experiments (SEC) (Figure 1C). Nup155^{CTD} also contains a binding site for Nup98, which is both a component of the symmetric core of the NPC and of the cytoplasmic filaments (Lin et al., 2016). We could form a stoichiometric complex between Nup155^{CTD} and Nup98$^{\Delta FG}$ (Figure S1A). However, when we attempted to reconstitute a trimeric complex of Nup155, Nup98, and Gle1 by adding Nup98$^{\Delta FG}$ to Nup155^{CTD}•Gle1^{N}, we found that complex formation between Nup155 and Nup98 coincided with displacement of Gle1^{N}, suggesting that the interactions were mutually exclusive (Figure

1C). Similarly, addition of Gle1^{N} to Nup155^{CTD}•Nup98$^{\Delta FG}$ dissociated that complex (Figure S1A).

To determine the molecular basis for the apparent mutual exclusivity of these interactions, we expanded upon mutational and structural analysis we previously performed using the *C. thermophilum* proteins. *C. thermophilum* Nup98 binds to two hydrophobic pockets on the *C. thermophilum* Nup155^{CTD} surface and point mutations in either binding partner are sufficient to disrupt this interaction (Lin et al., 2016). We tested variants of human Nup155^{CTD} containing homologous mutations for their ability to bind to Gle1^{N} or Nup98$^{\Delta FG}$ (Figure 1D, S1B-C). Mutations that abolished binding for the *C. thermophilum* proteins also disrupted binding between human Nup155^{CTD} and Nup98$^{\Delta FG}$, demonstrating that the mechanism of interaction between these two nucleoporins is evolutionarily conserved (Figure 1D, S1B). Moreover, residues that were critical for Nup98$^{\Delta FG}$ binding were also important for Gle1^{N} binding, indicating that Gle1^{N} and Nup98$^{\Delta FG}$ recognized the same surface on Nup155^{CTD} (Figure 1D, S1C). By testing the effect of mutations in Gle1^{N} on complex formation, we also identified several hydrophobic residues critical for Nup155^{CTD} binding, which would be consistent with utilization of the same hydrophobic pockets as Nup98$^{\Delta FG}$ (Figure 1E, S2A).

In summary, we found that Gle1^{N} and Nup98$^{\Delta FG}$ recognized the same surface on Nup155 in a mutually exclusive manner. The sequence of Gle1^{N} that binds to Nup155^{CTD} is present in many metazoans, but is not present in fungi, suggesting that this interaction emerged more recently than the interaction between Nup155^{CTD} and Nup98$^{\Delta FG}$. In the composite structure of the NPC symmetric core, Nup155 molecules bridge the inner ring to the cytoplasmic and nuclear outer rings (Lin et al., 2016). In the cytoplasmic outer ring, but not the nuclear outer ring, there is a volume of unexplained density directly adjacent to the

Nup155^{CTD} surface that binds Gle1^{N} (Figure 1F). Our data suggests this density contains the remainder of the Gle1 molecule and its binding partners.

Identification of a minimal Nup42 fragment sufficient for Gle1 binding

We next focused on the interaction between Gle1^{CTD} and Nup42. To identify the fragment of Nup42 that recognized Gle1, we utilized an observation made in yeast that Nup42 deletion strains exhibit a temperature-sensitive growth phenotype concurrent with mislocalization of Gle1 from the nuclear rim (Rollenhagen et al., 2004). We introduced a series of mCherry-tagged truncation constructs including full-length Nup42 (residues 1-430), Nup42$^{\Delta FG}$ (364-430), Nup42^{CTD} (397-430), Nup42^{minCTD} (405-430), and Nup42^{FG} (1-395) into *nup42*Δ*/gle1-GFP S. cerevisiae* cells and monitored the localization of Nup42 and Gle1. Consistent with previous reports, these strains were all viable and deletion of Nup42 did not affect Gle1 localization at 30 °C (Figure 2A, S3B) (Rollenhagen et al., 2004). However strains lacking an intact Nup42^{CTD} displayed a strong temperature-sensitive growth phenotype at 37 °C and mislocalization of Gle1 after heat shock at 42° C (Figure 2A, S3B). In addition, only truncations of Nup42 containing an intact Nup42^{CTD} displayed nuclear rim staining consistent with localization to the NPC (Figure 2A). From these results, we concluded that residues 397-430 were sufficient for Nup42 binding to Gle1 and that this interaction was required for Nup42 localization to the NPC.

We next recombinantly purified the minimal *S. cerevisiae* Gle1^{CTD}•Nup42^{CTD} complex and observed that the purified complex was less prone to aggregation than *apo* Gle1^{CTD}, which typically requires the addition of IP_6 in purification buffers for stability (Montpetit et al., 2011). This observation led us to test the effect of IP_6 and Nup42 on the stability of Gle1^{CTD} using differential scanning fluorimetry and a pelleting assay. In both experiments, IP_6 potently improved Gle1^{CTD} stability, with saturating amounts of IP_6 shifting

the melting temperature (T_m) from 22 °C to 36 °C (Figure 2B, S4A). Nup42 had an even more dramatic effect, increasing the T_m from 22 °C to 46 °C and the presence of both IP_6 and Nup42 shifted the T_m of $Gle1^{CTD}$ to 53 °C (Figure 2B, S4A). The thermostability of $Gle1^{CTD}$•IP_6 in the absence of Nup42 (T_m 36 °C) is consistent with the temperatures we observed phenotypes in our *nup42*Δ*/gle1-GFP* strain, suggesting that the mislocalization of Gle1 in our *in* vivo experiments could be due to misfolding of Gle1 and subsequent dissociation from the NPC (Figure 2A).

Human $Gle1^{CTD}$ has been difficult to purify recombinantly, but the dramatic effect of yeast Nup42 on Gle1 stability led us to test if human Nup42 exerted the same effect on human Gle1. We found that co-purification of a homologous human $Gle1^{CTD}$•$Nup42^{CTD}$ complex resulted in dramatic improvements in the yield and stability of $Gle1^{CTD}$. In thermostability experiments, the T_m of human $Gle1^{CTD}$ increased from 35 °C to 49 °C in the presence of $Nup42^{CTD}$ (Figure 2C, S4B). However, in contrast to the yeast proteins, IP_6 did not have a detectable effect on thermostability for human Gle1 (Figure 2C). In summary, our analysis identified an evolutionarily conserved C-terminal fragment of Nup42 that binds to Gle1 and has a profound effect on Gle1 stability in both yeast and human.

An evolutionary conserved mechanism of interaction between $Gle1^{CTD}$ and $Nup42^{CTD}$

To further understand the molecular basis for the interaction between $Gle1^{CTD}$ and $Nup42^{CTD}$ and the resulting stabilization of $Gle1^{CTD}$, we determined the x-ray crystal structure of *S. cerevisiae* $Gle1^{CTD}$•$Nup42^{CTD}$ at 1.75 Å resolution (Figure 2D, Table S1). $Nup42^{CTD}$ folded into a compact domain with a hydrophobic core that buried a solvent-exposed hydrophobic surface on $Gle1^{CTD}$ yielding a total interface area of ~835 $Å^2$ (Figure 2D, S7A). Specifically, the Nup42 hydrophobic core contained two proline residues (P420 and P423) that pointed their pyrrolidine rings inwards towards phenylalanine residues F409 and F414 (Figure 2H).

This core wrapped around and buried the exposed hydrophobic Gle1 residues W451, Y488, and L495 (Figure 2H). This hydrophobic interface was supplemented by polar interactions including a salt bridge between Nup42 residue D421 and Gle1 residue R456, as well as a network of hydrogen bonds between Gle1 residues Q491 and K494 with the backbone carbonyls of Nup42 residues I408, F409, A411, and L428 (Figure 2H). To confirm the relevance of the crystal structure, we purified a series of mutants designed to perturb the observed interactions or the hydrophobic core of Nup42. However, in SEC experiments, none of the individual mutations had a detectable effect, consistent with the extensive interaction surface (Figure S3B). Instead, more disruptive mutations that introduced negative charge into the hydrophobic core (F409D, F414D, F409D/F414D) were necessary to disrupt the interaction between Gle1 and Nup42 (Figure S3C). In agreement with these results, we found that mutant Nup42 constructs containing the F409D, F414D, or F409D/F414D mutations were unable to rescue deletion of Nup42 deletion (Figure S3D).

We next wondered if the mode of interaction observed for *S. cerevisiae* $Gle1^{CTD}$•$Nup42^{CTD}$ was conserved in other eukaryotes, but there was no structural information for any homologues of Gle1 or Nup42. To address this question, we determined the x-ray crystal structure of human $Gle1^{CTD}$•$Nup42^{CTD}$ at 2.8 Å resolution and of *C. thermophilum* $Gle1^{CTD}$•$Nup42^{CTD}$ in the presence and absence of IP_6 at 2.2 Å and 2.65 Å resolution, respectively (Figure 2E-2F, S10A, Tables S1, S2). The crystal structures revealed that Nup42 folded into a similar structure and recognized the same surface on Gle1 in both *H. sapiens* and *C. thermophilum*, with the critical hydrophobic residues nearly universally conserved (human Gle1 residues W602, Y637, M644 recognized by human Nup42 residues F401, F406, P412, and P416; *C. thermophilum* Gle1 residues W447, Y484, A491, and A451 recognized by *C. thermophilum* Nup42 residues W530, F539, P544, and P548) (Figure 2I-2J, S5, S8, S9). While human Nup42 adopted a nearly identical fold to *S. cerevisiae* Nup42,

C. thermophilum Nup42 contained an extended N-terminus that increased the area involved in the interaction (Figure 2F, S9A). This insertion appeared to be present at the sequence level among species within the *Pezizomycotina* subdivision of fungi (Figure S6). The overall structure of $Gle1^{CTD}$ was largely conserved, with minor differences resulting from small insertions or different loop sizes (Figure 2G).

Our combined structural data demonstrate that Nup42 folds into a compact domain that buries an exposed hydrophobic surface on Gle1 and that this mechanism of interaction is highly conserved between fungi and humans. Burial of the exposed hydrophobic residues and the thermodynamic favorability of Nup42 folding could explain the large effect Nup42 has on Gle1 stability. The remainder of Nup42 is comprised primarily of FG repeats. In yeast, deletion of the FG repeats was detrimental when combined with the deletion of other FG repeats in the cytoplasmic filaments (Adams et al., 2014). Thus, in addition to ensuring the stability of $Gle1^{CTD}$, $Nup42^{CTD}$ also has a role in anchoring FG repeats proximal to Gle1 in the NPC. To our knowledge, this is the first example where crystal structures of the same nucleoporin-nucleoporin interaction from three divergent species could be compared. We speculate that many of the critical interaction interfaces in the NPC will possess a similar degree of structural conservation.

The IP_6 binding pocket is not conserved in metazoan Gle1

In yeast, activation of Dbp5 ATPase activity requires the small molecule IP_6 (Alcazar-Roman et al., 2006; Weirich et al., 2006), which binds to a highly positively charged pocket in $Gle1^{CTD}$ adjacent to the Dbp5 binding surface and bridges the two proteins (Montpetit et al., 2011). In our structure of $Gle1^{CTD}$•$Nup42^{CTD}$, the Nup42 binding surface is opposite of the IP_6 and Dbp5 binding surfaces (Figure S7) (Montpetit et al., 2011). Thus, for Nup42 to impact activation of Dbp5 or DDX19, it would have to exert an allosteric effect on

Gle1. To identify potential conformational changes mediated by Nup42 binding, we superposed our Gle1^{CTD}•Nup42^{CTD} structure onto the previously determined structure of Gle1^{CTD}•IP_6•Dbp5, but observed minimal conformational differences between the two states (Figure S10B) (Montpetit et al., 2011). However, because there were no *apo* Gle1 structures available, we could not exclude the possibility that Nup42 binding caused Gle1 to adopt the same conformation as IP_6 and Dbp5 binding

To directly assess the conformational changes in Gle1 induced by IP_6 binding independent of Dbp5 binding, we compared the conformations of our *apo* and IP_6-bound structures *C. thermophilum* Gle1^{CTD}•Nup42^{CTD} (Figure S10A). We observed minimal conformational changes upon IP_6 binding, mostly limited to the loop directly adjacent to the IP_6 pocket (Figure S10A). Similar to the structures of *S. cerevisiae* Gle1^{CTD} determined in the presence of IP_6, several IP_6 molecules were present in the crystal structure IP_6-bound *C. thermophilum* Gle1^{CTD}•Nup42^{CTD}, but only the primary site was occupied in both species (Figure S10C) (Montpetit et al., 2011). Despite differences in the location of positively charged residues in the *S. cerevisiae* and *C. thermophilum* IP_6 pockets, the electrostatic potential of the pocket was conserved and IP_6 bound in a similar orientation, suggesting that IP_6 could function in a similar role in *C. thermophilum* as *S. cerevisiae* (Figure 3A-3B, S4D).

In contrast to the IP_6 binding pockets of *S. cerevisiae* and *C. thermophilum* Gle1^{CTD}, the same position on the *H. sapiens* molecule was significantly altered, resulting in a dramatic reduction in positive electrostatic potential (Figure 3A-3F). This was because several positively charged residues that were nearly invariant in fungi (S. cerevisiae residues K264, K286, K333, R374, C. thermophilum residues K225, R249, K327, K374) were not conserved in humans (V401, Q423, E482, H523) (Figure 3A-3C). In fact, this trend was true for all the metazoan sequences we inspected (Figure S11A), indicating that the reduced electrostatic potential in this pocket may be a general feature of metazoan Gle1. Sequence conservation

analysis also indicated that the positively charged residues in Dbp5 that bind IP_6 were also not conserved in metazoans (Figure S11B). However, we noted that two positions in human Gle1 contained lysine residues which were not present in the fungal proteins. In sum, our structural observations raised questions about what roles Nup42 and IP_6 had on DDX19 activation.

DDX19 activation in humans is IP_6 independent and the effect of Nup42 is tied to thermostability

We first tested if complex formation was dependent on IP_6 in SEC experiments. Consistent with previous reports, *S. cerevisiae* Dbp5 and $Gle1^{CTD}$•$Nup42^{CTD}$ formed a complex only in the presence of IP_6 (Figure 3G) (Montpetit et al., 2011). We observed the same pattern of IP_6 dependence for the interaction between *C. thermophilum* $Gle1^{CTD}$•$Nup42^{CTD}$ and Dbp5, consistent with the strong evolutionary conservation of the IP_6 binding residues among fungi (Figure 3H). In contrast, complex formation between human $Gle1^{CTD}$•$Nup42^{CTD}$ and DDX19 was more robust in the absence of IP_6 than in the presence IP_6 (Figure 3I). Importantly, we observed significant complex formation for the human proteins in the absence of IP_6, compared to the fungal proteins where there was no complex formation (Figure 3I).

To directly test the effect of Nup42 on Dbp5 activity, we measured steady state ATPase rates using conditions identical to those previously reported (Montpetit et al., 2012). For both yeast Dbp5 and human DDX19, we observed minimal differences in the stimulated levels of ATPase activity by $Gle1^{CTD}$ compared to $Gle1^{CTD}$•$Nup42^{CTD}$ (Figure 2K-2L). We therefore tested whether stimulation of DDX19 by $Gle1^{CTD}$•$Nup42^{CTD}$ was dependent on IP_6. We first confirmed that we could reproduce previously reported stimulation of Dbp5 by Gle1, IP_6, and RNA (Alcazar-Roman et al., 2006; Weirich et al., 2006) in conditions identical to

those previously reported (Figure 3J) (Montpetit et al., 2012). To summarize, *S. cerevisiae* Gle1^{CTD}•Nup42^{CTD} weakly stimulates Dbp5 in the absence of IP_6, but very strongly in the presence of IP_6 (Figure 3J). In contrast, when we performed ATPase assays with human DDX19 and Gle1^{CTD}•Nup42^{CTD}, we found that stimulation in the presence and absence of IP_6 was indistinguishable regardless of the presence of RNA (Figure 3K). We noted that human DDX19 was less active than *S. cerevisiae* Dbp5 and to ensure accurate measurements of ATPase activity, we performed our assays with 5-fold higher concentrations of human DDX19, Gle1, and IP_6. To address the possibility that the differences in the IP_6 binding pocket in human Gle1 and DDX19 weakened affinity for IP_6 rather than ablated binding, we tested whether higher concentrations of IP_6 would stimulate DDX19. However, we saw no effect up to concentrations of 160 μM (Figure 3L), which would be than the highest reported total concentrations in human tissues and 160-fold higher than used for the yeast proteins (Shears, 2001). Taken together, our results indicated that IP_6 binding was conserved in fungi, but not conserved in humans. Sequence analysis of other metazoan Gle1 and DDX19 sequences suggest that IP_6 binding may not be a feature of animal DDX19 activation in general, as the residues were not conserved across a diverse array of metazoan sequences (Figure S5A). Our results do not exclude the possibility that another small molecule may serve a similar function in humans as IP_6 does in fungi. However, we observed substantial complex formation between human Gle1^{CTD}•Nup42^{CTD} and DDX19, indicating that the human proteins were already sufficient for comparable levels of complex formation.

Structural basis for IP_6 independent DDX19 activation in humans

Given the observation that DDX19 activation was IP_6 independent in humans, we looked for other differences that emerged since the divergence of fungi and humans. Structural studies

have revealed that DDX19 possesses an autoinhibitory helix N-terminal to the $DDX19^{NTD}$ (residues 54-67) that can bind between the $DDX19^{NTD}$ and $DDX19^{CTD}$, preventing RNA binding or formation of a catalytically-competent active site (Collins et al., 2009). While both fungal Dbp5 and metazoan DDX19 sequences contain long disordered N-termini, we could only identify homologous helices in metazoans (Figure SX), further indicating that the mechanism of fungal and human DDX19 activation may be different. Furthermore, while crystal structures exist for most steps of the ATPase cycle (Montpetit et al., 2011; Napetschnig et al., 2009; von Moeller et al., 2009), a detailed analysis of the structural basis for activation has been hindered by the absence of a complete set of structures for the entire ATPase cycle from a single species. To address these questions, we determined x-ray crystal structures of human $Gle1^{CTD}$•$Nup42^{CTD}$ in complex with $DDX19^{\Delta N53}$ in the presence of ADP or AMP-PNP•Mg^{2+} to maximum resolutions of 3.6 and 3.4 Å resolution, respectively (Figure 4A-4B, Table S3). We also determined the crystal structure of *apo* $DDX19^{\Delta N53}$•AMP-PNP•Mg^{2+} to a resolution of 2.2 Å (Figure S6, Table S3). For all three structures, we used a construct of DDX19 that retained the autoinhibitory N-terminal helix (residues 54-479, $DDX19^{\Delta N53}$), but did not contain the flexible N-terminus.

The $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N53}$ complexes exhibited similar conformations in the presence of ADP or AMP-PNP•Mg^{2+}, with the autoinhibitory helix still bound between the $DDX19^{NTD}$ and $DDX19^{CTD}$ (Figure 4A, 4D). The surprising observation that ATP-bound DDX19 could adopt the inhibited conformation was confirmed by the structure of *apo* $DDX19^{\Delta N53}$•AMP-PNP•Mg^{2+}, which formed an identical conformation to *apo* $DDX19^{\Delta N53}$•ADP (Figure S13A). Comparison of the nucleotide-binding residues in both states revealed that the nucleotide-binding pocket readily accommodated the additional phosphate and Mg^{2+} ion, with minor changes in the sidechain conformations of K64 from the

inhibitory helix and E243 (Figure S13B). We conclude from our structural data that the inhibited state does not differ structurally in the presence of ADP or ATP. The tight packing of the inhibitory helix between the $DDX19^{NTD}$ and $DDX19^{CTD}$ suggests that this conformation may be the favored conformation of nucleotide-bound DDX19 absent other binding partners, consistent with a strong autoinhibitory effect.

The $Gle1^{CTD}$-$DDX19^{CTD}$ interface involved approximately 40 Gle1 residues and 35 DDX19 residues and buried almost 1300 $Å^2$ in surface area (Figure S8). Most the solvent exposed residues in helices α2 and α4 from $Gle1^{CTD}$ were buried by the interaction and we identified two major interaction interfaces (Figure 4B-C). In the first, Gle1 residues in helix α4 (H495, E491, and Q487) formed an extensive hydrogen bond network with each other and DDX19 main chain atoms (residues 380-382), as well as a potential salt bridge between Gle1 residue E491 and DDX19 residue K385 (Figure 4B). In the second major interface, the highly acidic C-terminal helix of DDX19 was recognized by lysine residues on the Gle1 surface, with potential salt bridges forming between DDX19 residues D470, D472, E475 and Gle1 residues K416, K419, and K479 (Figure 4C). We also noted several instances where DDX19 residues (G329, T332, A334, L471) packed directly against Gle1 helices, fitting in between the sidechain atoms to form hydrophobic interactions. Single point mutations at these interfaces were sufficient to abrogate Gle1-mediated stimulation of DDX19 (Figure 4E).

We next searched for structural differences between the human and yeast complexes that allowed the human complex to form in the absence of IP_6 (Figure 4F). In the structure of *S. cerevisiae* Gle1•IP_6•Dbp5•ADP, the interface between $Gle1^{CTD}$ and $Dbp5^{CTD}$ buried only 700 $Å^2$ in surface area, compared to 1300 $Å^2$ for the human complex, a difference presumably compensated by the presence of IP_6 (Figure S7-S8). The IP_6-binding lysine residues in *S.*

cerevisiae Dbp5 are not conserved in humans. Rather the C-terminal helix that contains those residues was shorter in the human structure and packed more closely to Gle1^{CTD} positioning D470, D472, E475 to form the salt bridges in the second site (Figure 4F, S11A). Similarly in human Gle1, helix α2 was curved which allowed K416 and K419 to form the salt bridges as well as for other residues to form more contacts with DDX19. A β-tongue insertion unique to human Gle1 between helices α3 and α4 also formed new contacts with DDX19 (Figure 4F). Altogether, the human proteins evolved to utilize a more extensive interface to compensate for the loss of IP_6 binding.

