Structure of the yeast Nup84-Nup133 complex details flexibility and reveals universal conservation of the membrane anchoring ALPS motif

Sarah A. Nordeen[1], Daniel L. Turman[1] & Thomas U. Schwartz[1*]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States
*corresponding author

Abstract

The hallmark of the eukaryotic cell is the complex endomembrane system that compartmentalizes cellular functions. Transport into and out of the nucleus, occurs through the Nuclear Pore Complex (NPC). The heptameric Nup84 or Y complex is an essential scaffolding component of the NPC. Here we report two nanobody-bound structures: the full-length Nup84-Nup133 C-terminal domain complex and the Nup133 N-terminal domain, both from *S. cerevisiae.* Together with previously published structures, this work enables the structural description of the entire 575 kDa Y complex, from one species. The structure of Nup84-Nup133$_{CTD}$ details the high flexibility of this dimeric unit of the Y complex. Further, the Nup133$_{NTD}$ contains a structurally conserved amphipathic lipid packing sensor (ALPS) motif, confirmed by liposome interaction studies. The new structures reveal important details about the function of the Y complex that affect our understanding of NPC structure and assembly.

In the eukaryotic cell, the genetic material is sequestered in the nucleus, separated from the cytoplasm by the double-membraned nuclear envelope (NE). Soluble transport across the NE occurs solely through nuclear pore complexes (NPCs), massive, proteinaceous assemblies that sit in circular openings where inner and outer nuclear membranes fuse, effectively perforating the NE. Regulated transport through NPCs plays a key role in many essential cellular processes, such as regulated gene expression, cell division, and ribosome assembly[1]. The NPC is a cylindrical assembly subdivided into three rings, based on cryo-electron tomographic analysis[2–4]. The cytoplasmic and nucleoplasmic rings each sit on the surface of the NE, sandwiching the inner ring in between. Despite the NPC's large mass (~60MDa in yeast), only about 30 different nucleoporins (nups) make up the 8-fold symmetric assembly of subcomplexes that together form the NPC[5–8]. These Nups can be classified into three main categories, (i) scaffold nups that make up the stable structure of the NPC, (ii) peripheral Nups that are biased to one face of the NPC, and (iii) phenylalanine-glycine (FG) nups that form the permeability barrier of the NPC.

One main scaffold nup subcomplex is the Nup84 or Y complex[5]. The Y complex is an essential component for NPC assembly and integrity, and it makes up approximately 20% of the mass of the NPC[9,10]. In *Saccharomyces cerevisiae*, the heptameric 575kDa Y complex comprises Nup84, Nup85, Nup120, Nup133, Nup145C, Sec13, and Seh1[11]. It can generally be described as having two short arms and a long stalk, connected in a central hub. Nup120 and a the Seh1-Nup85 complex form one short arm each, while Nup145C-Sec13, Nup84, and Nup133 sequentially emanate from the hub to form the long stalk (Fig. 1a). Great efforts to study the Y complex via X-ray crystallography yielded structures of individual Y complex Nups or overlapping subassemblies across multiple species to near completion[12–26]. However, there remains one elusive element of the Y complex never solved by X-ray crystallography from any species: the full-length structure of Nup84. While complex structures of an N-terminal fragment with Nup145C-Sec13, and a C-terminal fragment of Nup107 (the human homolog) with Nup133

are known, the missing connection between the two yields a significant problem in modeling a composite Y complex structure. Several studies suggested flexibility within the middle segment of Nup84/Nup107, but the degree of this and whether the modeling of the missing element is correct are limiting uncertainties[25–28]. Nup84 has an ancestral coatamer element (ACE1) fold, as do Nup85, Nup145C, and Nic96 of the inner ring complex[6]. The ACE1 fold is characterized by three elements, the 'crown', the 'trunk', and the 'tail', formed by a specific helical arrangement that forms a fold-back, U-shaped structure[13]. Flexibility has been observed at the interfaces between the three core elements[29]. However, for Nup84, the similarity with other scaffold nups suggests that it should not contain a significantly disordered region. Therefore, we reasoned that knowing the structure of full-length Nup84 would be critical to better model the entire Y complex. The structure of the assembled NPC, which remains a major goal in structural biology, can only be revealed by a hybrid approach, combining high-resolution component structures with lower-resolution cryo-ET maps capturing the NPC *in situ*[2,4,30–32].