To gain further insight into how Gle1^{CTD}•Nup42^{CTD} binding facilitated DDX19 activation, we compared the structures of Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$ to *apo* DDX19$^{\Delta N53}$. The most apparent difference was a rotation of the DDX19^{NTD} and autoinhibitory helix away from the DDX19^{CTD} domain despite the autoinhibitory helix still being bound between the two domains (Figure 5A). In the structure of *S. cerevisiae* Gle1•IP6•Dbp5•ADP, the Dbp5^{NTD} domain is rotated further resulting in the Dbp5^{NTD} also contacting Gle1 (Figure 4F) (Montpetit et al., 2011). In yeast, a triple mutation (V513D/A516D/I520D) in Gle1 at the Dbp5^{NTD} binding interface abolishes Gle1-mediated stimulation of Dbp5 (Montpetit et al., 2011). A homologous triple mutation in human Gle1^{CTD} (G666D/I669D/Q673D) similarly abolished Gle1-mediated stimulation of DDX19, suggesting that the human proteins also form the fully separated conformation observed for the *S. cerevisiae* complex (Figure S12C). We therefore considered our structure as an early intermediate state in which the DDX19^{NTD} and DDX19^{CTD} are partially separated, but the autoinhibitory helix remained bound that would form immediately after Gle1^{CTD} binding to DDX19, but before complete opening of the enzyme.

We next searched for the conformational changes in DDX19^{CTD} that might explain how Gle1-binding would stimulate ATPase activity by comparing the Gle1-bound and *apo* structures of DDX19. There were several differences that could be attributed to Gle1 binding. The first was ordering of the C-terminal helix, which positioned the acidic residues of the helix to interact with lysines on the Gle1 surface (Figure 4C, 5A). In our structure of *apo* DDX19$^{\Delta N53}$•AMP-PNP the C-terminal helix was already partially ordered, likely because it forms crystal contacts (Figure S11). The second was a conformational rearrangement in the adjacent loop, the "trigger loop", containing residues 328-335, allowing I331 to pack against the C-terminal helix, T332 to pack against Gle1, and moving Q335 out of the way of the neighboring "anchor loop" (Figure 5B). This rearrangement was reinforced by the extensive hydrogen bond network formed with Gle1 (Figure 4B, 5B). Third, there was a large rearrangement of the anchor loop (residues 390-403) that lead to a shift in register (see residue C393) and movement away from the autoinhibitory helix (Figure 5C). This resulted in the loss of several contacts between the autoinhibitory helix and DDX19^{CTD} domain, most notably a salt bridge between D398 and R67 and several hydrophobic interactions (Figure 5C), explaining the partial separation of the DDX19^{NTD} and DDX19^{CTD}. Fourth, the loop containing DEAD-box motif VI (residues 429-435), which was disordered in both *apo* DDX19$^{\Delta N53}$ structures, became ordered. DEAD-box motif VI contains several residues that directly bind to nucleotide in the closed, active conformation (R429, R432, F433), and these residues adopt very similar conformations to those observed in the structures of DDX19•AMP-PNP•RNA (Figure 5D). In summary, Gle1 binding to the DDX19^{CTD} caused a cascade of conformational changes that partially released the autoinhibitory helix and prepared the residues responsible for nucleotide binding to form the closed, active conformation.

Gle1^{CTD}•Nup42^{CTD} relieves autoinhibition from N-terminal regulatory sequences

Our structural data suggested that one of the primary mechanisms for Gle1^{CTD} stimulation of DDX19 activity was through relief from inhibition by the N-terminal autoinhibitory helix. Removal of the first 90 residues of *S. cerevisiae* Dbp5 increased unstimulated Dbp5 activity, but did not increase the fully stimulated activity (Montpetit et al., 2011). A previous study of human DDX19 found that removing inhibitory sequences from the N-terminus also increased DDX19 activity, but could not assess the fully-stimulated activity (Collins et al., 2009). However, the autoinhibitory helix appears to be a feature of only metazoan DDX19 sequences and both studies used truncations that removed sequences that formed conserved secondary structure elements in the DDX19^{NTD} domain, potentially disrupting the fold of the domain or otherwise perturbing activity. To obtain a better understanding of the role of various regulatory elements in the N-terminus of DDX19, we tested the activation of a series of truncation mutants that included DDX19$^{\Delta N53}$, the crystallized construct, which still contained the autoinhibitory helix; DDX19$^{\Delta N67}$, which additionally removed the autoinhibitory helix; and DDX19$^{\Delta N91}$, which further removed mobile residues that were a part of the DDX19^{NTD} and was analogous to the yeast Dbp5 truncation construct. Based on our crystal structures of DDX19, we also designed a mutant variant of full-length DDX19, DDX19$^{S60D/K64D}$, containing two aspartate substitutions in the autoinhibitory helix, S60D/K64D, that we predicted would disfavor formation of the inhibited state due to electrostatic repulsion.

We found that all four variants had higher basal ATPase activity than wild-type DDX19, consistent with inhibitory roles for the removed sequences. Removal of the N-terminal extension yielded ~2.5-fold higher basal ATPase activity, which was similar to the reported ~3-fold higher basal ATPase activity upon removal of the entire Dbp5 N-terminal extension (Montpetit et al., 2011)(Figure 6A). In contrast DDX19$^{\Delta N67}$, DDX19$^{\Delta N91}$, and

DDX19$^{S60D/K64D}$ exhibited ~7.5-fold higher basal activity (Figure 6A). The addition of Gle1^{CTD}•Nup42^{CTD}, which stimulated wild-type DDX19, inhibited the activity of all four variants, resulting in comparable activities to wild-type DDX19 (Figure 6A). We observed dramatic stimulation by RNA in the absence of Gle1^{CTD}•Nup42^{CTD} for the DDX19 variants, yielding rates between ~2-fold to ~10-fold faster than wild-type DDX19 (Figure 6A). For wild-type DDX19, addition of both RNA and Gle1^{CTD}•Nup42^{CTD} results in dramatically more stimulation than either alone, but for the variants that perturbed the autoinhibitory helix, addition of Gle1^{CTD}•Nup42^{CTD} yielded modest further stimulation over RNA-mediated levels of most variants, resulting in fully stimulated activities ~2-fold to ~3-fold faster than wild-type DDX19 (Figure 6A). In contrast, the truncation mutant that still possessed an intact autoinhibitory helix, DDX19$^{\Delta N53}$, exhibited a ~2-fold slower fully stimulated activity than wild-type.

This analysis explains several features of DDX19 activation. First, while human DDX19 has much lower ATPase activity than Dbp5, a large part of this difference can be attributed to the autoinhibitory helix present in human DDX19 but not yeast Dbp5, as disruption of this helix yields activities only ~2-fold slower than fully stimulated Dbp5. Second, there appears to be a distinct, but subtle role for the N-terminal 53 residues of DDX19, possibly similar to the role of the N-terminal residues in yeast Dbp5. Third our results indicate that Gle1-mediated stimulation occurs cooperatively with RNA-mediated stimulation, as Gle1 inhibits hyperactive mutants in the absence of RNA, whereas the weak Gle1-mediated stimulation of the hyperactive mutants in the presence of RNA suggests that relief from autoinhibition is indeed a major mechanism for Gle1-mediated stimulation.

Role of mRNA export factors in the context of Gle1 activation

We next analyzed the effect of the mRNA export factors Nup214 and Nxf1•Nxt1 on DDX19 activation in the context of $Gle1^{CTD}$•$Nup42^{CTD}$. The interaction between $Nup214^{NTD}$ and $DDX19^{NTD}$ is required for steady-state localization of DDX19 at the nuclear rim (Hodge et al., 1999; Napetschnig et al., 2009; Schmitt et al., 1999; Weirich et al., 2004). Previously, $Nup214^{NTD}$ was shown to bind to the $DDX19^{NTD}$ in a mutually exclusive manner with RNA and consequently inhibit DDX19 ATPase activity (von Moeller et al., 2009). However, it remains unclear what the precise role of $Nup214^{NTD}$ is in DDX19 activity as studies in yeast have also shown that it can also enhance ADP release from Dbp5 but that $Gle1^{CTD}$ stimulates Dbp5 the same in the presence and absence of $Nup214^{NTD}$ (Montpetit et al., 2011; Noble et al., 2011).

Addition of $Nup214^{NTD}$ did not affect basal DDX19 ATPase activity or activity in the presence of $Gle1^{CTD}$•$Nup42^{CTD}$ (Figure 6B). Consistent with previous reports and structural data, we observed an inhibition of RNA-mediated stimulation by $Nup214^{NTD}$ (Figure 6B) (Napetschnig et al., 2009; von Moeller et al., 2009). However, when DDX19 was stimulated by both $Gle1^{CTD}$•$Nup42^{CTD}$ and RNA, addition of $Nup214^{NTD}$ further increased ATPase activity (Figure 6B). This effect was concentration dependent and more pronounced when $Gle1^{CTD}$•$Nup42^{CTD}$ was limiting (Figure 6C). To better characterize the effect $Gle1^{CTD}$•$Nup42^{CTD}$ and $Nup214^{NTD}$ have on DDX19 activity, we utilized an electrophoretic mobility shift assay to probe their effects on RNA binding. In the conditions we tested, $Gle1^{CTD}$•$Nup42^{CTD}$ enhanced DDX19 binding to RNA while $Nup214^{NTD}$ inhibited DDX19 binding to RNA (Figure 6D). $Gle1^{CTD}$•$Nup42^{CTD}$ rescued RNA binding in the presence of $Nup214^{NTD}$, but incompletely compared to DDX19 binding alone, suggesting that the stimulatory effects of $Nup214^{NTD}$ may be unrelated to RNA binding (Figure 6D). Previous structural work has shown that $Nup214^{NTD}$ binding to DDX19 is incompatible with

the closed, catalytically-competent conformation of DDX19 (Napetschnig et al., 2009; von Moeller et al., 2009). When we previously crystallized Nup214^{NTD} with full-length DDX19, only the DDX19^{NTD} was ordered, suggesting that Nup214^{NTD} binding may separate the DDX19^{NTD} and DDX19^{CTD}, possibly to facilitate the previously observed acceleration of nucleotide release (Napetschnig et al., 2009; Noble et al., 2011). Our findings indicate that under certain conditions, Nup214^{NTD} can be stimulatory for DDX19 ATPase activity, but further analysis in the context a fully assembled complex is necessary.

The effect of Nxf1•Nxt1, the substrate for DDX19 remodeling, on DDX19 activity has not previously been studied. When we introduced saturating amounts of Nxf1•Nxt1, we found that Nxf1•Nxt1 significantly inhibited RNA-mediated stimulation of DDX19 in the absence of Gle1^{CTD}•Nup42^{CTD} (Figure 6B). We attributed this effect to Nxf1•Nxt1 binding and sequestering RNA, effectively reducing the concentration of available RNA. However, DDX19 stimulation was fully restored in the presence of Gle1^{CTD}•Nup42^{CTD}, likely because Gle1^{CTD} increases the effective affinity of DDX19 for RNA. We anticipate that reconstitution of a fully assembled complex with FG domains may reveal further insight into the interplay of all these components.

Human disease mutations in Gle1

Several human diseases have been linked to Gle1 dysfunction, including amyotrophic lateral sclerosis (ALS), Huntington's disease, and the related disorders lethal congenital contracture syndrome 1 (LCCS1) and lethal arthrogryposis with anterior horn cell disease (LAAHD), which lead to spinal cord motor neuron atrophy and premature death, often prior to birth (Gasset-Rosa et al., 2017; Kaneb et al., 2015; Nousiainen et al., 2008). For ALS and LCCS1/LAAHD, specific mutations have been mapped to Gle1^{CTD}. In ALS, three mutations have been associated with Gle1: a nonsense mutation that results in a severely truncated

protein (70 aa versus 698), a splice site mutation that replaces the last 44 amino acids of $Gle1^{CTD}$ with a novel 88 amino acid sequence, and a missense mutation R697C (Kaneb et al., 2015). LCCS1 and LAAHD are related disorders, in which the more phenotypically severe LCCS1 patients are typically homozygous for a 3 amino acid insertion after residue 144 in the coiled-coil domain of Gle1 (Fin_{Major}) and the less phenotypically severe LAAHD patients are heterozygous for the Fin_{Major} mutation but also possess a mutation in $Gle1^{CTD}$ (V617M, I684T) (Nousiainen et al., 2008). The mutation R569H in the $Gle1^{CTD}$ was identified in a patient diagnosed with the more severe LCCS1.

When we mapped the four point mutations onto the structure of $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N53}$, we found that none of the mutated residues directly participated in binding to Nup42 or DDX19 (Figure S15A, S15D-G). We next purified $Gle1^{CTD}$•$Nup42^{CTD}$ variants containing each of the mutations to evaluate their ability to stimulate DDX19 activity and thermostability. We did not observe a defect in ATPase stimulation or thermostability for the ALS-associated R697C mutation (Figure S15B-S15C), but we speculate that because the residue is solvent exposed (Figure S15G), this mutation may lead to Gle1 aggregation and mislocalization when exposed to the high levels of oxidative stress associated with ALS. We measured modest defects in ATPase stimulation for the R569H and V617M mutants (Figure S15B). Strikingly, we observed altered thermostability for all three LCCS1/LAAHD mutations, with the R569H variant exhibiting the largest difference in T_m, decreasing from 48 °C to 36 °C (Figure S15C). The more severe thermostability defect was consistent with the more severe phenotype associated with the R569H mutation versus the less severe phenotypes for the V617M and I684T mutations. R569 was involved in intramolecular interactions with E561, E489 and the carbonyl oxygen of G552 (Figure S15E), whereas V617 and I684 are both buried in hydrophobic cores (Figure

S15D, S15F). Thus our structural and biochemical analysis provides a molecular basis for how these mutations would destabilize the protein fold. Notably, the effects of these mutations may be even more severe in the absence of Nup42^{CTD}. Altogether, our data reveals a strong link between Gle1 stability and human disease.

DISCUSSION

Through reconstitution of the human nucleoporin machinery that regulates mRNA export termination, we have uncovered several new features of DDX19 activation in humans. Mapping of the Gle1-binding site on Nup155 provided a critical spatial restraint for Gle1 and the entire cytoplasmic filament subcomplex adjacent to the bridging Nup155 molecules. The specific spatial organization of Nup42 and Nup214 as well as their FG domains around Gle1 likely also influence the DDX19 cycle. However, a complete characterization of these effects will require the reconstitution of a fully assembled complex with FG domains.

Our analysis revealed that human DDX19 activation differs in important ways from yeast Dbp5 activation. First, human DDX19 activation by Gle1 does not require IP_6. This was supported by multiple lines of evidence: (1) we observed no difference in ATPase stimulation of DDX19 by Gle1 in the presence and absence of IP_6 (2) Gle1 and DDX19 formed a complex in SEC experiments without IP_6 or any other factor, (3) alterations in the IP_6 binding pocket in human Gle1 dramatically reduced the electrostatic potential of the pocket, (4) IP_6 coordinating residues were not conserved in human DDX19, and (5) a crystal structure determined using crystals of $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N53}$ grown with 500 μM IP_6 did not reveal any IP_6 binding. Second, human DDX19 activity was largely dictated by a metazoan-specific autoinhibitory helix. The autoinhibited state appeared to be a favored resting conformation regardless of nucleotide state, as revealed by a novel crystal structure of *apo* DDX19 bound to AMP-PNP.

Crystal structures of $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N53}$ revealed how Gle1 induced rearrangements in the $DDX19^{CTD}$ to stimulate activity. Rearrangement of the anchor loop removed several contacts from the autoinhibitory helix leading to a partial separation of $DDX19^{NTD}$ and $DDX19^{CTD}$. These movements also resulted in ordering of the loop containing critical ATP-binding residues Motif VI, positioning them to bind ATP in the

closed, enzymatically-competent conformation. Previously determined crystal structures of *S. cerevisiae* Gle1^{CTD}•IP$_6$•Dbp5$^{\Delta 90}$ captured what is likely the next conformational step in DDX19 activation – complete separation of the DDX19^{NTD}and DDX19^{CTD} domains leading to complete eviction of the autoinhibitory helix so that the two domains can then close in a catalytically-competent state (Montpetit et al., 2011).

These results allowed us to propose a model of the DDX19 catalytic cycle that reflects the importance of the autoinhibitory helix (Figure 7). DDX19 predominantly exists in an autoinhibited state, which can form in the presence of either ATP or ADP, although we cannot exclude that the ATP- and ADP-bound states exhibit different dynamics for release of the autoinhibitory helix. Gle1 binding facilitates release from autoinhibition by inducing conformational rearrangements in DDX19^{CTD}, eventually leading to complete separation of the DDX19^{NTD} and DDX19^{CTD}, at which point the DDX19^{NTD} contacts Gle1 in the conformation observed of *S. cerevisiae* Gle1^{CTD}•IP$_6$Dbp5. DDX19 can then either form a closed, catalytically competent state, which RNA-binding would favor, or the autoinhibitory helix could rebind, resulting in reformation of the conformation observed in our Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$ structures. In crystal structures of RNA-bound DDX19, the N-terminal loop adopts a conformation that would be incompatible with Gle1-binding, suggesting that Gle1 and DDX19 may dissociate upon formation of a catalytically-competent closed state. After ATP hydrolysis and subsequent dissociation of ADP, P_i, and RNA, DDX19 would then recycle back to the inhibited state and another cycle of ATP hydrolysis could then occur. The precise step in which Nup214^{NTD} is involved remains unclear, although our results indicate that it can further accelerate this process *in vitro*.

Lastly, the effect of Nup42 on Gle1 stability highlights protein instability as a potent modifier of Gle1 function. The Gle1^{CTD} fold is unstable in the absence of cofactors and the

stabilizing effect of Nup42 on Gle1^{CTD} thermostability facilitated our analysis of the role of human Gle1 in the DDX19 cycle. Importantly, the effects of mutations associated with human disease on thermostability correlate with phenotypic severity. Thus, our results reinforce the strong link between mislocalization or aggregation of proteins involved in RNA metabolism and nucleocytoplasmic transport with human disease.

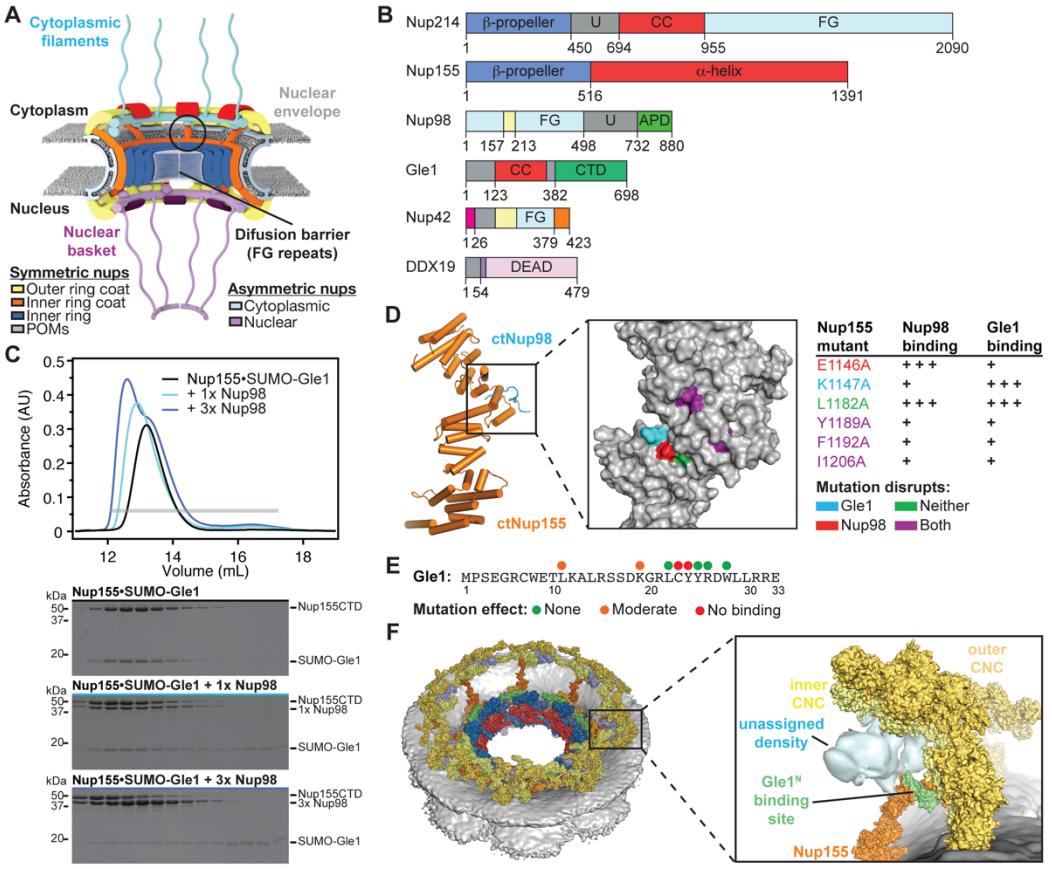


Fig. 1.