In addition to the lack of knowledge about Nup84, we have little structural information about yeast Nup133 aside from the C-terminal 'heel' domain, which only has modest similarity to its human homolog[22,23]. Also, completing the structure of the Y complex from a single species should help understand the differences in the NPC structure across species[32–34]. As a member of the cytoplasmic and nucleoplasmic rings in the NPC, the Y complex coats the peripheral inner and outer nuclear membrane (INM, ONM, respectively) bordering the circular openings in the NE[35]. It remains unclear how the yeast Y complex interacts with the INM and ONM and how it links to the inner NPC ring. The ArfGAP1 lipid packing sensor (ALPS) motif in the human homolog of Nup133 is thought to anchor the Y complex to the INM and ONM[21,36,37]. The hsNup133 ALPS motif is critical for interphase assembly in metazoa, a process which is thought to be similar to NPC assembly in organisms with closed mitosis like *S. cerevisiae*[37,38]. Whether or not the ALPS motif is conserved in *S. cerevisiae* has been debated in the literature, due to the lack of high-resolution structural or functional studies on yeast Nup133[21,24,36].

Here we report the structures of the Nup84-Nup133 C-terminal α-helical domain and Nup133 N-terminal β-propeller *from S. cerevisiae*. The structures were obtained using nanobodies, single domain antibodies derived from alpacas[39], as crystallization chaperones. This completes the entire structure of the Y complex from *S. cerevisiae* and allowed us to create a new composite model of the Y complex assembly, the first complete composite model from a single species. Additionally, we show Nup133 has a functional ALPS motif through liposome interaction studies. This new model of the Y complex delineates key hinge points and possible motion ranges in the Y complex stalk and establishes the position of Nup133 with its ALPS motif placed adjacent to the membrane in the NPC assembly.

Results

Structure of the Nup84-Nup133$_{CTD}$-VHH-SAN8 complex

We solved the structure of full-length yeast Nup84-Nup133 C-terminal domain (Nup133$_{CTD}$, residues 521-1157) bound by a nanobody (VHH-SAN8) by single-wavelength anomalous dispersion (SAD) to 6.4 Å resolution using an Anderson-Evans polyoxotungstate derivative (Table 1, Fig. 1b). We observed clear elongated, helical density in the initial, solvent-flattened experimental map. After multiple rounds of refinement and successively placing helices into the density, we arrived at a map with clear enough density to employ both real-space docking and molecular replacement to place fragments of the structure that were previously known[17,22,23] (Supplementary Fig. 1). Following placement of roughly half of Nup84 and the C-terminal heel domain of Nup133, we were able to manually build the remaining helices. Due to the limited resolution of the data, we did not build side chains for the model. However, we are confident in assigning the sequence for both, Nup84 and Nup133 (see Material and Methods for details).

Placement of the nanobody VHH-SAN8 was unambiguous as we observed large difference density near the Nup84 crown element (Supplementary Fig. 2). However, the electron

density around the three complementarity determining regions (CDRs) which presumably make up the majority of the interface with Nup84 are not clear enough for chain tracing. Therefore, we only rigid body positioned a nanobody model without CDRs in our structure. Additionally, we solved the structure of Nup84-Nup133$_{CTD}$ bound to both VHH-SAN8 and a second nanobody, VHH-SAN9, at 7.4 Å resolution (Table 1 and Supplementary Fig. 3). The doubly bound Nup84-Nup133$_{CTD}$-VHH8/9 structure crystallized in the same crystal form as the singly bound Nup84-Nup133$_{CTD}$-VHH8 structure. VHH-SAN9 binds on the opposite face as VHH-SAN8 at nearly the same vertical position along Nup84. Again, this resolution precluded the building of the CDR loops for both nanobodies, and a nanobody model without CDRs was rigid body positioned into the map at the site of clear difference density for VHH-SAN9 (Supplementary Fig. 2).