Gle1 is anchored to the nuclear pore complex through a competitive interaction with Nup98. (A) Cartoon schematic of the human nuclear pore complex. The region of interest is indicated by a circle on the cytoplasmic side of the nuclear pore complex. (B) Domain schematics for nucleoporins used in this study. Protein names and boundaries correspond to the human proteins. FG, phenylalanine-glycine repeats; CC, coiled-coil domain; U, unstructured; APD, autoproteolytic domain; CTD, C-terminal domain. (C) Size exclusion chromatography analysis of the interaction between Nup155, SUMO-Gle1^{N}, and Nup98$^{\Delta FG}$. Purified Nup155•SUMO-Gle1^{N} complex was mixed with the indicated amounts of Nup98$^{\Delta FG}$ and loaded onto a Superdex 200 10/300 GL size exclusion column. The gray bar indicates the fractions visualized with Coomassie-stained SDS-PAGE gels. (D) Table summarizing size exclusion chromatography analysis of Nup155 mutants for Nup98$^{\Delta FG}$ and SUMO-Gle1^{N} binding. See also Figure S1. The homologous positions were colored on the *C. thermophilum* Nup170•Nup145N structure, indicating that the same binding surface is recognized by Nup98$^{\Delta FG}$ and SUMO-Gle1^{N}. (E) Summary of the effect of alanine mutants in Gle1^{N} on Nup155^{CTD} binding. Colored dots above the sequence of Gle1^{N} indicate the effect of the mutation. See also Figure S2. (F) Identification of the Gle1 binding site suggests that the unassigned cytoplasmic density adjacent to bridging Nup155 molecules would contain Gle1 and its binding partners. left, Cartoon representation of the composite structure of the NPC. right, Zoom view of unassigned cytoplasmic density, with Nup170 shown in orange, the Gle1^{N} binding site colored in green, and the coat nucleoporin complexes shown in yellow.

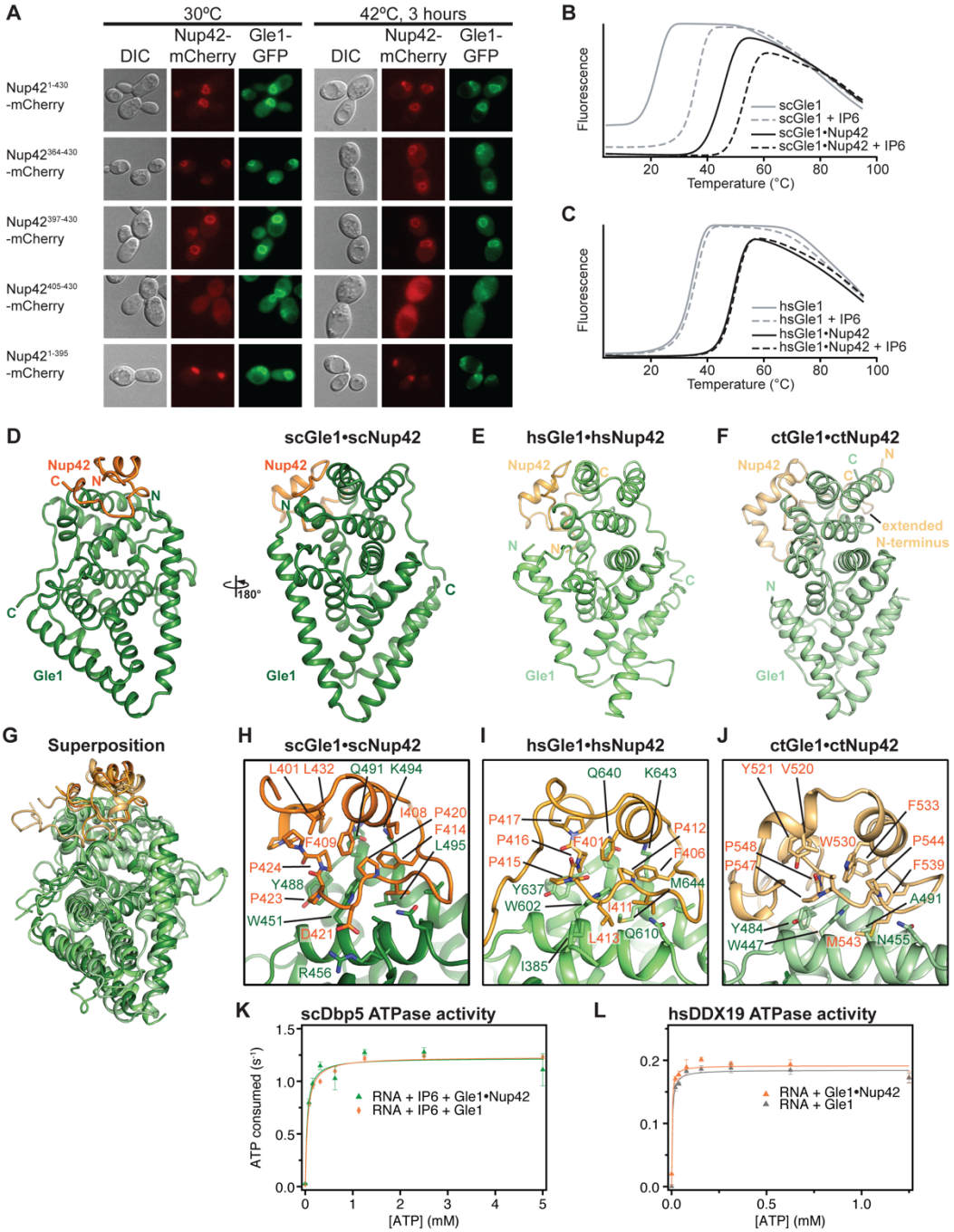


Fig. 2.

A conserved mechanism for Gle1•Nup42 complex formation. (A) *in vivo* localization analysis in *S. cerevisiae* of Gle1-GFP and Nup42-mCherry variants. (B) Differential scanning fluorimetry analysis of *S. cerevisiae* Gle1^{CTD} in the presence and absence of Nup42^{CTD} and IP_6. The curves represent the average of 3 separate experiments. (C) Differential scanning fluorimetry analysis of *H. sapiens* Gle1^{CTD} in the presence and absence of *H. sapiens* Nup42^{CTD} and IP_6. The curves represent the average of 3 separate experiments. (D-F) Crystal structures of (D) *S. cerevisiae,* (E) *H. sapiens*, or (F) *C. thermophilum*

Gle1^{CTD}•Nup42^{CTD}. (G) Superposition of the structures of *S. cerevisiae, H. sapiens*, and *C. thermophilum* Gle1^{CTD}•Nup42^{CTD}, with same coloring as in (D-F). (H-J) Zoom view of (H) *S. cerevisiae,* (I) *H. sapiens*, or (J) *C. thermophilum* Gle1^{CTD}•Nup42^{CTD} interactions with relevant residues labeled. (K) Steady state colorimetric ATPase assay with *S. cerevisiae* Dbp5 performed at 30 °C with either purified *S. cerevisiae* Gle1^{CTD} or Gle1^{CTD}•Nup42^{CTD}. (L) Steady state colorimetric ATPase assay with *H. sapiens* DDX19 performed at 37 °C with either purified *H. sapiens* Gle1^{CTD} or Gle1^{CTD}•Nup42^{CTD}.

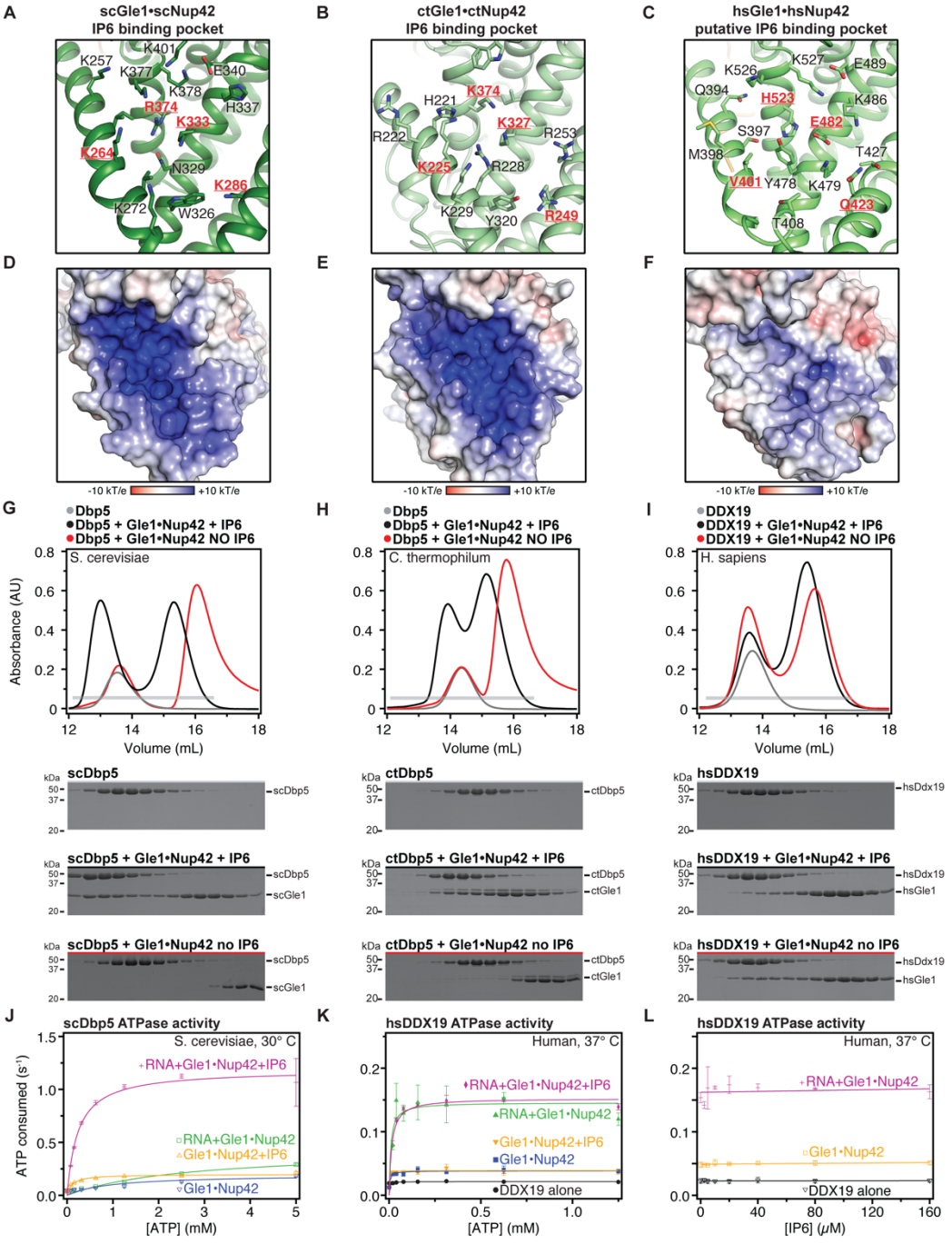


Fig. 3.

Human DDX19 activation is IP6 independent. (A-C) Zoom view of the IP_6 binding pocket of (A) *S. cerevisiae*, (B) *C. thermophilum*, and (C) *H. sapiens* Gle1. Residues that are conserved in fungi but not metazoans are highlighted in bold, red and underlined. (D-F) Surface electrostatic potential analysis of IP_6 binding pockets for (D) *S. cerevisiae*, (E) *C. thermophilum*, and (F) *H. sapiens*. The same view as (A-C) is shown in surface representation, colored by electrostatic potential, from red (-10 k_BT/e) to white (0 k_BT/e) to blue (+10 k_BT/e). (G-I) Size exclusion chromatography analysis of the interaction between

Dbp5/DDX19 and Gle1^{CTD}•Nup42^{CTD} for (G) *S. cerevisiae*, (H) *C. thermophilum*, and (I) *H. sapiens*, in the presence of absence of IP_6. The elution profiles for Dbp5/DDX19 alone are shown in grey, Dbp5/DDX19 with Gle1^{CTD}•Nup42^{CTD} and IP_6 are shown in black, and Dbp5/DDX19 with Gle1^{CTD}•Nup42^{CTD} without IP_6 are shown in red. The gray horizontal bar indicates fractions visualized with Coomassie-stained SDS-PAGE gels shown below. (J) Steady state colorimetric ATPase assay performed with increasing concentrations of ATP at 30 °C for *S. cerevisiae* Dbp5 with the indicated factors added. (K) Steady state colorimetric ATPase assay performed with increasing concentrations of ATP at 37 °C for *H. sapiens* DDX19 with the indicated factors added. (L) Steady state ATPase assay performed with increasing concentrations of IP_6 at 37 °C for *H. sapiens* DDX19.

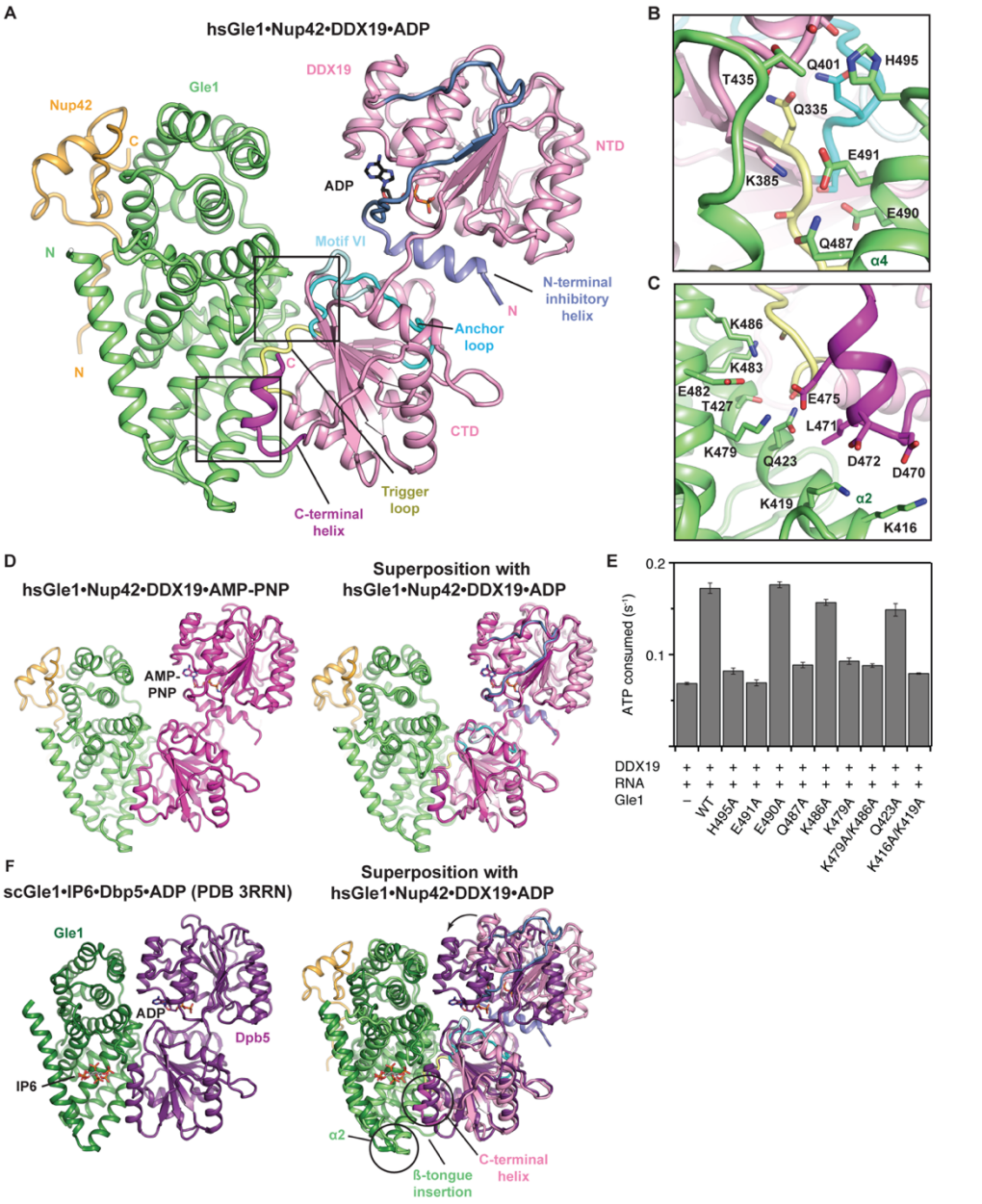


Fig. 4.

Structure of the human Gle1•Nup42•DDX19 complex. (A) Crystal structure of *H. sapiens* Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$•ADP**.** Motifs of interest are colored and labeled. Boxes indicate the regions shown in a zoom view. (B and C) Zoomed views of critical interactions for complex formation (D) left, Crystal structure of *H. sapiens* Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$•AMP-PNP•Mg^{2+}. DDX19 is colored magenta for clarity. right, Superposition of the ADP and AMP-PNP•Mg^{2+} bound structures. (E) Analysis of the effect of point mutations on Gle1-mediated stimulation of DDX19. Values reported are the average of at least three experiments. Error bars indicate standard deviation. (F) left, Crystal structure of *S. cerevisiae* Gle1^{CTD}•IP6•Dbp5$^{\Delta N}$•ADP (PDB code 3RRN). right, Superposition of the *S. cerevisiae* and *H. sapiens* structures. The arrow indicates the rotation relating the conformations observed in the two crystal structures.

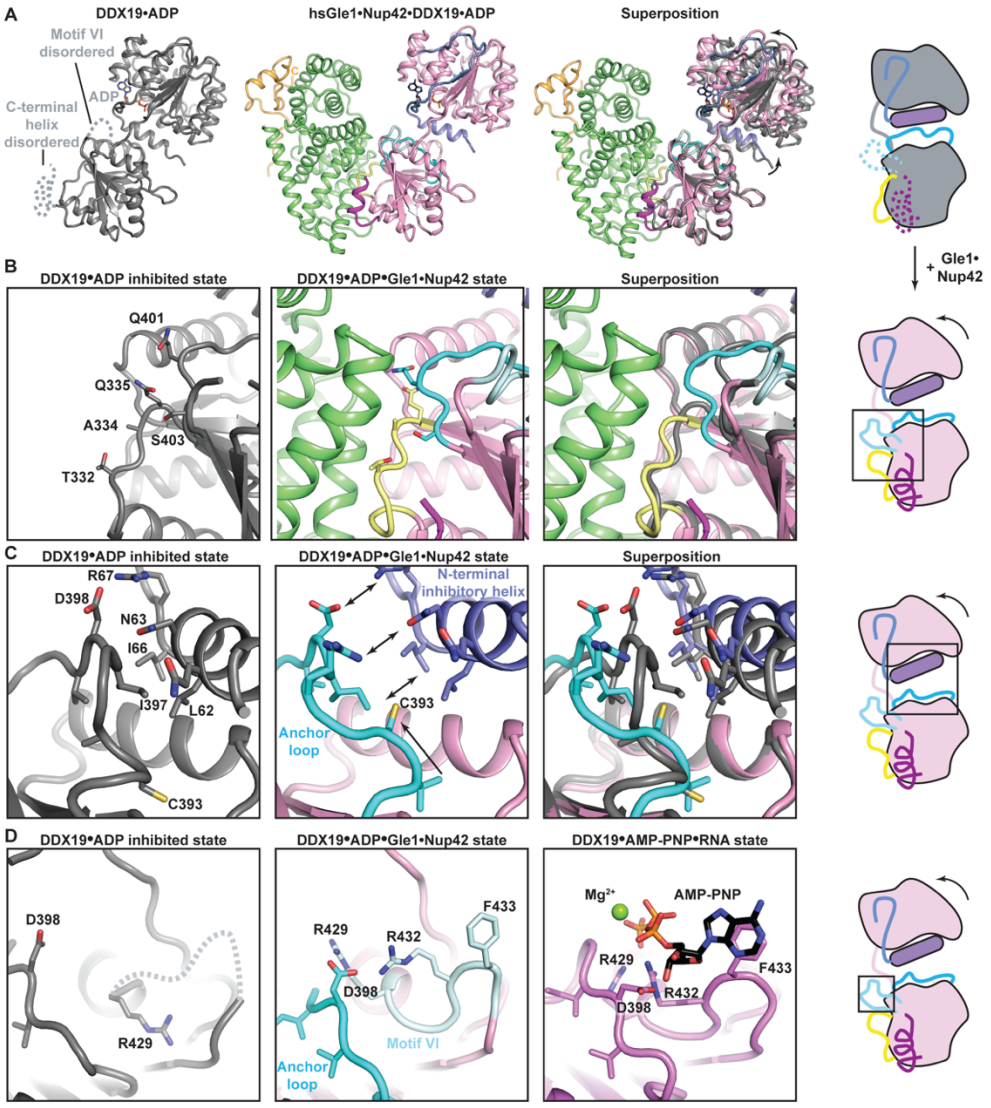


Fig. 5.

Conformational changes in DDX19 induced by Gle1 binding. (A) left, Crystal structure of *H. sapiens* DDX19$^{\Delta N53}$•ADP (PDB code 3EWS). Disordered regions (C-terminal helix and motif VI) are indicated with dashed lines. middle, Crystal structure of *H. sapiens* Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$•ADP**.** middle, shown in the same orientation as Figure 4A. right, Superposition of the two structures. Arrows indicate the rotation relating the conformations of DDX19$^{\Delta N53}$•ADP in the presence and absence of Gle1^{CTD}•Nup42^{CTD}. Cartoon on the right schematizes the transition from the inhibited state to the Gle1-bound state (B) Zoom view of the trigger loop in DDX19 in (left) the inhibited state (PDB code 3EWS), (middle) the Gle1-bound state, and (right) their superposition. The cartoon on the right indicates the part of the protein shown. (C) Zoom view of the anchor loop and autoinhibitory helix in DDX19 in the (left) inhibited state (PDB code 3EWS), (middle) the Gle1-bound state, and (right) a superposition of the two. The cartoon on the right indicates the part of the protein shown. (D) Zoom view of Motif VI in DDX19 in (left) the inhibited state (PDB code 3EWS), (middle) the Gle1-bound state, and (right) the RNA-bound state (PDB code 3G0H). The cartoon on the right indicates the part of the protein shown.