Nup84 forms a continuous α-helical stack domain which, expectedly, belongs to the ACE1 architecture. Its 28 α-helices form a rectangular block with dimensions of ~124 x 37 x 46 Å, with α1-7 zig-zagging in one direction and α8-28 folding over and zig-zagging the opposite direction to the C-terminal end of the molecule (Fig. 1). The C-terminal helices interact with Nup133$_{CTD}$, which itself forms an elongated, highly arched stack of 26 α-helices (Fig. 1). The Nup84-Nup133 binding interface is compact, with ~770 Å^2 buried surface area, comprising α26-28 of Nup84 and α15-17 of Nup133. Typically, scaffold Nups are poorly conserved on the sequence level, despite a high degree of structural conservation[29]. However, the interface between the human homologs of Nup107-Nup133 showed a high degree of sequence conservation[20]. To see if this held true in our structure, we calculated conservation via ConSurf[40] and mapped the degree of conservation onto the surface of the molecule (Supplementary Fig. 4). To generate the surface, we added side chains in ChimeraX using the Dunbrack rotamer library[41]. From this analysis, we can observe an enrichment of highly conserved residues at the interface between Nup84 and Nup133, similar to that observed in the human homologs[20] (Supplementary Fig. 4).

Full length Nup84 has an ACE1 fold

With the full-length structure of Nup84, we can unambiguously detect the complete topology of the ACE1 fold. We assign the crown to consists of α-helices 4-10, the trunk to α-helices 1-3 paired with α-helices 11-18, and the tail to α-helices 19-28 (Fig. 1). To find potentially undetected ACE1 folds in *S. cerevisiae,* we analyzed the structural homology of Nup84 by Backphyre[42] across the genome and had relatively few hits (Supplementary Table 1). As expected, these hits included Nup145C and Nup85 from the Y complex, Nic96 from the inner ring complex, and Sec31 and Sec16 from COPII vesicle coats and assembly. However, Backphyre also identified Sea4, a component of the SEA complex, which also includes Sec13 and Seh1[43]. The SEA complex is thought to associate with the vacuolar membranes and play a role in autophagy intracellular trafficking[44]. While no high-resolution structural information is available for Sea4, it is predicted to have an N-terminal β-propeller domain, followed by an α-helical stack[43]. The Backphyre analysis strongly suggests that Sea4 is another member of the ACE1 family.

When superimposing each ACE1 structure, one can easily detect their structural similarity (Fig. 2). Most of the trunk helices run perpendicular to the long axis of the protein, while the ‘crown’ helices tip upwards. The typically longer ‘tail’ helices lean downwards, away from the ‘trunk’. Comparing the ‘tail’ modules across Nups, they are visibly rotated to different degrees relative to each ‘trunk’, highlighting the flexibility at the ‘trunk-tail’ interface. While Sec31 and Sec16 both do not have a ‘tail’, they follow the same helical topology in the ‘trunk’ and ‘crown’ as the nups.

Interestingly, VHH-SAN8 binds at the interface between the crown and trunk elements, thereby potentially rigidifying this portion of the protein (Fig. 1b). We hypothesize that this interaction may stabilize a significant crystal packing interaction between Nup84 and a symmetry related copy of Nup133 (Supplementary Fig. 5). In support of this hypothesis, previous attempts at crystallizing this complex without the nanobody yielded poorer diffracting

crystals, albeit in the same space group and with identical unit cell dimensions. With the solved structure, we could now phase these nanobody-free Nup84-Nup133$_{CTD}$ crystals. We observe exceptionally poor density for the Nup84 crown, supporting our hypothesis of the nanobody stabilizing a conformationally dynamic area within Nup84 (data not shown).

For future *in vivo* studies, we tested by size-exclusion chromatography (SEC) whether the nanobodies VHH-SAN8/9 interfere with Nup145C binding, Nup84's anchor to the Y-complex. Analysis of Nup145C-Sec13-Nup84 incubated with VHH-SAN8 and VHH-SAN9 established that both nanobodies co-eluted with Nup145C-Sec13-Nup84 upon SEC (Supplementary Fig. 6). This suggests that the nanobodies do not interfere with Nup145C binding.

scNup133$_{CTD}$ is conformationally distinct from its human homolog

Nup133$_{CTD}$ consists of 3 modules: the 'blade' (α1- α9), the 'arch' (α10-17), and a 'heel' (α18-26) (Fig. 1). The first module is formed with a block of elongated helices that form a wide 'blade'. The arch helices α10-17 are much shorter and zig-zag back and forth, curving away from the 'blade'. The C-terminal 'heel' domain is irregular and compact, with core helices α20-24 surrounded by α18, α19, α25, and α26. The human homolog retains much of the same helical fold[20,22]. The largest difference between human and yeast Nup133 is in the C-terminal 'heel' domain (Fig. 3a). The hsNup133 'heel' lacks equivalent helices to α19, α24 and α25, making it smaller than the scNup133 'heel', and having a different helical topology[23]. When we superpose scNup133 and hsNup133 on the α-helices that interact with Nup84 (α15-17), we can observe a large difference between the two 'heel' domains (Fig. 3a). However, the helical topology at the scNup84/hsNup107 binding interface overlays reasonably well, consistent with the higher degree of sequence conservation in this part of the protein (RMSD 3.52 Å over 47 Cα positions) (Supplementary Fig. 4).