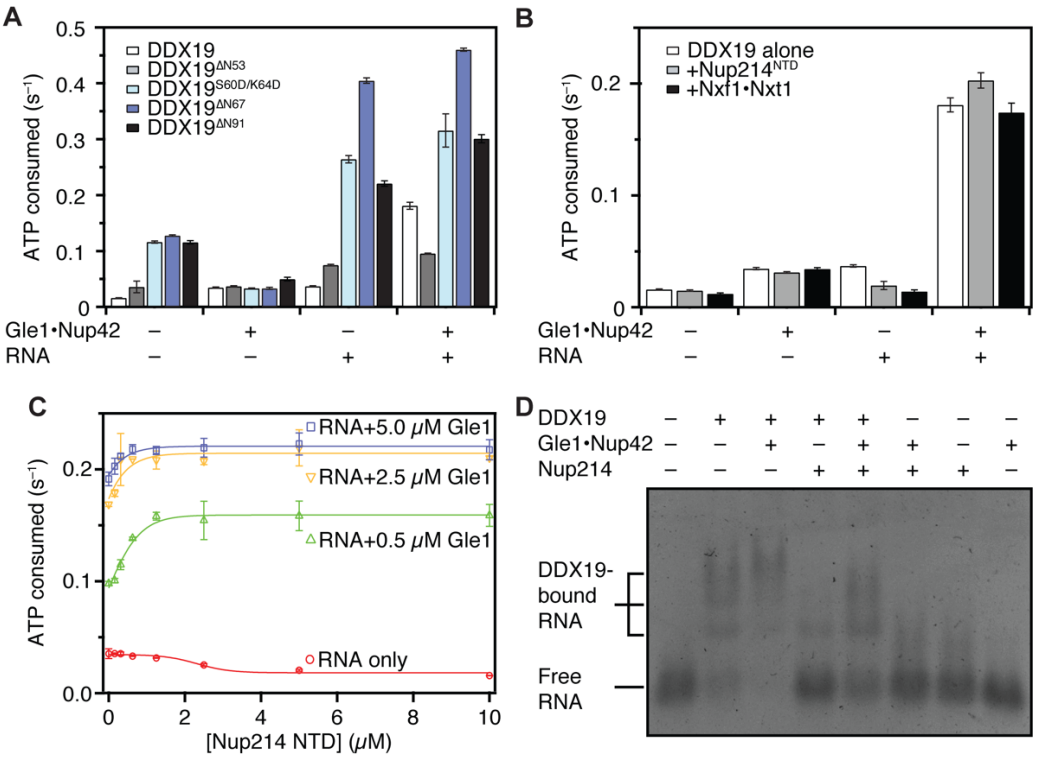


Fig. 6.

Biochemical analysis of DDX19 activity. (A) Analysis of DDX19 variant ATPase steady-state stimulation by RNA and Gle1^{CTD}•Nup42^{CTD}. Values reported are the average of at least three experiments. Error bars indicate standard deviation. (B) Analysis of the effect of Nup214^{NTD} and Nxf1$^{\Delta N}$ •Nxt1 on DDX19 stimulation. (C) Analysis of the dose-dependence of the effect of Nup214^{NTD} on DDX19 stimulation in the presence of RNA and Gle1^{CTD}•Nup42^{CTD}. (D) Electrophoretic mobility shift assay analysis of the effect of Gle1^{CTD}•Nup42^{CTD} and Nup214^{NTD} on DDX19 binding to RNA.

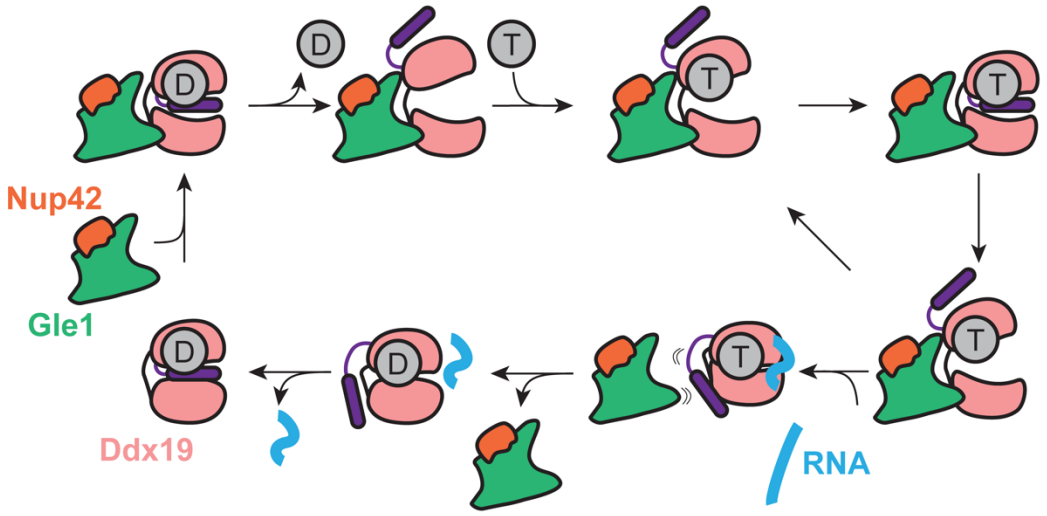


Fig. 7.
Model for IP6-independent activation of DDX19. Gle1 is shown in green, Nup42 in orange, DDX19 in pink with the autoinhibitory helix in purple, RNA in cyan, and nucleotide as a grey circle.

METHODS

Protein expression and purification. Proteins were expressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) in Luria-Bertani media and induced at an OD_{600} of ~0.6 with 0.5 mM IPTG. Unless otherwise noted, cells were harvested by centrifugation and resuspended in a buffer containing 20 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), and 15 mM imidazole, supplemented with complete EDTA-free protease inhibitor cocktail (Roche) and flash frozen in liquid nitrogen. Cells were supplemented with 1 mg deoxyribonuclease I (Roche), lysed with a cell disruptor (Avestin), and centrifuged at 4°C and 30,000g for 1 hour. Supernatants were loaded onto a Ni-NTA affinity column (GE Healthcare) equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), and 15 mM imidazole and eluted with a linear gradient of imidazole concentration to 500 mM. Eluted proteins were dialyzed overnight with a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), and 15 mM imidazole and subsequently purified through affinity, ion exchange, and size exclusion chromatography.

hsGle1^{CTD}•His_6-hsNup42^{CTD} variants, scGle1^{CTD}•His_6-scNup42^{CTD}, and His_6-scGle1^{CTD} were grown for 18 hours at 18°C. His6-ctGle1CTD was co-expressed with GST-ctNup42^{CTD} and grown for 18 hours at 18°C. After elution from the Ni-NTA column, the hexahistidine tags were removed by cleavage with PreScission protease concurrent with dialysis. Dialyzed protein was loaded onto a Ni-NTA column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), and 15 mM imidazole. Protein-containing fractions were pooled and loaded onto a HiTrap Heparin HP column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT and eluted with a linear gradient of NaCl concentration to 2 M. Protein containing fractions were concentrated and loaded onto a HiLoad Superdex

75 16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT.

His_6-SUMO-hsGle1^{CTD} was grown for 18 h at 18°C and purified by similarly to hsGle1^{CTD}•His_6-hsNup42^{CTD}, but the His_6-SUMO was removed using Ulp1 cleavage.

His_6-hsDDX19 variants, His_6-scDbp5 and His_6-SUMO-ctDbp5 were grown for 18 hours at 18°C. Cells were harvested by centrifugation and resuspended in a buffer containing 20 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), 15 mM imidazole, and 5 % (v/v) glycerol, supplemented with complete EDTA-free protease inhibitor cocktail (Roche) and flash frozen in liquid nitrogen. After lysis and centrifugation, the supernatant was loaded onto a Ni-NTA column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), 15 mM imidazole, and 5 % (v/v) glycerol. Protein was eluted with a linear gradient of imidazole concentration to 500 mM and protein-containing fractions were dialyzed overnight in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), 15 mM imidazole, and 5 % (v/v) glycerol. The hexahistidine tag was cleaved using PreScission protease concurrent with dialysis. The hexahistidine-SUMO tag was cleaved with ULP1 protease and was immediately desalted into the dialysis buffer after elution from the Ni-NTA column. The dialyzed/desalted protein was loaded onto a Ni-NTA column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), 15 mM imidazole, and 5 % (v/v) glycerol. Protein-containing flowthrough fractions were loaded onto a HiTrap Q HP column in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT, and 5 % (v/v) glycerol and eluted with a linear gradient of NaCl concentration to 2 M. Protein-containing fractions were concentrated and loaded onto a HiLoad Superdex 200 16/60 PG column equilibrated in a

buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT, and 5 % (v/v) glycerol.

His_6-SUMO-hsGle1^{N} and variants were grown at 30 °C for 2 hours. After elution from the Ni-NTA column, proteins were dialyzed overnight in 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT. After dialysis, proteins were loaded onto a HiTrap Q HP column and eluted with a linear gradient of NaCl concentration up to 2 M. Protein containing fractions were concentrated and loaded onto a Superdex 200 10/300 GL column equilibrated in 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT.

His_6-SUMO-scNup42^{CTD}-His_6 and variants were grown at 37 °C for 2 hours. After elution from the Ni-NTA column, the His_6-SUMO tag was removed by cleavage with Ulp1 concurrent with dialysis into a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT. After dialysis, proteins were run over a HiTrap Q HP column and collected in flowthrough. The flowthrough fractions were concentrated and loaded onto a Superdex Peptide 10/300 GL column.

His_6-Nup155^{CTD} and His_6-Nup214^{NTD} was grown for 18 hours at 18°C. After elution from the Ni-NTA column, proteins were dialyzed overnight in 20 mM TRIS (pH 8.0), 500 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), and 15 mM imidazole. Hexahistidine tags were removed by cleavage with PreScission protease concurrent with dialysis. Dialyzed protein was loaded onto a Ni-NTA column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 4 mM 2-mercaptoethanol (β-ME), and 15 mM imidazole. Protein-containing fractions were pooled and loaded onto a HiTrap Q HP column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT and eluted with a linear gradient of NaCl concentration to 2 M. Protein containing fractions were concentrated and loaded onto a HiLoad Superdex 200

16/60 PG column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, and 5 mM DTT.

His_6-SUMO-Nup98$^{\Delta FG}$ was purified similarly to Nup155^{CTD} and Nup214^{NTD} except the His_6-SUMO tag was removed by cleavage with Ulp1.

The expression plasmid for Nxf1•Nxt1$^{\Delta N}$ was a gift from Murray Stewart, and was purified as previously described(Aibara et al., 2015).

Crystallization and structure determination

S. cerevisiae Gle1^{CTD}•Nup42^{CTD} was crystallized at 23 °C with the hanging drop method using 1 µl of protein solution and 1 µl of reservoir solution (0.1 M HEPES pH 8.2, 11 % (w/v) PEG 3350, and 0.2 M L-Proline). Crystals were cryoprotected with a solution identical to the reservoir solution, but supplemented with 30 % (v/v) ethylene glycol. X-ray diffraction data were collected at the Stanford Synchotron Radiation Lightsource (SSRL), BL 12-2 and processed with the XDS package (Kabsch, 2010). The structure was solved by molecular replacement with Phaser, using the structure of Gle1^{CTD} (PDB code 3RRN) as a search model(McCoy et al., 2007; Montpetit et al., 2011).

H. sapiens Gle1^{CTD}•Nup42^{CTD} was crystallized using a reservoir solution of 0.2 M sodium potassium phosphate pH 7.6 and 26 % (w/v) PEG 3350. Crystals were cryoprotected by gradual supplementation of ethylene glycol in 5 % steps to a final concentration of 30 % (v/v). X-ray diffraction data were collected at the Advanced Photon Source (APS), 23-ID-D and processed with the XDS package. The structure was solved by molecular replacement with Phaser, using the structure of S. cerevisiae Gle1^{CTD} as a search model.

C. thermophilum Gle1^{CTD}•Nup42^{CTD} was crystallized at 23°C with the hanging drop method using 1µl of protein solution and 1µl of reservoir solution (0.1M MES pH 6.3 and

12% (w/v) PEG 20,000). Crystals were cryoprotected by gradual supplementation of ethylene glycol in 5 % steps to a final concentration of 25 % (v/v). X-ray diffraction data were collected at the Advanced Photon Source (APS), 23-ID-D and processed with the XDS package. The structure was solved by molecular replacement with Phaser, using the structure of *H. sapiens* $Gle1^{CTD}$ as a search model.

C. thermophilum $Gle1^{CTD}$•$Nup42^{CTD}$•IP_6 was crystallized at 23°C with the hanging drop method using 1µl of protein solution and 1µl of reservoir solution (0.01M zinc sulphate heptahydrate, 0.1M MES pH 6.3 and 18% (w/v) Polyethylene glycol monomethyl ether 550). Crystals were improved by microseeding. Crystals were cryoprotected by gradual supplementation of ethylene glycol in 5 % steps to a final concentration of 25 % (v/v). SeMet labeled crystals were grown using the same conditions. X-ray diffraction data were collected at the Stanford Synchotron Radiation Lightsource (SSRL), BL 12-2 and processed with the XDS package. The structure was solved using Crank2, using the structure of *H. sapiens* $Gle1^{CTD}$ as an initial search model(Skubak and Pannu, 2013).

For crystallization of the *H. sapiens* $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N1}$, the complex was reconstituted by mixing equimolar amounts of purified $Gle1^{CTD}$•$Nup42^{CTD}$ and $DDX19^{\Delta N1}$ to form a stock solution of 400 µM. The stock solution was supplemented with equimolar ADP and Mg^{2+} or AMP-PNP and Mg^{2+}. Protein stock solutions were diluted with a buffer containing 20 mM TRIS (pH 8.0), 100 mM sodium chloride, 5 mM DTT, and 1mM ADP and Mg^{2+} or AMP-PNP and Mg^{2+}. Crystals of $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N1}$•ADP were grown in 13 % (w/v) PEG 3350 and 0.2 M sodium potassium phosphate. Crystals of $Gle1^{CTD}$•$Nup42^{CTD}$•$DDX19^{\Delta N1}$• AMP-PNP were grown in 15 % (w/v) PEG 3350 and 0.3 M sodium potassium phosphate pH 6.6. Both crystals were cryoprotected by gradual supplementation of ethylene glycol to a final concentration of 25 % (v/v).

Diffraction data for Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N1}$•ADP were collected at APS 23ID-D and processed with XDS. Crystals diffracted anisotropically, with diffraction limits along the three principal components of 3.6, 3.8, and 4.6 Å. The crystal structure was solved by molecular replacement with Phaser, using the structure of *H. sapiens* Gle1^{CTD}•Nup42^{CTD} and *H. sapiens* DDX19 (PDB code 3EWS) as search models(Collins et al., 2009).

Diffraction data for Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N1}$•AMP-PNP were collected at the National Synchotron Lightsource (NSLSII), FMX. The crystals grew in space group $P2_1$ but were non-merohedrally twinned with the twin domains related by a 180° rotation along *a*. Diffraction data were processed using DIALS, and the structure was solved by molecular with Phaser, using the structure of *H. sapiens* Gle1^{CTD}•Nup42^{CTD} and *H. sapiens* DDX19 (PDB code 3EWS) as search models. The diffraction data are anisotropic, with diffraction limits along the three principal components of 3.4, 3.8, and 4.5 Å. Refinement was performed in Phenix using anisotropically truncated and scaled data generated with the UCLA-DOE anisotropy server(Strong et al., 2006).

Crystals of DDX19$^{\Delta N1}$•AMP-PNP grew during crystallization trials of Gle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N1}$•AMP-PNP in heavy precipitate. The reservoir solution contained 0.1 M MIB buffer pH 5.0 (malonate, imidazole, borate) and 13 % (w/v) PEG 1500. Crystals were cryoprotected with the reservoir solution supplemented with 20 % (v/v) ethylene glycol. X-ray diffraction data was collected at APS 23-ID-D and the data was processed using XDS. The crystal structure was solved by molecular replacement using the structure of DDX19$^{\Delta N1}$•ADP (PDB code 3EWS)(Collins et al., 2009).

Analytical size-exclusion chromatography

Protein-protein interaction experiments were carried out on a Superdex 200 10/300 GL gel filtration column equilibrated in a buffer containing 20 mM TRIS (pH 8.0), 100mM NaCl and 5mM DTT (when present, 0.5 mM IP_6 was added to the buffer). The different combinations were mixed and incubated for 30min on ice using a 2-fold molar excess of the smaller component. Complex formation was evaluated by comparing the mobility on the gel filtration column of pre-incubated proteins versus individual proteins. Complex formation was confirmed by SDS-PAGE of the protein containing fractions, followed by Coomassie brilliant blue staining.

Yeast strain generation

The *nup42Δ/gle1-GFP* strain was generated in a BY4741 parental strain by first introducing the *natNT2* cassette by homologous recombination into the nup42 gene followed by three rounds of selection on yeast extract peptone dextrose (YPD) plates containing Nourseothricin (GoldBio). Subsequently, a GFP-*kanMX* cassette was inserted into the C-terminus of Gle1 followed by three rounds of selection on YPD plates containing G418. Nup42-mCherry variants were introduced using a modified pRS411 plasmid followed by two rounds of selection on plates containing leucine-depleted synthetic dextrose complete medium (SDC-Leu).

Yeast live cell fluorescence

Cells were grown in SDC-Leu medium to mid-log phase at 30 ºC and shifted to 42 ºC for three hours. For fluorescence imaging, cells were washed once with water, resuspended, and imaged using a Carl Zeiss Observer Z.1 equipped with a Hamamatsu camera C10600 Orca-R2.

Yeast growth assay

For growth analysis, cells were grown in SDC-Leu medium at 30 ºC to an OD of 0.2. 15 µl of a ten-fold dilution series was spotted onto plates containing SDC-Leu medium, which were incubated at 30 ºC and 37 ºC.

Differential scanning fluorimetry assay

Differential scanning fluorimetry was performed using a previously described protocol(Niesen et al., 2007). Using a real-time PCR instrument (Bio-Rad C1000 96 well Thermal Cycler), fluorescence of a mixture of 10 µM purified protein and 5x SYPRO orange dye (Invitrogen) was measured once per minute while the temperature was increased 1°C/min from 4 °C to 95 °C. Reactions were performed in 20 mM TRIS (pH 8.0), 100mM NaCl, and 5mM DTT. When present, IP_6 was supplemented to 20 µM for yeast proteins and 100 µM for human proteins.

Pelleting thermostability assay

50 µl samples of purified protein (10 µg each) were incubated for 30 minutes at 35, 40, 45, 50, or 55 °C. Soluble and pellet fractions were isolated by centrifugation at 30,000 x g for 35 minutes at 4°C. Reactions were performed in 20 mM TRIS (pH 8.0), 100mM NaCl, and 5mM DTT. When present, IP_6 was included at 20 µM. Protein bands were resolved by SDS-PAGE and visualized with Coomassie brilliant blue staining.

NADH-coupled ATPase assay

Steady-state ATPase activity rates were determined at 30°C for scDbp5 and 37°C for hsDDX19 using previously established conditions(Montpetit et al., 2012). The reaction

mixture (80 µl) contained purified scDbp5 and hsDDX19 (WT and mutants) at 0.5 µM and 2.5 µM, respectively. Unless otherwise noted, all other *S. cerevisiae* and *H. sapiens* proteins (Gle1, Gle1•Nup42, $Nup214^{NTD}$, or Nxf1•Nxt1) were present at concentrations of 1 µM and 5 µM, respectively. When present, polyA RNA (GE Healthcare) was added to a final concentration of 0.1 mg/ml. When present, IP_6 was supplemented to a final concentration of 1 µM and 5 µM for yeast and human, respectively, unless otherwise noted. The reaction mixture contained 30 mM HEPES pH 7.5 (Sigma), 100 mM NaCl (Sigma), 2 mM $MgCl_2$ (Sigma), 1 mM DTT (Gold Biotechnology), 6 mM PEP (Alfa Aesar), 1.2 mM NADH (Sigma), and 1.6 µl PK/LDH (Sigma).

Protein components, IP_6, polyA RNA, and buffer (HEPES, NaCl, and $MgCl_2$) were mixed (20 µl total) and incubated on ice. A mixture of buffer, DTT, PEP, NADH, ATP, and PK/LDH (60 µl total) was dispensed into a 96-well plate and the reaction was initiated by addition of the protein mixture. Plates were centrifuged at 4000 x g for 2 min at 4°C prior to being loaded in the pre-warmed plate reader. A_{340} was measured every 30 sec for 30 min using a FlexStation 3 microplate reader. Rates were calculated by fitting the linear portion of the reaction.

Electrophoretic mobility shift assay

The electrophoretic mobility of free or DDX19-bound RNA was evaluated on native 1.4 % (w/v) agarose gels using a 20 nucleotide RNA probe (poly(U)). The RNA was prepared by *in vitro* transcription using T7 polymerase and according to the MEGAscript protocol (Ambion). The transcribed RNA was loaded on a 8 % denaturing polyacrylamide-urea gel (19:1 acryl:bisacryl ratio and 8.3 M Urea). The band corresponding to the correct size RNA was cut, eluted and extracted with phenol:chloroform. The RNA was than precipitated with ethanol and resuspended in 10 mM TRIS (pH 8.5), its concentration was calculated by

measuring absorbance at 260 nm. The RNA-protein reactions were carried out in 20 mM TRIS (pH 8.0), 100 mM NaCl, 2 mM MgCl2, 1mM DTT, 10 % (v/v) glycerol and 1 mM nucleotide. The recombinant proteins used in each assay were incubated in binding buffer containing either ADP or AMP-PNP and then mixed with 25 ng of RNA. The final reaction volume was 10 µl and the final protein concentrations are described in each experiment. After mixing, samples were incubated on ice for 20 minutes before being loaded onto a 1.4% native-agarose gel (0.25 x TBE). Electrophoresis was carried out at a constant voltage of 9V/cm at room temperature in 0.25 x TBE for 40 minutes. The gels were stained with SYBR gold (1/10,000 dilution in 0.25 x TBE) for 3 minutes and imaged.

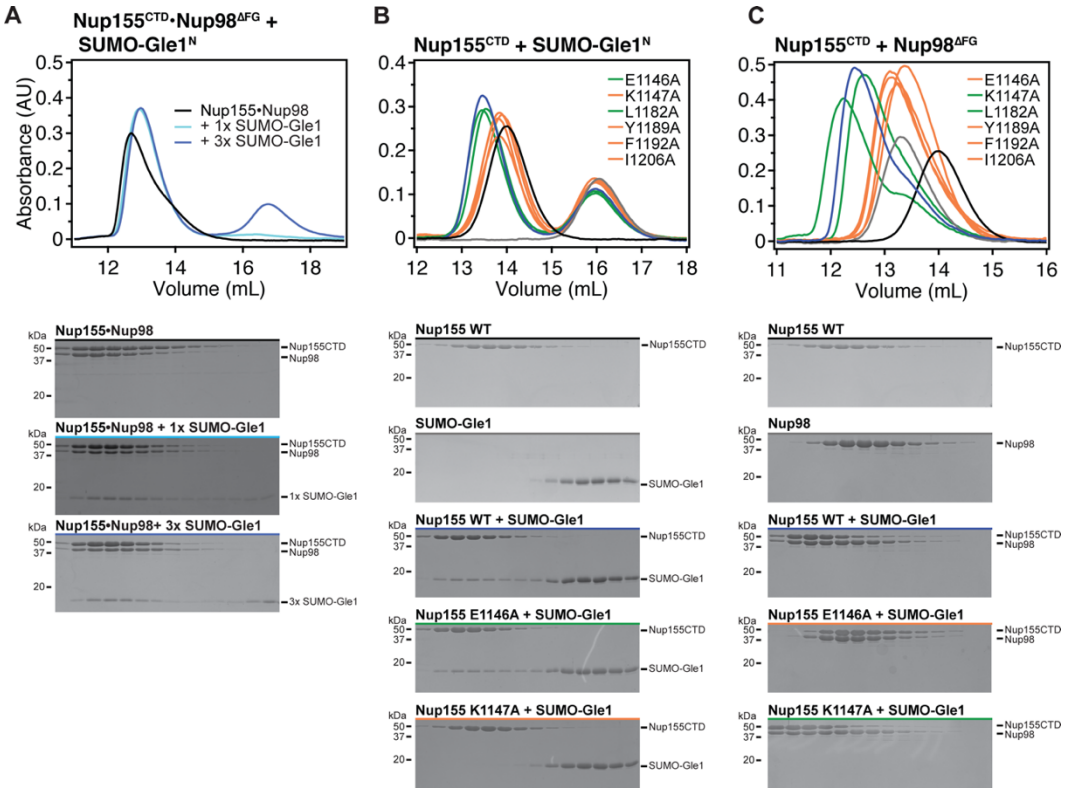


Fig. S1.

The interactions of Nup155 with Nup98$^{\Delta FG}$ or Gle1^{N} are mutually exclusive. Identification of Nup155 residues responsible for Nup98$^{\Delta FG}$ and Gle1^{N} binding. (A) Size exclusion chromatography analysis (SEC) of the interactions between Nup155•Nup98$^{\Delta FG}$ and SUMO-Gle1^{N}. Purified Nup155•Nup98$^{\Delta FG}$ complex was mixed with the indicated amounts of SUMO-Gle1^{N} and loaded on a Superdex 200 10/300 GL size exclusion chromatography column. (B) and (C) Size exclusion chromatography analysis of Nup155 mutants for SUMO-Gle1^{N} and Nup98$^{\Delta FG}$ binding, respectively. Control SEC profiles of Nup155^{CTD} (black), SUMO-Gle1^{N} or Nup98$^{\Delta FG}$ (grey) and preincubated (blue) are shown. SEC profiles of Nup155 alanine mutants preincubated with either SUMO-Gle1^{N} or Nup98$^{\Delta FG}$ are colored according to the measured effect, no effect (green), reduced binding (orange) and complete disruption (red). The gray bar indicates the fractions visualized with Coomassie-stained SDS-PAGE gels.

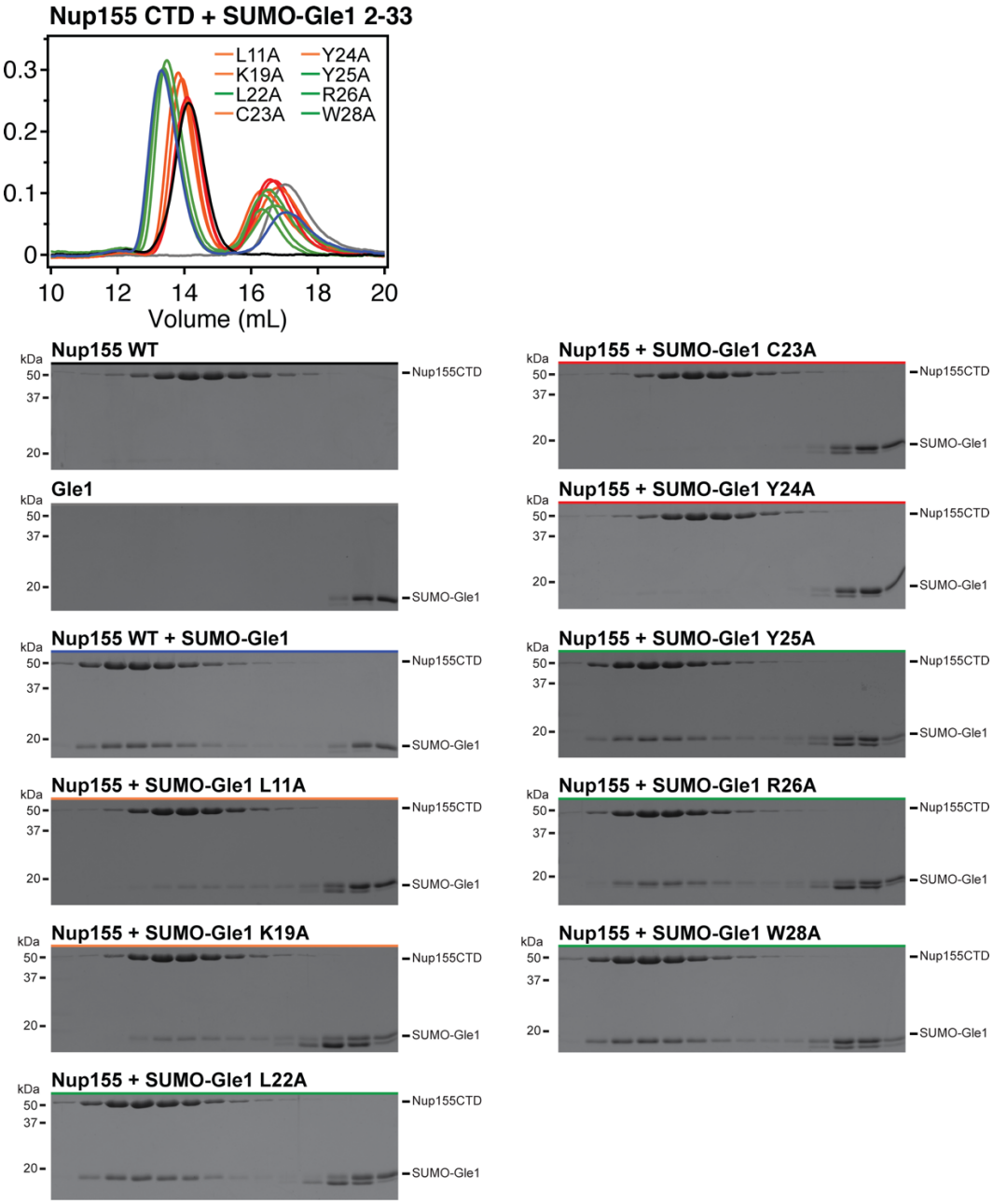


Fig. S2.

Identification of Gle1^{N} residues responsible for Nup155 binding. Size exclusion chromatography analysis of the interactions between SUMO-Gle1^{N} alanine mutants and Nup155^{CTD}. Purified Nup155^{CTD} was mixed with the indicated SUMO-Gle1^{N} mutants and loaded on a Superdex 200 10/300 GL size exclusion chromatography column. SEC profiles of Nup155^{CTD} (black), SUMO-Gle1$^{N(WT)}$ (grey) and Nup155^{CTD} preincubated with SUMO-Gle1$^{N(WT)}$ (blue) are shown as controls. SEC profiles of SUMO-Gle1 alanine mutants preincubated with Nup155^{CTD} are colored according to the measured effect, no effect (green), reduced binding (orange) and complete disruption (red). The gray bar indicates the fractions visualized with Coomassie-stained SDS-PAGE gels.

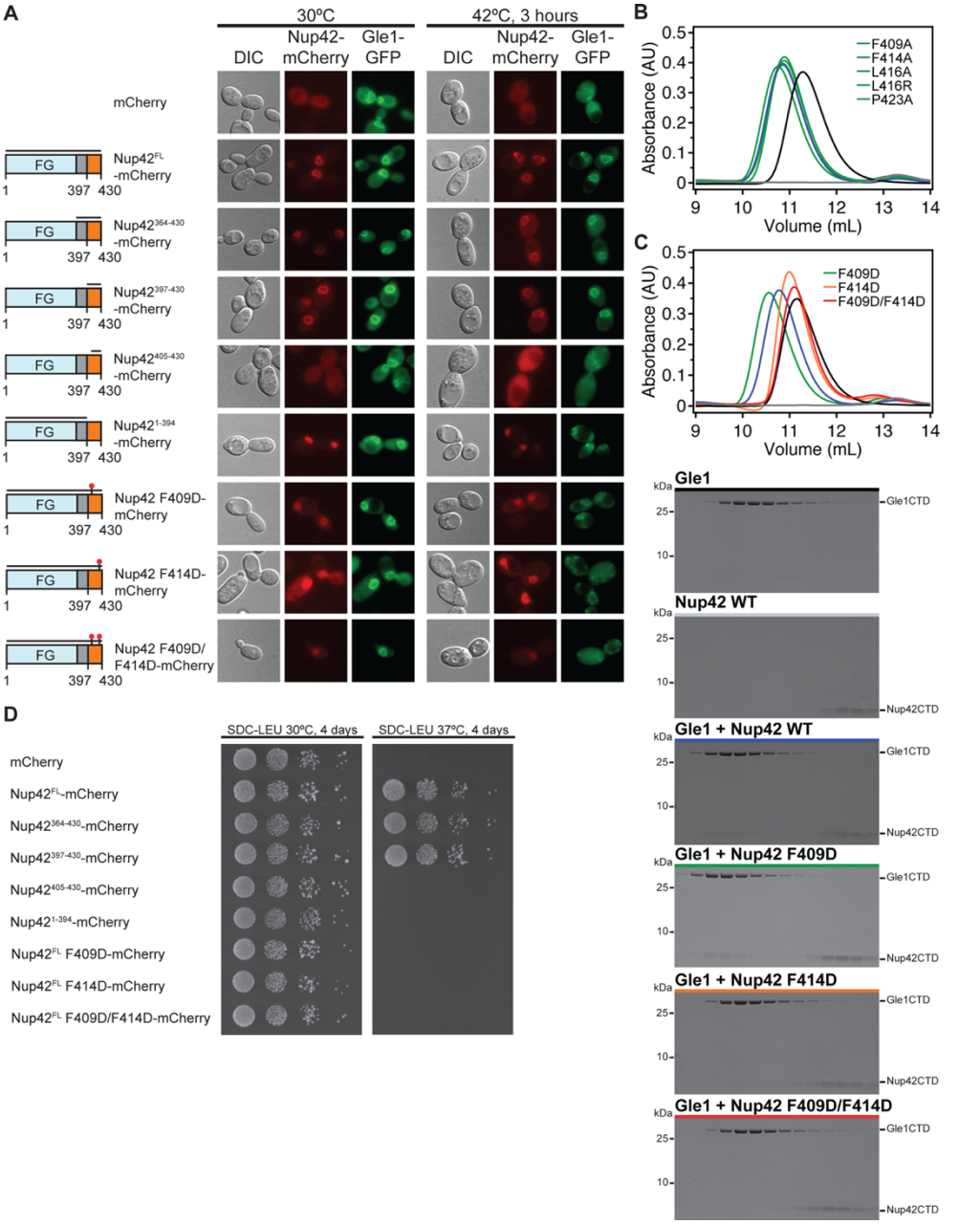


Fig. S3.

Analysis of interaction between Nup42^{CTD} and Gle1^{CTD}. (A) *in vivo* localization analysis in *S. cerevisiae* of Gle1-GFP and Nup42-mCherry variants. Constructs are schematized on the left. (B, C) Size exclusion chromatography analysis of the effect of mutations in Nup42^{CTD} on Gle1^{CTD} binding. SEC profiles of Gle1^{CTD}(black), Nup42$^{CTD(WT)}$ (grey) and Gle1 preincubated with Nup42$^{CTD(WT)}$ (blue) are shown as controls. SEC profiles of Nup42^{CTD} mutants preincubated with Gle1^{CTD} are colored according to the measured effect, no effect (green), reduced binding (orange) and complete disruption (red). The gray bar indicates the fractions visualized with Coomassie-stained SDS-PAGE gels. (D) Growth analysis of *S. cerevisiae* strains carrying the indicated Nup42-mCherry variants. Serial dilutions of the respective cells were spotted onto SDC-LEU plates and grown for 4 days at 30 and 37°C.

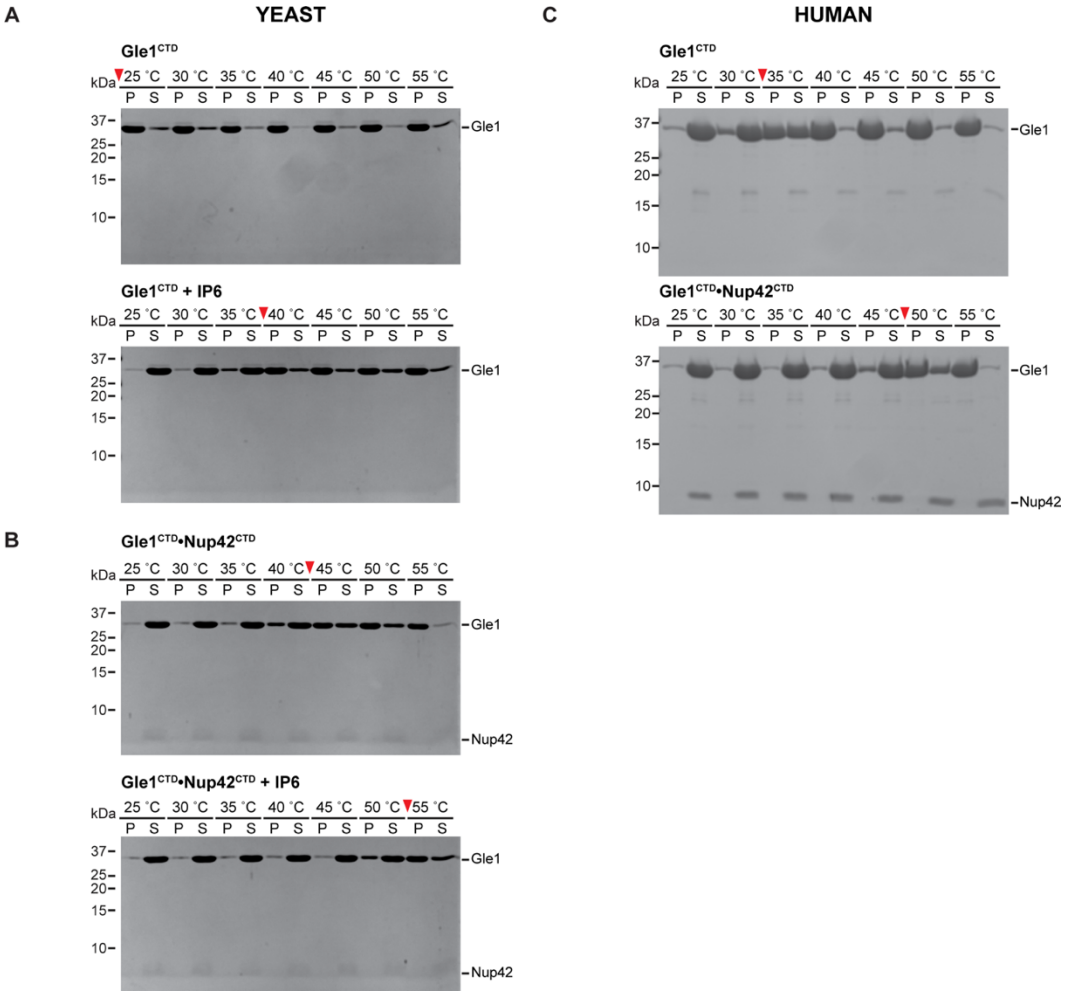


Fig. S4.

Thermostability assay. (A) *S. cerevisiae* Gle1^{CTD} was incubated at the indicated temperatures for 30 minutes prior to centrifugation in the absence or presence of IP6. (B) *S. cerevisiae* Gle1^{CTD} Nup42^{CTD} in the absence or presence or absence of IP6 was incubated at the indicated temperatures for 30 minutes prior to centrifugation. (C) *H. sapies* Gle1^{CTD} in the absence or presence of Nup42^{CTD} was incubated at the indicated temperatures for 30 minutes prior to centrifugation. Pelleted (P) and soluble (S) fractions were analyzed by SDS-PAGE and visualized by Coomassie staining. Red arrows indicate the temperature increment at which more than 50% of total Gle1 pelleted *in vitro*.

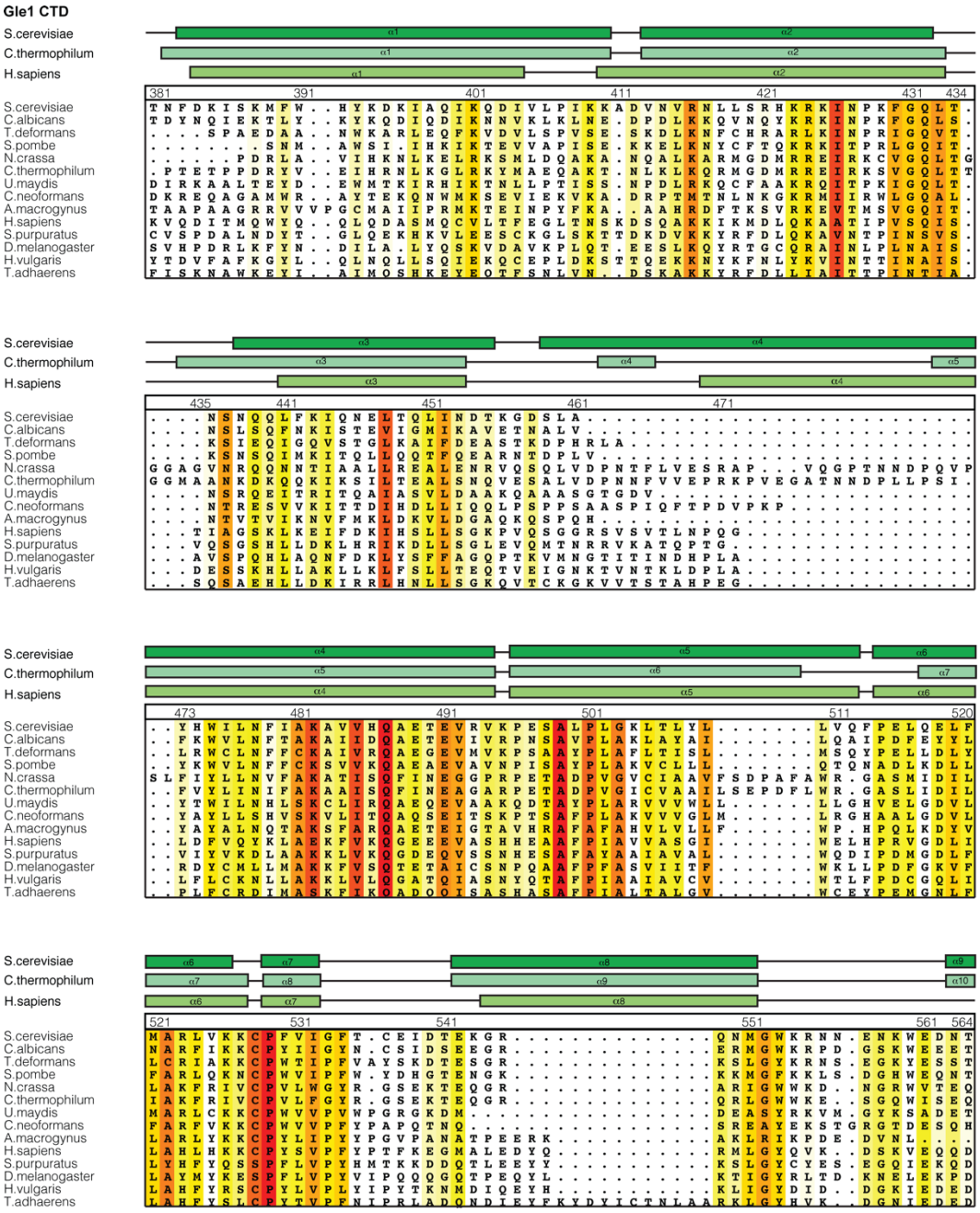


Fig. S5.