The 'blade' module of Nup133 is quite similar between yeast and human, only differing in two additional α-helices at the N terminus of the human homolog. However, the heel position

188 relative to the arch module varies across the two structures, with the human blade module
189 swung 'out' ~90° relative to the yeast blade. By superposing the two blade modules, this motion
190 becomes more apparent (Fig. 3b). This creates a swing of ~41 Å at the C terminus of the
191 protein and rotates both the arch and heel domains the same 90° rotation. This rotational and
192 translational motion, in addition to the flexible linker between $Nup133_{CTD}$ and $Nup133_{NTD}$ make it
193 the most flexible nup in the Y complex.

194 ***Structure of the Nup133 N-terminal β-propeller bound by VHH-SAN4***

195 We solved the structure of the yeast Nup133 N-terminal domain ($Nup133_{NTD}$, residues
196 55-481) at 2.1 Å resolution by molecular replacement, using the *Vanderwaltozyma polyspora*
197 $Nup133_{NTD}$ as a template[24] (Table 1, Fig. 4). $Nup133_{NTD}$ is a β-propeller consisting of seven β-
198 sheets (blades) that are radially arranged around a solvent-accessible core (Fig. 4). Blades β1-6
199 each contain 4 β-strands that start near the core and trace outwards. Namely, strand 'A' is the
200 innermost strand, where strand 'D' is the outermost. Blade β7 is the exception, forming a 5-
201 stranded blade made up of both, the N and C terminus, of the protein. In blade β7, the
202 outermost strands 'D' and 'E' are from the N-terminal part of the β-propeller and strands 'A-C'
203 are from the C-terminal portion. This "Velcro" closure is common in many β-propeller domains
204 and is thought to stabilize the closed fold[45]. In comparison to the $Nup133_{NTD}$ structures from
205 other species, the domain is structurally highly similar to *V. polyspora* (RMSD 1.36 Å over 381
206 Cα positions) and quite similar to human (RMSD 3.06 Å over 377 Cα positions) (Fig. 5a). All
207 three β-propellers have seven blades with two significant elaborations, an α-helical hairpin
208 between blades 4 and 5 (α2-3) and a 20-residue, mostly disordered loop between strand 3D
209 and strand 4A (DA_{34} loop) (Fig. 4). Additionally, this structure has an elongated loop between
210 stands 'A' and 'B' in blade β1 that is ordered due to crystal packing contacts. The core structure
211 of $Nup133_{NTD}$ is conserved across the three species.

The nanobody, VHH-SAN4, binds along strand 5D, via its large CDR3 loop (Fig. 4). The opposite end of the nanobody facilitated packing with three symmetry related copies of $Nup133_{NTD}$, critical for lateral assembly within the crystal. Additionally, we solved the structure of $Nup133_{NTD}$ bound by a second nanobody, VHH-SAN5, to 2.8Å resolution (Table 1 and Supplementary Fig. 7). VHH-SAN5 binds an adjacent epitope to VHH-SAN4, also binding along strand 5D using primarily its CDR3. There are additional contacts between CDR2 and the α-helical hairpin α2-3. The N-terminal β-strand of VHH-SAN5 also makes extended contact with strand 'D' of blade β6 on an adjacent copy of $Nup133_{NTD}$, ultimately forming a hexameric packing assembly in the core of the asymmetric unit (Supplementary Fig. 7). Two such $Nup133_{NTD}$-VHH-SAN5 heterohexamers stack, creating the striking dodecameric asymmetric unit (Supplementary Fig. 7). Despite the difference in packing between the two structures, the six copies of $Nup133_{NTD}$ in the VHH-SAN5 bound structure are highly similar both to each other (RMSD 0.17-0.22 Å) and to the VHH-SAN4 bound structure (RMSD 0.68-0.81 Å). The majority of the differences come in conformations of multiple, extended loops that are generally more ordered in VHH-SAN4 bound $Nup133_{NTD}$, presumably because of better crystal packing. Specifically, the loops containing residues 72-86, 143-160, and 404-412 are disordered in the majority of VHH-SAN5 bound $Nup133_{NTD}$ copies along with the α-helical hairpin α2-3 (residues 176-190) that are visible in the VHH-SAN4 bound map. On the other hand, loop residues 430-440 are better ordered in the VHH-SAN5 bound structure.