Multispecies sequence alignment of Gle1^{CTD}. Sequences from fourteen diverse species were aligned and colored by sequence similarity according to the BLOSUM62 matrix from white (less than 40% similarity), to yellow (55% similarity), to red (100% identity). The numbering is according to the *H. sapiens* protein. The secondary structure is indicated above the sequences as rectangles (α-helices) and lines (unstructured regions) for the *S. cerevisiae*, *C. thermophilum* and *H. sapiens* proteins. Secondary structure elements are colored according to Figure 2.

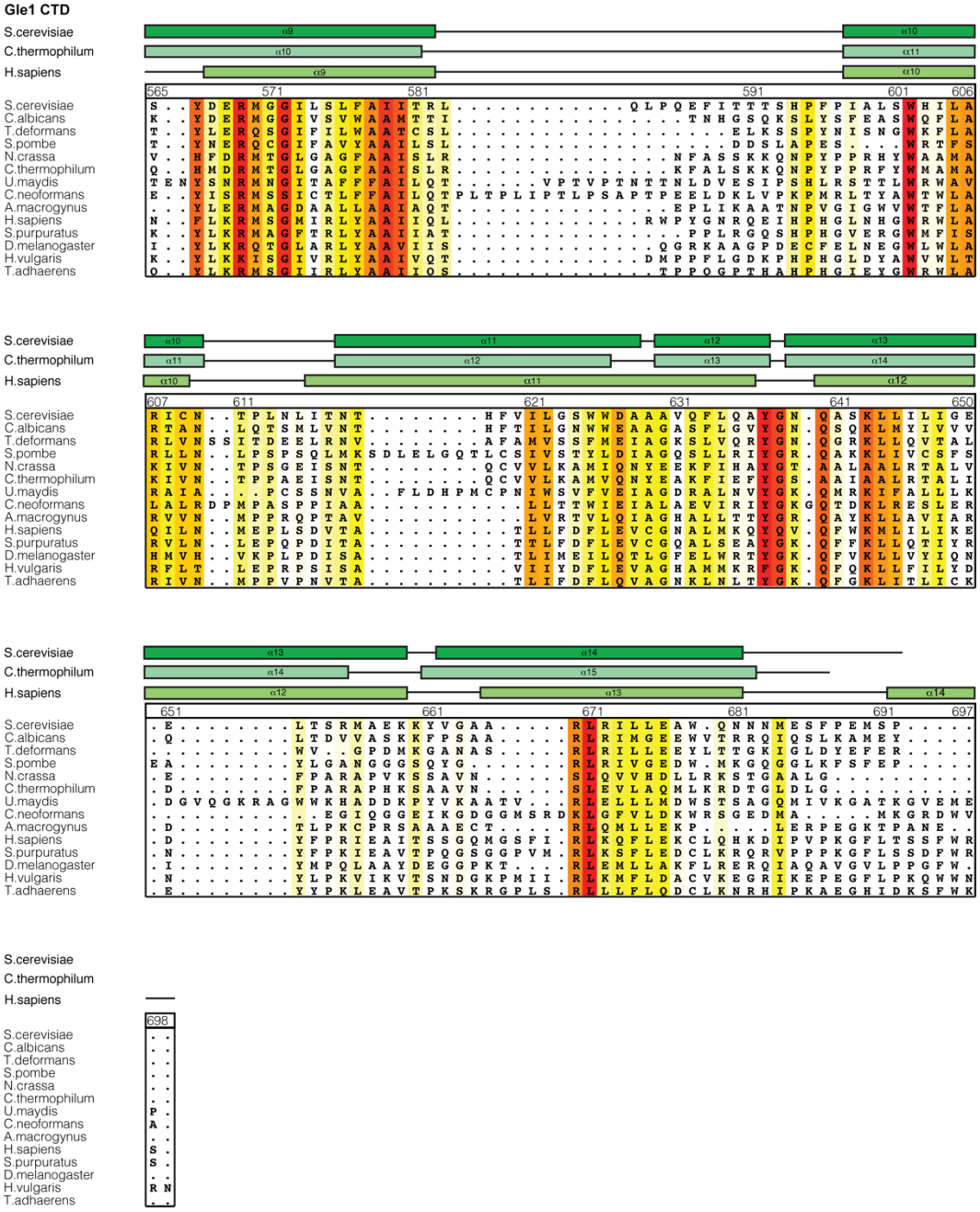


Fig. S5 continued.

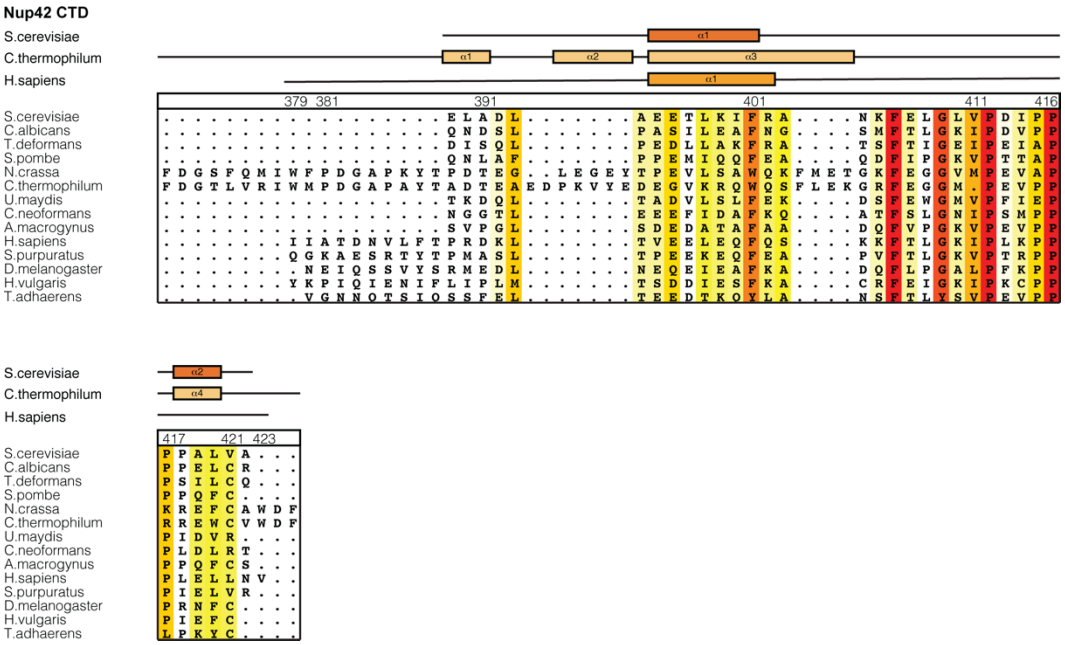


Fig.S6.

Multispecies sequence alignment of Nup42^{CTD}. Sequences from fourteen diverse species were aligned and colored by sequence similarity according to the BLOSUM62 matrix from white (less than 40% similarity), to yellow (55% similarity), to red (100% identity). The numbering is according to the *H. sapiens* protein. The secondary structure is indicated above the sequences as rectangles (α-helices) and lines (unstructured regions) for the *S. cerevisiae*, *C. thermophilum* and *H. sapiens* proteins. Secondary structure elements are colored according to Figure 2.

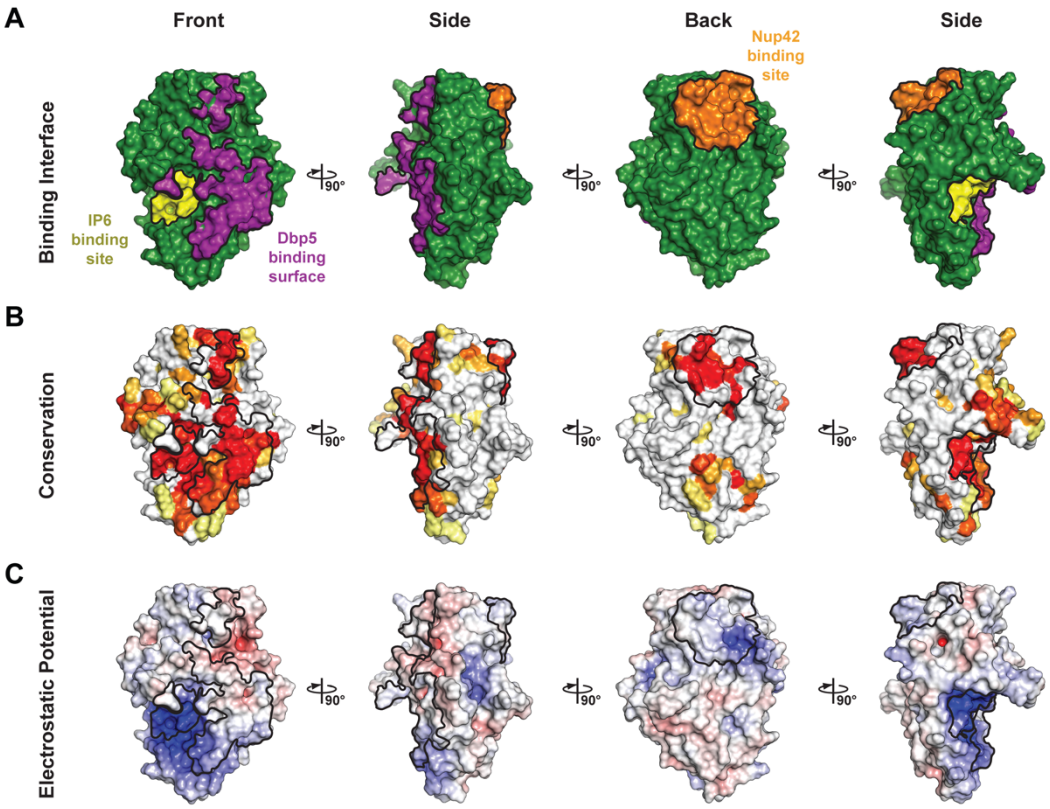


Fig. S7.

Surface properties of *S. cerevisiae* Gle1^{CTD}. Surface representations of Gle1^{CTD} in four orientations related by 90° rotations. The Nup42^{CTD}, IP6, and Dbp5 binding interfaces are outlined in black. (A) Identification of *S. cerevisiae* Gle1^{CTD} binding surfaces. IP6 binding site is colored in yellow, Dbp5 binding interface is colored in purple and Nup42CTD interface is colored in orange. (B) Surface representation colored according to sequence conservation for Saccharomycotina and Schizosaccharomycetes using an alignment containing the species *S. cerevisiae*, *Z. rouxii*, *K. lactis*, *C. albicans*, *Y. liplytica, T. deformans, S.complicata,* and *S. pombe*. (C) Surface representation colored according to electrostatic potential from -10k_B T/e (red) to +10k_B T/e (blue).

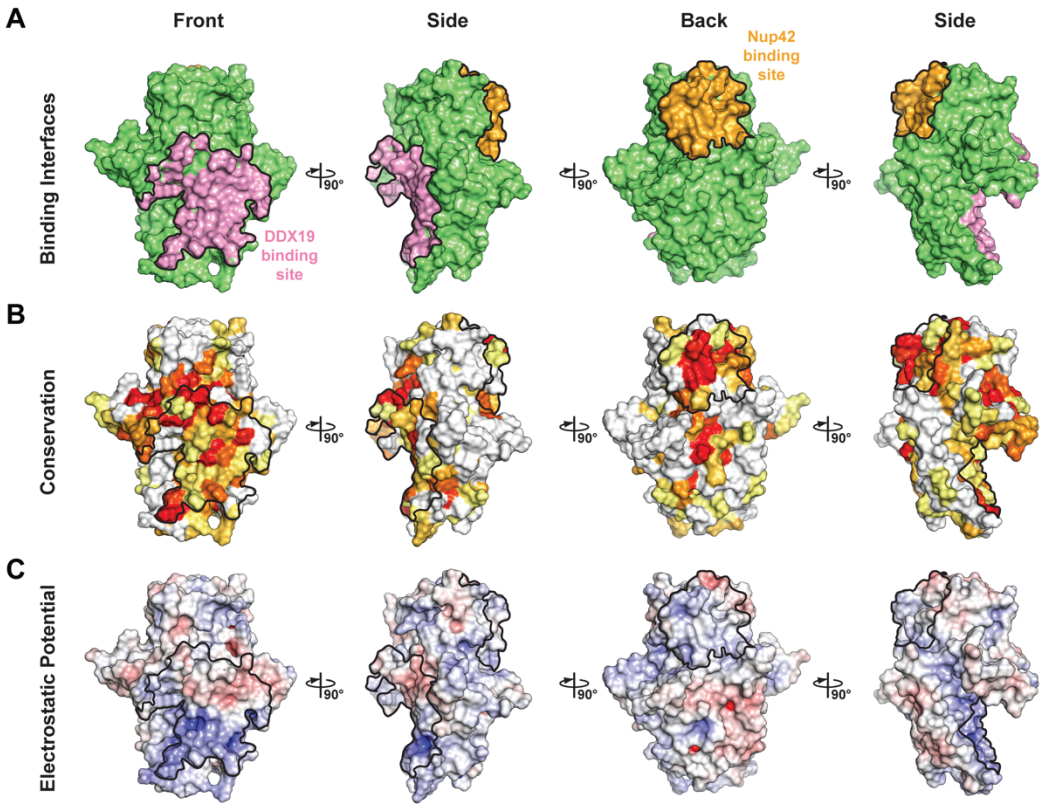


Fig. S8.
Surface properties of *H. sapiens* Gle1^{CTD}. Surface representations of Gle1^{CTD} in four orientations related by 90° rotations. The Nup42^{CTD} and DDX19 binding interfaces are outlined in black. (A) Identification of *H. sapiens* Gle1^{CTD} binding surfaces. DDX19 binding interface is colored in pink and Nup42^{CTD} interface is colored in orange. (B) Surface representation colored according to sequence conservation for metazoans using an alignment containing the species *H. sapiens*, *X. tropicalis*, *D. rerio*, *S. purpuratus*, *T. castaneum*, *D. melanogaster*, *C. teleta*, *H. vulgaris*, *A. digitifera*, and *T. adhaerens*. (C) Surface representation colored according to electrostatic potential from -10k_B T/e (red) to +10k_B T/e (blue).

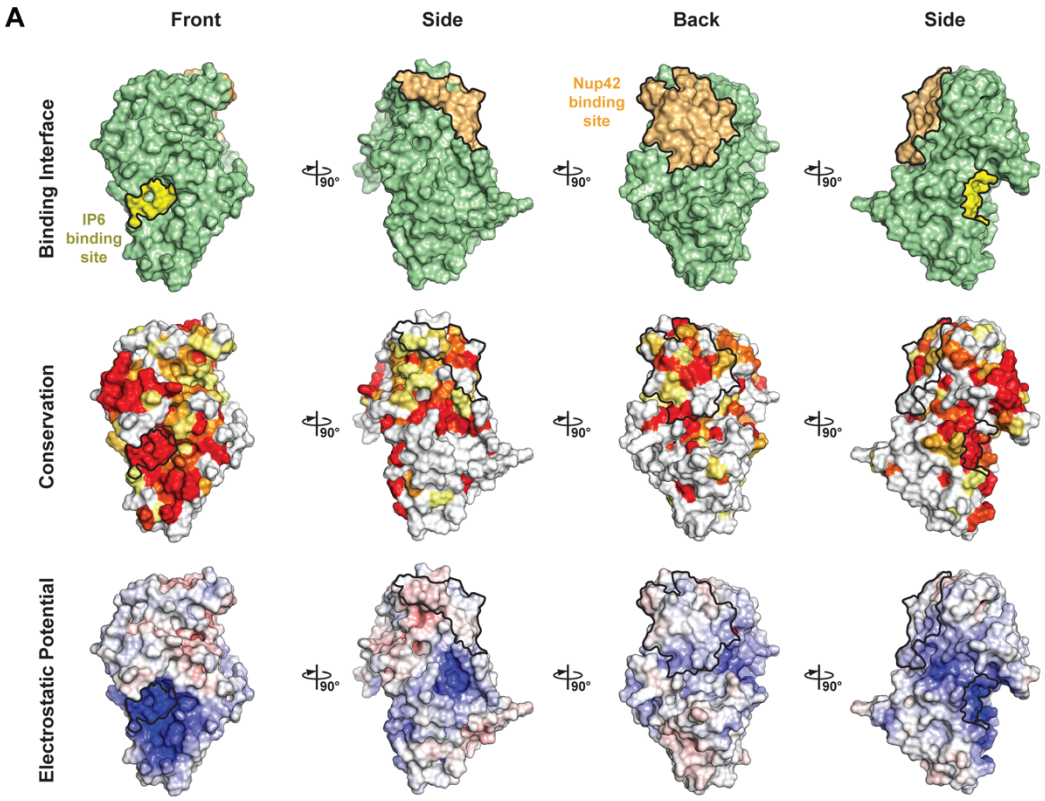


Fig. S9.
Surface properties of *C. thermophilum* Gle1^{CTD}. Surface representations of Gle1^{CTD} in four orientations related by 90° rotations. The Nup42^{CTD} and IP6 binding interfaces are outlined in black. (A) Identification of *C. thermophilum* Gle1^{CTD} binding surfaces. IP6 binding site is colored in yellow and Nup42^{CTD} interface is colored in wheat. (B) Surface representation colored according to sequence conservation for Pezizomycotina using an alignment containing the species *C. thermophilum*, *X. heveae, S. sclerotiorum, C. militaris, P. digitatum,* and *F. verticilliodes*. (C) Surface representation colored according to electrostatic potential from -10k_B T/e (red) to +10k_B T/e (blue).

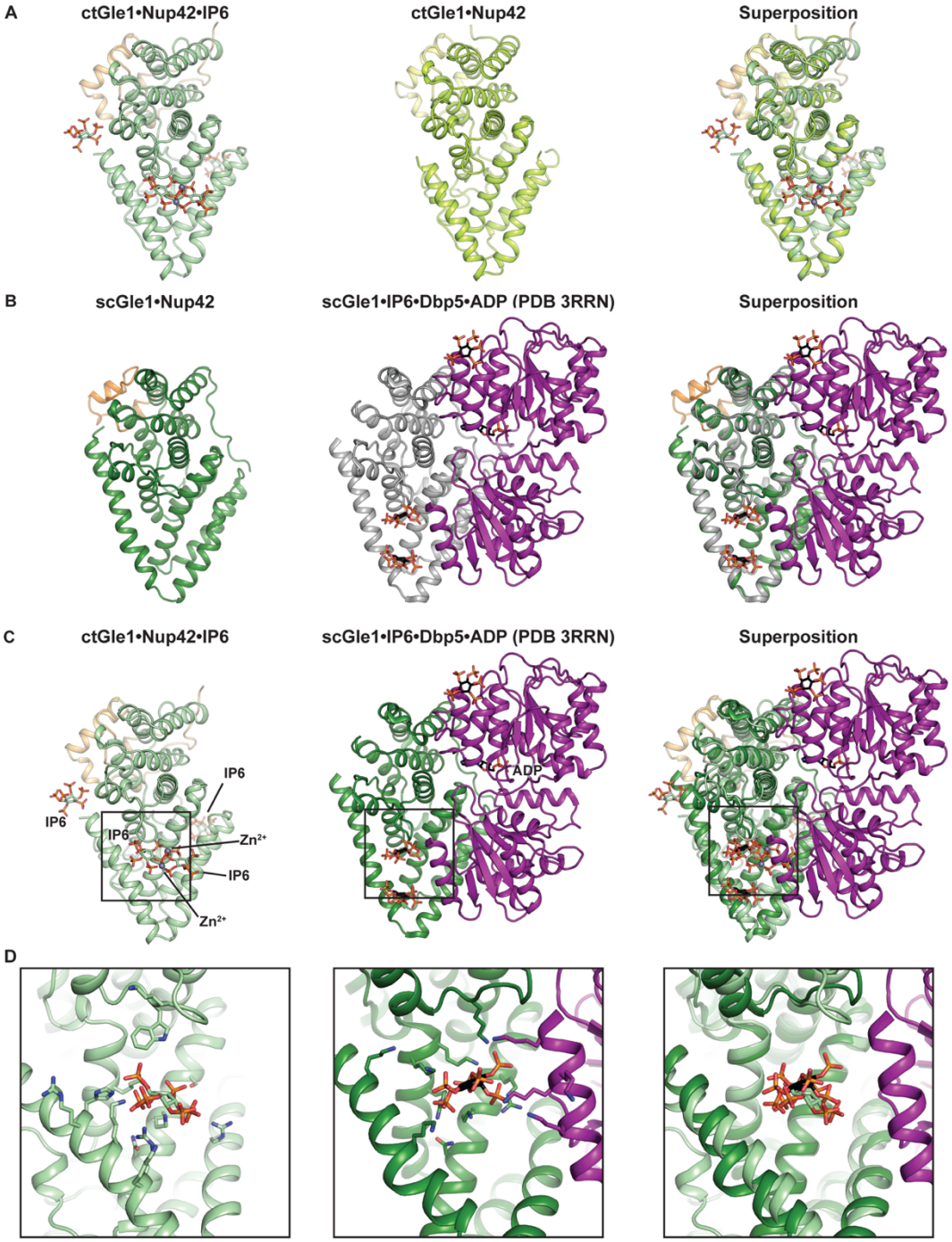


Fig. S10.

Analysis of the effect of IP_6 binding on $Gle1^{CTD}$ in fungi. (A) Comparison of the *C. thermophilum* $Gle1^{CTD}$•$Nup42^{CTD}$•IP_6 and $Gle1^{CTD}$•$Nup42^{CTD}$ structures. (B) Comparison of the *S. cerevisiae* $Gle1^{CTD}$•$Nup42^{CTD}$ and *S. cerevisiae* $Gle1^{CTD}$•IP6•Dbp5•ADP, highlighting the fact that $Nup42^{CTD}$ has a minimal effect on the conformation of Gle1 and does not have a direct involvement in the interaction with Dbp5. (C) and (D) Comparison of the *C. thermophilum* $Gle1^{CTD}$•$Nup42^{CTD}$•IP_6 and *S. cerevisiae* $Gle1^{CTD}$•IP6•Dbp5•ADP structures highlighting the conservation of IP6 binding pocket and Dbp5 recognition mode in fungi.

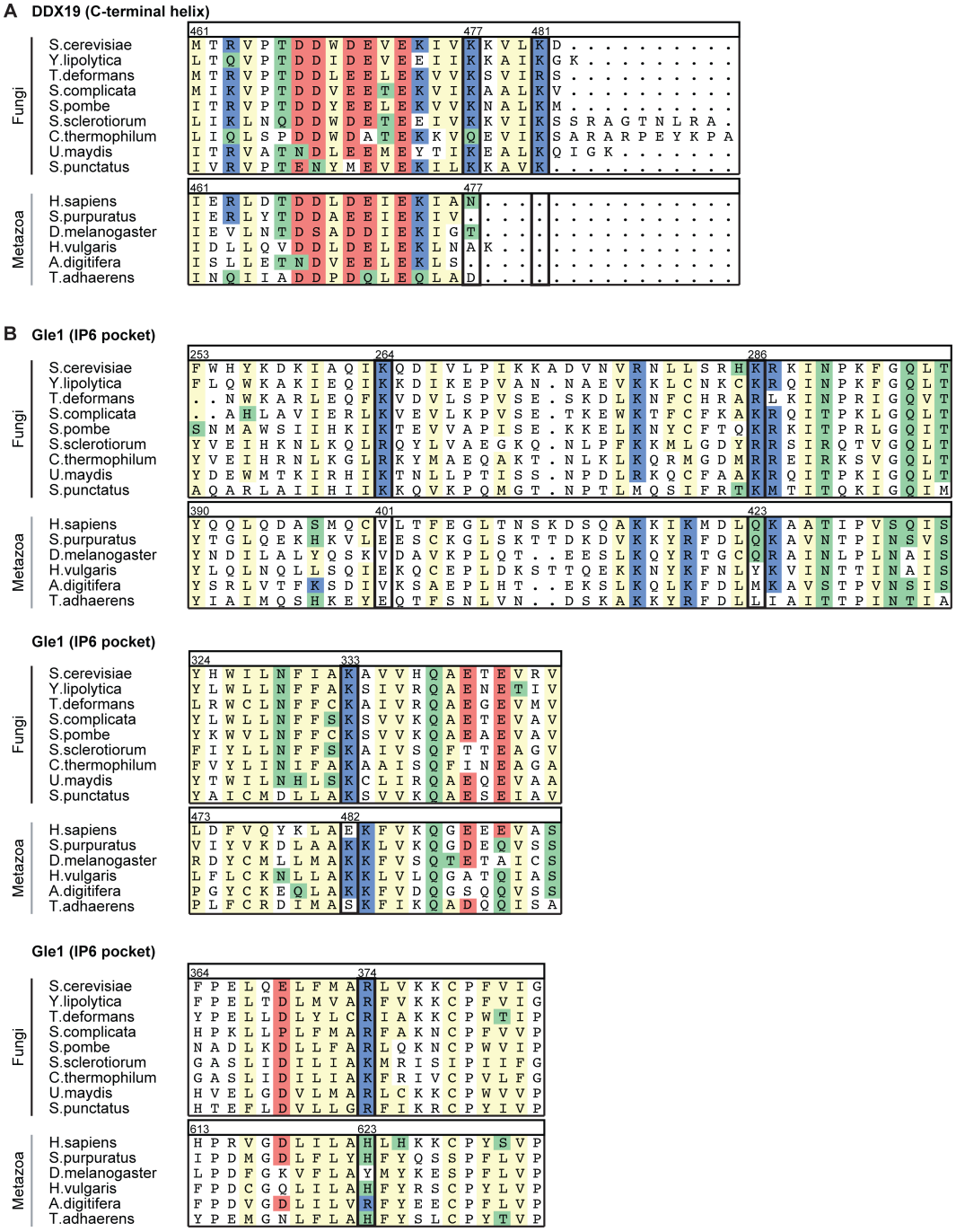


Fig. S11.

Multispecies sequence alignment of DDX19 C-terminal helix and Gle1 IP6 pocket. Basic residues conserved in fungi are outlined with black boxes. Amino acid residues are colored according to Clustal X color scheme.

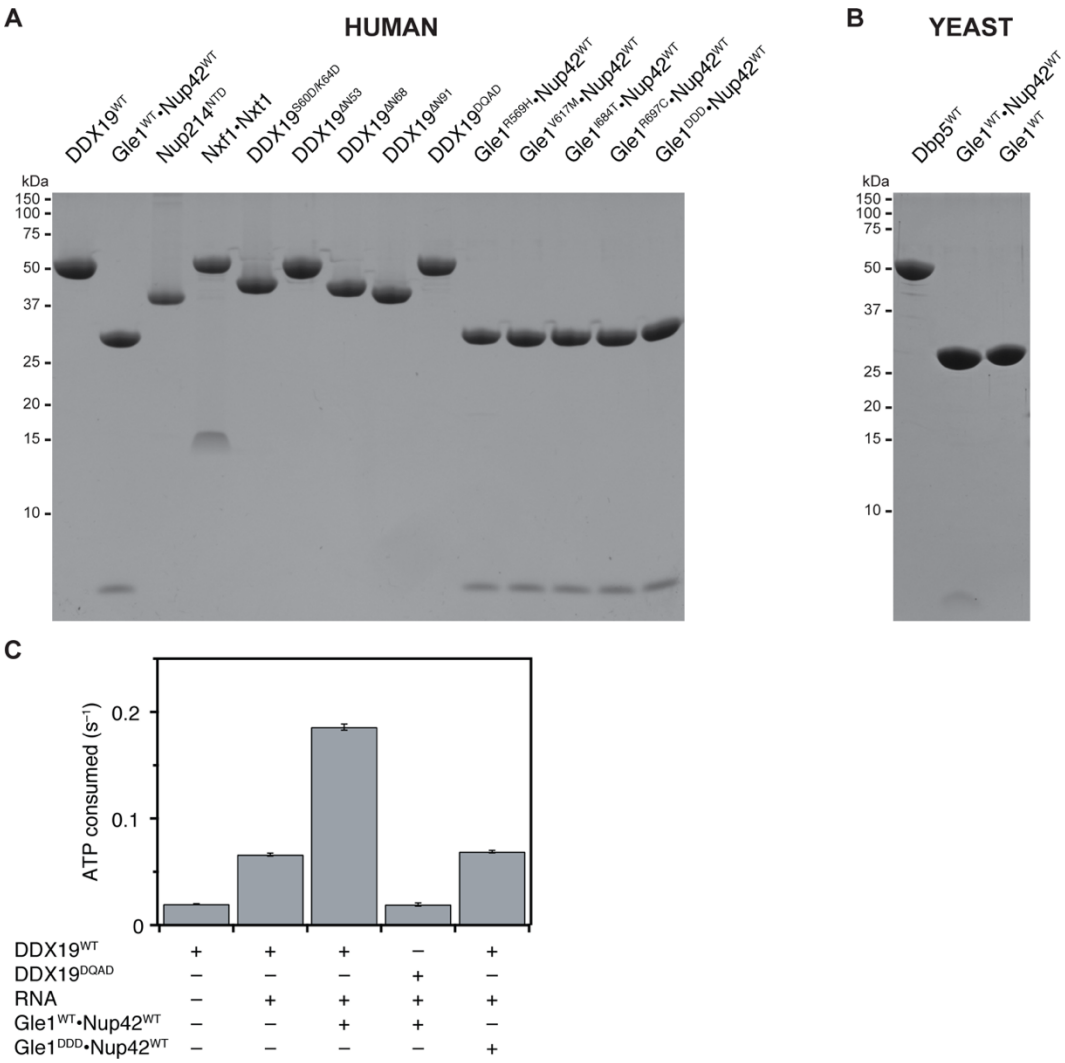


Fig. S12.

Purified proteins used in the steady state colorimetric ATPase activity assays. (A) and (B) SDS-PAGE analysis of the purified proteins used in the activity assays. (C) Steady-state ATPase rates of previously established mutants that abolish stimulation in DDX19 or Gle1^{CTD}, indicating that no activity can be attributed to contaminating factors.

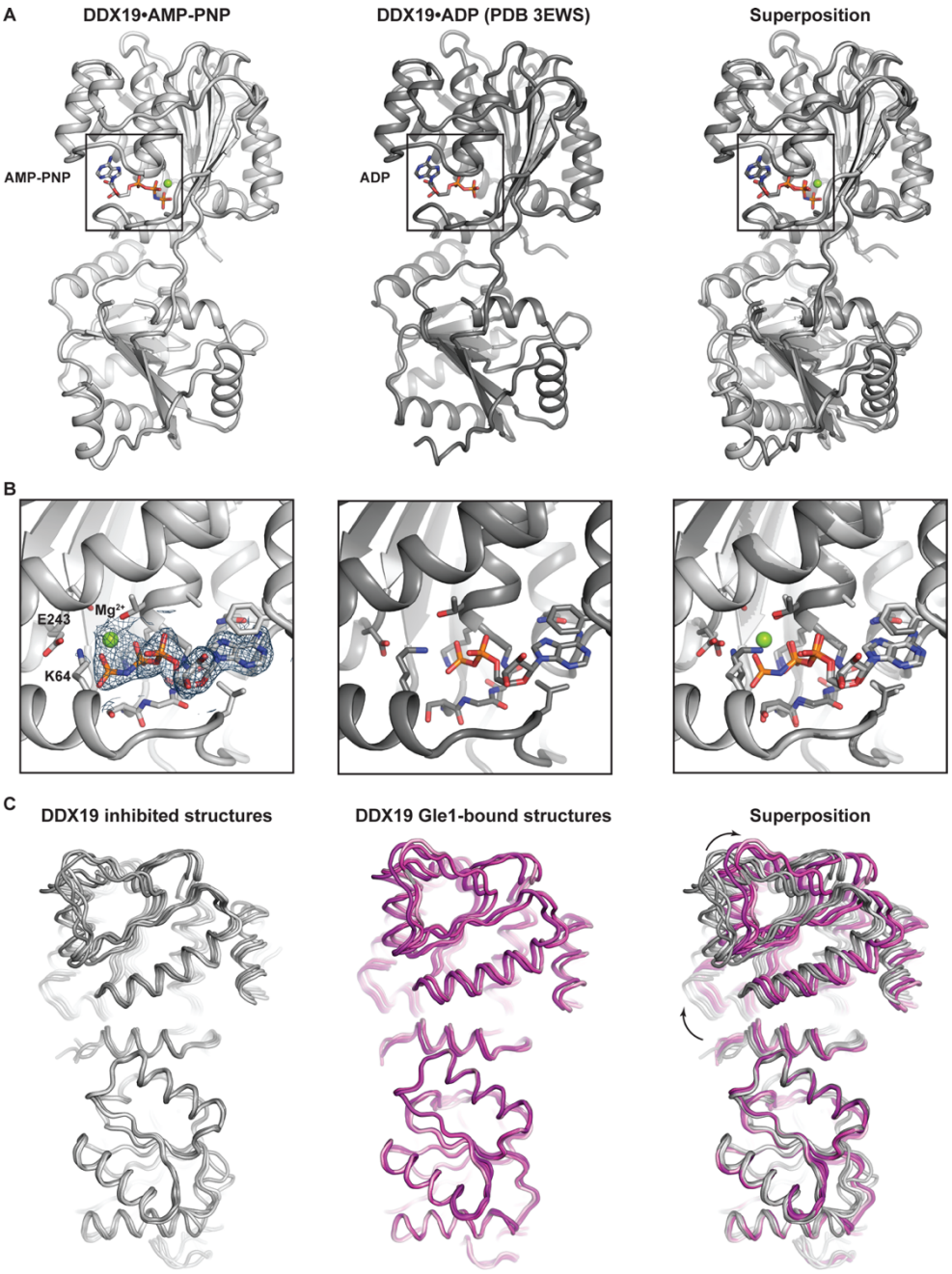


Fig. S13.
Structure of DDX19 AMP-PNP and analysis of DDX19 conformations. (A) Left, Crystal structure of *H. sapiens* DDX19 AMP-PNP, middle, Crystal structure of *H. sapiens* DDX19 ADP (PDB code 3EWS), right, superposition of the two structures highlighting the conformational changes associated to the binding of the different nucleotides. (B) Zoom view of the nucleotide binding pocket of the structures presented in (A). (C) Left, Superposition of DDX19 inhibited structures, middle DDX19 Gle1-bound structures, right, superposition of DDX19 inhibited structures (grey) and DDX19 Gle1-bound structures (pink) highlighting the movement of the N-terminal RecA domain upon Gle1 binding. Arrows indicate the rotation relating the different conformations.

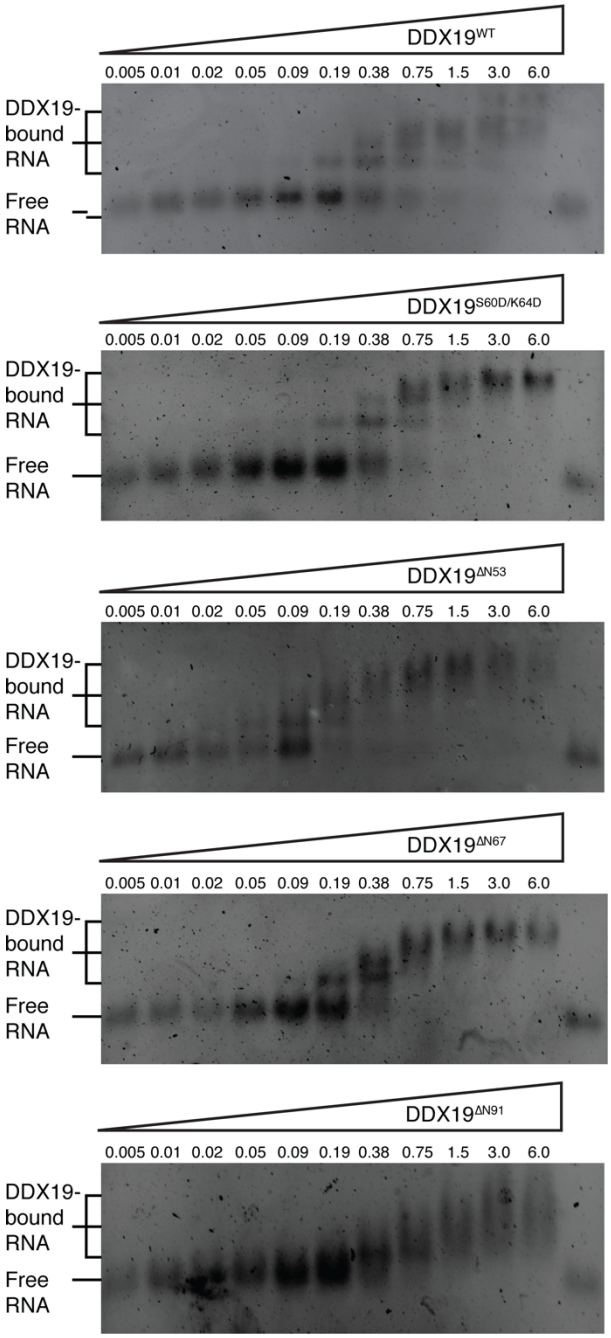


Fig. S14.
RNA binding affinity of DDX19 variants. Electrophoretic mobility shift analysis of RNA upon DDX19 binding. Impact of DDX19 N-terminal extension and DDX19 inhibitory helix on RNA binding affinity.

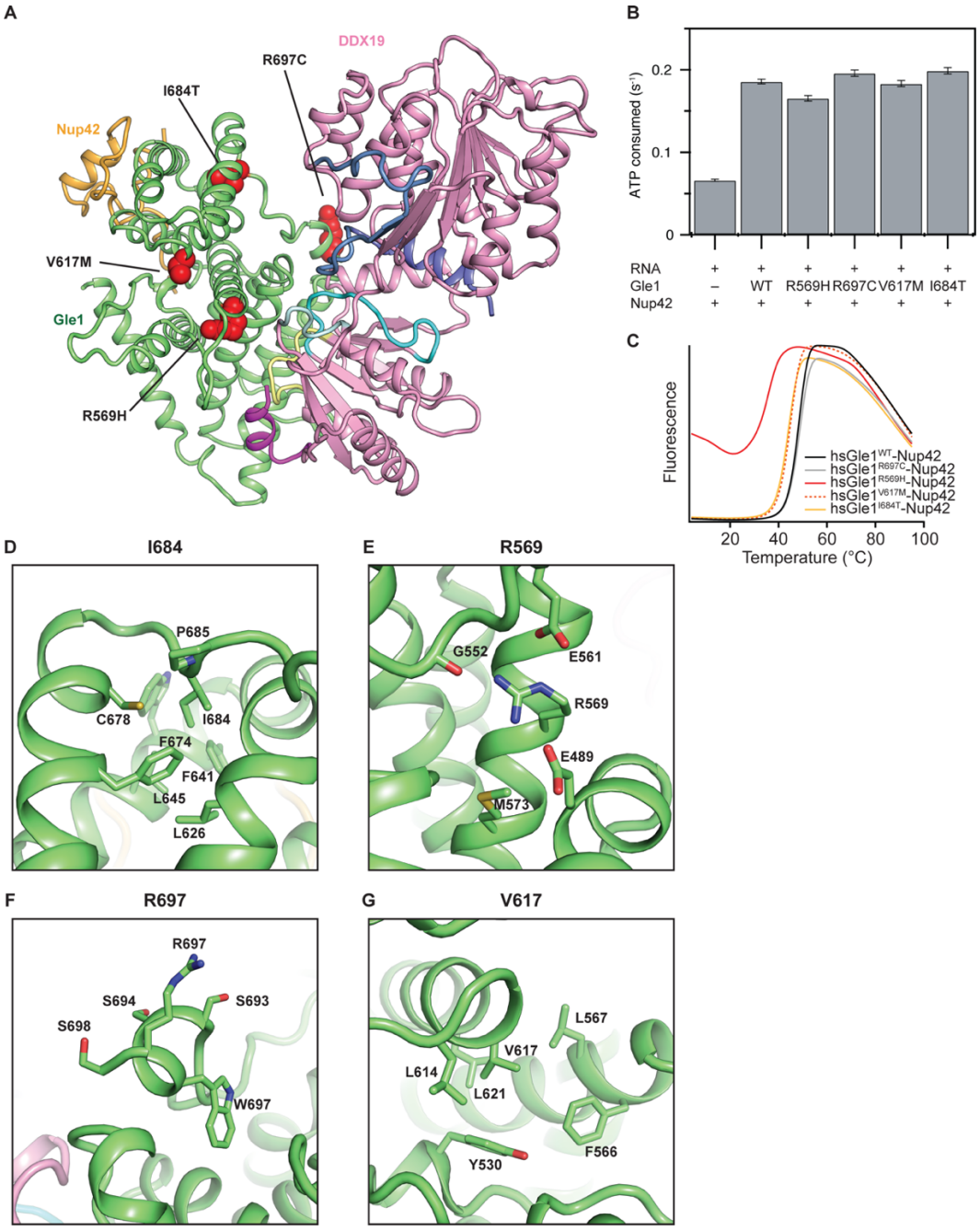


Fig. S15.

Impact of Gle1 disease-related mutations. (A) Mapping of Gle1 disease related mutations onto the structure of *H. sapiens* Gle1 Nup42^{CTD} DDX19$^{\Delta N1}$ ADP (colors as in Figure 4A), (B) Impact of Gle1 disease related mutations on Gle1's ability to stimulate DDX19 ATPase activity. (C) Effect of Gle1 disease related mutations on Gle1-Nup42^{CTD} thermostability evaluated by differential scanning fluorimetry. (D-G) Zoom view of the local environments of residues mutated in disease.

Table S1.

X-ray crystallography analysis of *S. cerevisiae* and *H. sapiens* Gle1 ^{CTD}•Nup42^{CTD}

Data collection		
Protein	scGle1^{CTD}•Nup42^{CTD}	hsGle1^{CTD}•Nup42^{CTD}
Synchrotron	SSRL[a]	APS[b]
Beamline	BL12-2	23-ID-D
Space group	$P4_32_12$	C2
Cell dimensions		
a, *b*, *c* (Å)	64.5, 64.5, 361.7	163.7, 69.2, 93.0
α, β, γ (°)	90.0 90.0 90.0	90.0, 90.6, 90.0
Wavelength	0.9795	0.9794
Resolution (Å)	47.3–1.75 (1.81–1.75)	46.5–2.8 (2.9–2.8)
R_{meas} (%)[c]	12.4 (127.0)	10.7 (144.3)
R_{pim} (%)[c]	2.5 (39.8)	5.4 (73.5)
$CC_{1/2}$[c]	99.9 (57.3)	99.8 (53.6)
$< I / \sigma I >$[c]	16.4 (1.3)	10.1 (1.1)
Completeness (%)[c]	99.3 (93.5)	99.2 (97.8)
No. of observations	1,862,860 (76, 034)	98,150 (8,916)
No. of unique reflections[c,d]	85,335 (7,866)	25,969 (2,551)
Redundancy[c]	21.8	3.8 (3.5)
Refinement		
Resolution (Å)	47.3-1.75	46.8-2.8
No. of reflections	85,335	20,826
No. of reflections test set		1,067 (5.1%)
R_{work} / R_{free}	19.9/22.3	25.2/28.6
No. atoms (non-hydrogen)	6,258	5,785
Protein	5,569	5,775
Water	667	-
Ligand/Ions	12	10
B-factors	36.4	58.2
Protein	35.7	58.2
Water	42.1	-
Ligand/Ions	38.1	75.4
RMSD		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	.780	0.800
Ramachandran plot[f]		
Favored (%)	96.9	93.2
Additionally allowed (%)	2.2	6.7
Outliers (%)	0.9	0.1
MolProbity		
Clashscore[d]	3.28	7.22
Molprobity score[d]	1.59	2.27

[a]APS, Advanced Photon Source
[b]SSRL, Stanford Synchotron Radiation Lightsource
[c]Highest-resolution shell is shown in parentheses
[d]As determined by MolProbity

Table S2.

X-ray crystallography analysis of *C. thermophilum* Gle1 ^{CTD}•Nup42^{CTD}

Data collection			
Protein	ctGle1^{CTD}•Nup42^{CTD}	ctGle1^{CTD}•Nup42^{CTD}•IP_6	ctGle1^{CTD}•Nup42^{CTD}•IP_6
Synchrotron	APS[a]	SSRL[b,d]	SSRL[b]
Beamline	23-ID-D	BL12-2	BL12-2
Space group	$P2_12_12_1$	C2	C2
Cell dimensions			
a, *b*, *c* (Å)	84.5, 93.1, 229.7	118.7, 72.8, 117.7.7	119.2 73.3 117.7
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 94.2, 90.0
		Se Peak	
Wavelength	1.0332	0.9792	1.0332
Resolution (Å)	48.9-2.7	41.6-3.2	35.0-2.1
R_{meas} (%)[c]	9.2 (93.5)	14.1 (78.3)	12.4 (174.7)
R_{pim} (%)[c]	2.5 (28.3)	4.5 (24.7)	3.6 (66.1)
$CC_{1/2}$[c]	100.0 (68.5)	100.0 (68.5)	99.9 (70.9)
$< I / \sigma I >$[c]	21 (2.3)	11.6 (3.0)	10.7 (1.3)
Completeness (%)[c]	99.8 (97.6)	98.3 (98.6)	98.2 (96.1)
No. of observations	700,489	158,483	389,707
No. of unique reflections[c,d]	53,444 (5,143)	16,420 (1,658)	58,076 (5,657)
Redundancy[c]	13.1	9.7	6.7
Refinement			
Resolution (Å)	48.4-2.6		35.0-2.2
No. of reflections	53,433		52,738
No. of reflections test set	4,114 (3.74%)		2,587 (4.9%)
R_{work} / R_{free}	25.2/28.8		20.5/23.4
No. atoms	11,485		6,684
Protein	11,350		6,071
Water			
Ligand/Ions	1,431		316
B-factors	88.5		48.1
Protein	88.5		46.3
Water	-		-
Ligand/Ions	88.1		46.9
RMSD			
Bond lengths (Å)	0.005		0.012
Bond angles (°)	0.800		0.510
Ramachandran plot[f]			
Favored (%)	94.0		98.0
Additionally allowed (%)	4.2		2.1
Outliers (%)	1.3		0.0
MolProbity			
Clashscore[e]	8.50		3.11
Molprobity score[e]	1.99		1.13

[a]APS, Advanced Photon Source
[b]SSRL, Stanford Synchotron Radiation Lightsource
[c]Highest-resolution shell is shown in parentheses
[d]Friedel pairs were merged
[e]As determined by MolProbity

Table S3.

X-ray crystallography analysis of DDX19$^{\Delta N1}$ and complexes

Data collection				
Protein	hsGle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$•ADP[d,f]		hsGle1^{CTD}•Nup42^{CTD}•DDX19$^{\Delta N53}$•AMP-PNP	DDX19$^{\Delta N53}$•AMP-PNP
Synchrotron	APS[a]		NSLS-II[b]	APS[a]
Beamline	23-ID-D		17-ID-2	23-ID-D
Space group	$P2_1$		$P2_1$	$P2_1$
Cell dimensions				
a, *b*, *c* (Å)	87.7, 74.7, 146.8		87.6, 73.4, 145.3	83.4, 45.6, 127.6
α, β, γ (°)	90.0, 94.8, 90.0		90.0, 95.1, 90.0	90.0, 97.0, 90.0
Wavelength	1.0333	1.0333	0.9793	1.0333
Resolution (Å)	50.0–3.6 (3.7–3.6)	50.0–3.6 (3.8–3.6)	48.2–3.4 (3.5–3.4)	42.2-2.2 (2.3-2.2
R_{meas} (%)[c]	15.9 (223.2)	13.8 (114.2)	29.8 (406.9)	10.7 (168.5)
R_{pim} (%)[c]	6.1 (83.3)	5.3 (43.6)	9.3 (124.9)	6.7 (2.4)
$CC_{1/2}$[c]	99.7 (47.9)	99.7 (68.3)	99.7 (48.3)	99.9 (47.4)
$< I / \sigma I >$[c]	8.5 (1.0)	9.6 (2.0)	6.6 (0.9)	12.0 (1.2)
Completeness (%)[c]	98.6 (98.9)	86.5 (21.0)	100.0 (100.0)	99.0 (97.0)
No. of observations	148,572 (15,143)	129,022 (3,035)	261,742 (26,462)	335,582 (29,666)
No. of unique reflections[c,d]	21,685 (2,157)	18,971 (454)	25,636 (2,540)	49,144 (4,795)
Redundancy[c]	6.9 (7.0)	6.8 (6.7)	10.2 (10.4)	6.8
Refinement				
Resolution (Å)		46.1–3.6	45.8-3.4	47.6-2.2
No. of reflections		18,966	21,290	49,087
No. of reflections test set		1,742 (9.2%)	1,979 (9.3%)	2,000(4.1%)
R_{work} / R_{free}		24.4/29.1	26.0/31.2	22.9/27.1
No. atoms (non-hydrogen)		12,517	12,583	6,812
Protein		12,445	12,491	6,651
Water		-	-	-
Ligand/Ions		72	92	72
B-factors		136.9	105.9	81.9
Protein		136.9	105.9	82.6
Water		-	-	-
Ligand/Ions		122.4	100.5	55.2
RMSD				
Bond lengths (Å)		0.004	0.004	0.003
Bond angles (°)		0.740	0.722	0.590
Ramachandran plot[f]				
Favored (%)		93.2	93.9	96.0
Additionally allowed (%)		5.6	5.8	4.1
Outliers (%)		1.2	0.3	0.2
MolProbity				
Clashscore[e]		5.67	6.00	3.62
Molprobity score[e]		2.48	2.48	1.45

[a]APS, Advanced Photon Source
[b]NSLSII, National Synchotron Light Source II
[c]Highest-resolution shell is shown in parentheses
[d]Refinement was performed with ellipsoidally truncated data
[e]As determined by MolProbity
[f]As a reference, two different high-resolution cutoffs are shown

Table S4.

Bacterial expression constructs and expression conditions

#	Protein	Residues	Expression vector	Restriction sites 5', 3'	N-terminal overhang	C-terminal overhang	Expression conditions
1	hsNup155 CTD	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
2	hsGle1 N	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
3	hsNup98 ΔFG	440-884	pET28a-SUMO	BamHI, NotI	S	-	18 °C / 18 hours
4	hsNup155 CTD E1146A	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
5	hsNup155 CTD K1147A	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
6	hsNup155 CTD L1182A	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
7	hsNup155 CTD Y1189A	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
8	hsNup155 CTD F1192A	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
9	hsNup155 CTD I1206A	870-1391	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
10	hsGle1 N L11A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
11	hsGle1 N K19A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
12	hsGle1 N L22A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
13	hsGle1 N C23A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
14	hsGle1 N Y24A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
15	hsGle1 N R25A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
16	hsGle1 N W28A	2-33	pET28a-SUMO	BamHI, NotI	Smt3p-S	-	30 °C / 2 hours
17	scGle1 CTD	244-538	pETDuet PreS	NdeI, XhoI	M	-	18 °C / 18 hours
18♦	scGle1 CTD scNup42 CTD	244-538 397-430	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
19	hsGle1 CTD	382-698	pET28a-SUMO	BamHI, NotI	S	-	18 °C / 18 hours
20♦	hsGle1 CTD hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
21♦	ctGle1 CTD	216-519	pET28a-PreS	NdeI, NotI	GPHM	-	37C °C / 3 hours co-expressed with ctNup42
22♦	ctNup42 CTD	494-558	pGex6P1-PreS	EcoRI, XhoI	GPLGSPEF		37C °C / 3 hours co-expressed with ctGle1
23	scNup42 CTD	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
24	scNup42 CTD F409A	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
25	scNup42 CTD F414A	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
26	scNup42 CTD L416A	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
27	scNup42 CTD L416R	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
28	scNup42 CTD P423A	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
29	scNup42 CTD F409D	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
30	scNup42 CTD F414D	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
31	scNup42 CTD F409D/F414D	397-430	pET28a-SUMO	BamHI, XhoI	S	YALEHHHHHH	37 °C / 2 hours
32	scDbp5	1-482	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
33♦	hsDDX19	1-479	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
34	ctDbp5	1-477	pETMCN-SUMO	BamHI, NotI	S	-	18 °C / 18 hours
35	ctDbp5 E219Q	1-477	pETMCN-SUMO	BamHI, NotI	S	-	18 °C / 18 hours
36	hsDDX19 ΔN53	54-479	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
37	hsGle1 CTD H495A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
38	hsGle1 CTD E491A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
39	hsGle1 CTD E490A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
40	hsGle1 CTD Q487A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
41	hsGle1 CTD K486A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
42	hsGle1 CTD K479A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
43	hsGle1 CTD K479A/K486A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
44	hsGle1 CTD Q423A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
45	hsGle1 CTD K416A/K419A hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
46	hsDDX19 ΔN67	68-479	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
47	hsDDX19 ΔN91	92-479	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
48	hsDDX19 S60D/K64D	1-479	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
49	hsNup214 NTD	1-450	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours
50	hsNxf1 ΔN hsNxt1	110-619 1-140	pETDuet	BamHI, EcoRI NdeI, KpnI	S	- -	18 °C / 18 hours
51	hsGle1 CTD G666D/I669D/Q673D hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
52	hsGle1 CTD R569H hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	-- -	18 °C / 18 hours

#	Protein	Residues	Expression vector	Restriction sites 5', 3'	N-terminal overhang	C-terminal overhang	Expression conditions
53	hsGle1 CTD V617M hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
54	hsGle1 CTD I684T hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
55	hsGle1 CTD R697C hsNup42 CTD	382-698 379-423	pETDuet PreS	NdeI, XhoI BamHI, NotI	M GPGS	- -	18 °C / 18 hours
56	hsDDX19 E243Q	1-479	pET28a-PreS	NdeI, NotI	GPHM	-	18 °C / 18 hours

♦Constructs that were used for crystallization

Table S5.

Yeast constructs

Plasmid	Protein	Residues (Mutations)	Vector	Restriction Sites 5', 3'	Selection
pRS411-P_{Nop1}-mCherry	N/A	N/A	pRS411	N/A	LEU2
pRS411-P_{Nop1}-NUP42-mCherry	Nup42	1-430	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}-nup42 ΔFG-mCherry	Nup42	364-430	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}-nup42 CTD-mCherry	Nup42	397-430	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}-nup42 minCTD-mCherry	Nup42	405-430	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}-nup42 ΔCTD-mCherry	Nup42	1-397	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}-nup42 F409D-mCherry	Nup42	1-430 (F409D)	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}- nup42 F414D mCherry	Nup42	1-430 (F414D)	pRS411	NdeI, SpeI	LEU2
pRS411-P_{Nop1}- nup42 F409D/F414D-mCherry	Nup42	1-430 (F409D/F414D)	pRS411	NdeI, SpeI	LEU2

REFERENCES

Adams, R.L., Terry, L.J., and Wente, S.R. (2014). Nucleoporin FG domains facilitate mRNP remodeling at the cytoplasmic face of the nuclear pore complex. Genetics *197*, 1213-1224.

Aditi, Folkmann, A.W., and Wente, S.R. (2015). Cytoplasmic hGle1A regulates stress granules by modulation of translation. Mol Biol Cell *26*, 1476-1490.

Aibara, S., Katahira, J., Valkov, E., and Stewart, M. (2015). The principal mRNA nuclear export factor NXF1:NXT1 forms a symmetric binding platform that facilitates export of retroviral CTE-RNA. Nucleic Acids Res *43*, 1883-1893.

Alcazar-Roman, A.R., Bolger, T.A., and Wente, S.R. (2010). Control of mRNA export and translation termination by inositol hexakisphosphate requires specific interaction with Gle1. J Biol Chem *285*, 16683-16692.

Alcazar-Roman, A.R., Tran, E.J., Guo, S., and Wente, S.R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. Nat Cell Biol *8*, 711-716.

Bolger, T.A., and Wente, S.R. (2011). Gle1 is a multifunctional DEAD-box protein regulator that modulates Ded1 in translation initiation. J Biol Chem *286*, 39750-39759.

Collins, R., Karlberg, T., Lehtio, L., Schutz, P., van den Berg, S., Dahlgren, L.G., Hammarstrom, M., Weigelt, J., and Schuler, H. (2009). The DEXD/H-box RNA helicase DDX19 is regulated by an {alpha}-helical switch. J Biol Chem *284*, 10296-10300.

Dossani, Z.Y., Weirich, C.S., Erzberger, J.P., Berger, J.M., and Weis, K. (2009). Structure of the C-terminus of the mRNA export factor Dbp5 reveals the interaction surface for the ATPase activator Gle1. Proc Natl Acad Sci U S A *106*, 16251-16256.

Folkmann, A.W., Collier, S.E., Zhan, X., Aditi, Ohi, M.D., and Wente, S.R. (2013). Gle1 functions during mRNA export in an oligomeric complex that is altered in human disease. Cell *155*, 582-593.

Folkmann, A.W., Noble, K.N., Cole, C.N., and Wente, S.R. (2011). Dbp5, Gle1-IP6 and Nup159: a working model for mRNP export. Nucleus *2*, 540-548.

Freibaum, B.D., Lu, Y., Lopez-Gonzalez, R., Kim, N.C., Almeida, S., Lee, K.H., Badders, N., Valentine, M., Miller, B.L., Wong, P.C.*, et al.* (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature *525*, 129-133.

Gasset-Rosa, F., Chillon-Marinas, C., Goginashvili, A., Atwal, R.S., Artates, J.W., Tabet, R., Wheeler, V.C., Bang, A.G., Cleveland, D.W., and Lagier-Tourenne, C. (2017). Polyglutamine-Expanded Huntingtin Exacerbates Age-Related Disruption of Nuclear Integrity and Nucleocytoplasmic Transport. Neuron *94*, 48-57 e44.

Grima, J.C., Daigle, J.G., Arbez, N., Cunningham, K.C., Zhang, K., Ochaba, J., Geater, C., Morozko, E., Stocksdale, J., Glatzer, J.C.*, et al.* (2017). Mutant Huntingtin Disrupts the Nuclear Pore Complex. Neuron *94*, 93-107 e106.

Hodge, C.A., Colot, H.V., Stafford, P., and Cole, C.N. (1999). Rat8p/Dbp5p is a shuttling transport factor that interacts with Rat7p/Nup159p and Gle1p and suppresses the mRNA export defect of xpo1-1 cells. EMBO J *18*, 5778-5788.

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Abou Merhi, R., and Maiorano, D. (2017). An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. EMBO J.

Hoelz, A., Debler, E.W., and Blobel, G. (2011). The structure of the nuclear pore complex. Annu Rev Biochem *80*, 613-643.

Jovicic, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S.B., Paul, J.W., 3rd, Sun, S., Herdy, J.R., Bieri, G.*, et al.* (2015). Modifiers of C9orf72 dipeptide repeat

toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci *18*, 1226-1229.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kaneb, H.M., Folkmann, A.W., Belzil, V.V., Jao, L.E., Leblond, C.S., Girard, S.L., Daoud, H., Noreau, A., Rochefort, D., Hince, P.*, et al.* (2015). Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. Hum Mol Genet *24*, 1363-1373.

Kendirgi, F., Barry, D.M., Griffis, E.R., Powers, M.A., and Wente, S.R. (2003). An essential role for hGle1 nucleocytoplasmic shuttling in mRNA export. J Cell Biol *160*, 1029-1040.

Kendirgi, F., Rexer, D.J., Alcazar-Roman, A.R., Onishko, H.M., and Wente, S.R. (2005). Interaction between the shuttling mRNA export factor Gle1 and the nucleoporin hCG1: a conserved mechanism in the export of Hsp70 mRNA. Mol Biol Cell *16*, 4304-4315.

Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A.L., Wan, W., Bui, K.H., Hagen, W.J., Briggs, J.A., Glavy, J.S.*, et al.* (2016). Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science *352*, 363-365.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N.*, et al.* (2016). Architecture of the symmetric core of the nuclear pore. Science *352*, aaf1015.

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

Mikhailova, T., Shuvalova, E., Ivanov, A., Susorov, D., Shuvalov, A., Kolosov, P.M., and Alkalaeva, E. (2017). RNA helicase DDX19 stabilizes ribosomal elongation and termination complexes. Nucleic Acids Res *45*, 1307-1318.

Montpetit, B., Seeliger, M.A., and Weis, K. (2012). Analysis of DEAD-box proteins in mRNA export. Methods Enzymol *511*, 239-254.

Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature *472*, 238-242.

Napetschnig, J., Kassube, S.A., Debler, E.W., Wong, R.W., Blobel, G., and Hoelz, A. (2009). Structural and functional analysis of the interaction between the nucleoporin Nup214 and the DEAD-box helicase Ddx19. Proc Natl Acad Sci U S A *106*, 3089-3094.

Neumann, B., Wu, H., Hackmann, A., and Krebber, H. (2016). Nuclear Export of Pre-Ribosomal Subunits Requires Dbp5, but Not as an RNA-Helicase as for mRNA Export. PLoS One *11*, e0149571.

Niesen, F.H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc *2*, 2212-2221.

Noble, K.N., Tran, E.J., Alcazar-Roman, A.R., Hodge, C.A., Cole, C.N., and Wente, S.R. (2011). The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1. Genes Dev *25*, 1065-1077.

Nousiainen, H.O., Kestila, M., Pakkasjarvi, N., Honkala, H., Kuure, S., Tallila, J., Vuopala, K., Ignatius, J., Herva, R., and Peltonen, L. (2008). Mutations in mRNA export mediator GLE1 result in a fetal motoneuron disease. Nat Genet *40*, 155-157.

Rayala, H.J., Kendirgi, F., Barry, D.M., Majerus, P.W., and Wente, S.R. (2004). The mRNA export factor human Gle1 interacts with the nuclear pore complex protein Nup155. Mol Cell Proteomics *3*, 145-155.

Rollenhagen, C., Hodge, C.A., and Cole, C.N. (2004). The nuclear pore complex and the DEAD box protein Rat8p/Dbp5p have nonessential features which appear to facilitate mRNA export following heat shock. Mol Cell Biol *24*, 4869-4879.

Saavedra, C.A., Hammell, C.M., Heath, C.V., and Cole, C.N. (1997). Yeast heat shock mRNAs are exported through a distinct pathway defined by Rip1p. Genes Dev *11*, 2845-2856.

Schmitt, C., von Kobbe, C., Bachi, A., Pante, N., Rodrigues, J.P., Boscheron, C., Rigaut, G., Wilm, M., Seraphin, B., Carmo-Fonseca, M.*, et al.* (1999). Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. EMBO J *18*, 4332-4347.

Shang, J., Yamashita, T., Nakano, Y., Morihara, R., Li, X., Feng, T., Liu, X., Huang, Y., Fukui, Y., Hishikawa, N.*, et al.* (2017). Aberrant distributions of nuclear pore complex proteins in ALS mice and ALS patients. Neuroscience *350*, 158-168.

Shears, S.B. (2001). Assessing the omnipotence of inositol hexakisphosphate. Cell Signal *13*, 151-158.

Shi, K.Y., Mori, E., Nizami, Z.F., Lin, Y., Kato, M., Xiang, S., Wu, L.C., Ding, M., Yu, Y., Gall, J.G.*, et al.* (2017). Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. Proc Natl Acad Sci U S A *114*, E1111-E1117.

Skubak, P., and Pannu, N.S. (2013). Automatic protein structure solution from weak X-ray data. Nat Commun *4*, 2777.

Stewart, M. (2010). Nuclear export of mRNA. Trends Biochem Sci *35*, 609-617.

Strong, M., Sawaya, M.R., Wang, S., Phillips, M., Cascio, D., and Eisenberg, D. (2006). Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. Proc Natl Acad Sci U S A *103*, 8060-8065.

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

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

Weirich, C.S., Erzberger, J.P., Flick, J.S., Berger, J.M., Thorner, J., and Weis, K. (2006). Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. Nat Cell Biol *8*, 668-676.

Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P., Daley, E.L., Miller, S.J., Cunningham, K.M., Vidensky, S.*, et al.* (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature *525*, 56-61.