Yeast Nup133 has an ALPS motif

After observing the same disordered DA_{34} loop in our structure that is present in the human $Nup133_{NTD}$, we wondered if the yeast Nup133 also has a functional ArfGAP1 lipid packing sensing (ALPS) motif. A computationally determined ALPS motif has been identified in $scNup133_{NTD}$ via homology modeling[24]. However, the hydrophobic moment of this loop is weaker than the human DA_{34} loop due in part to the presence of an asparagine and lysine within the hydrophobic half of the amphipathic helix (Fig. 5a). The increase in hydrophilicity in the

yeast Nup133$_{NTD}$ DA$_{34}$ loop led some to conclude that this feature may not be conserved between metazoa and yeast[36]. To directly test whether yeast Nup133$_{NTD}$ has a functional ALPS motif, we performed liposome floatation assays. This assay determines whether a protein interacts with liposomes through the floatation of a liposome-protein mixture to the top of an iso-osmotic gradient during ultra-centrifugation[46]. We observed that WT-Nup133$_{NTD}$ 'floats' with liposomes comprising yeast polar lipids, but pellets without liposomes (Fig. 5b).

We also wondered if we could visualize any remodeling of liposomes by negative stain electron microscopy, as other yeast Nups (Nup1, Nup60, Nup53) with ALPS motifs have done to various extents[47,48]. Indeed, WT Nup133$_{NTD}$ produces small 'fringe' like protrusions on the surface of liposomes (Fig. 5c). This is in contrast to human Nup133$_{NTD}$, which did not alter liposomes[47]. However, this difference could be due to the varying liposome composition and size across these experiments. With an ALPS mutant where we replaced the DA$_{34}$ loop with an isosteric linker (GGGGSGGGGS) (Nup133$_{NTD}$ΔALPS), these protrusions did not occur (Fig. 5c). The results of these disparate assays demonstrate that the yeast Nup133$_{NTD}$ has an ALPS motif in its DA$_{34}$ loop, a conserved motif with human Nup133, that can bind to and modify curved biological membranes.

A complete composite model of the yeast Y complex

With our additional structures in hand, we constructed a new composite model of the yeast Y complex, the first complete Y complex of a single species. The assembly of this new model was straightforward, as both the structure we describe here and the previously published 'hub' structure overlap significantly within Nup84, and they superpose well (RMSD 2.1 Å over 454 Cα positions). While the structures of both, the Nup84-Nup133$_{CTD}$ element and the Y complex hub, are of modest resolution, many fragments of the assembly are known at much higher resolution. In comparing our new model to our previous homology-based model[25], we observe a number of conformational differences (Fig. 6). Both models have the same dimensions overall, but superpositioning the two models using Nup145C, the most rigid part of

the Y complex, reveals hinge points in both the short arms and the stalk of the Y. In the short arm of Nup85-Seh1, we notice significant motion around the trunk-to-tail transition of Nup85. Additionally, Nup120 is bent more outward from the hub in our previous model. The stalk has multiple points of flexibility. The Nup84 tail is positioned about ~42Å away from its previous position, with the majority of the movement occurring around the hinge at the trunk-tail interface. This motion causes the Nup133 heel domain to shift ~30Å. The Nup133 arch between the heel and the blade moves the blade domain an additional ~30Å, with the motion from the long stalk totaling 73Å. Measuring motion between the $Nup133_{NTD}$ and $Nup133_{CTD}$ seems pointless, as the β-propeller is tethered by a flexible, unstructured element to the α-helical stack[24]. While both models provide snapshots of Y complexes dictated by the packing environment of the underlying crystal structures, electron microscopy studies of the Y complex support the notion of defined motions within the complex[27,28]. Overall, these two composite models capture different conformations, highlighting the range of motion and the hinge points of the Y complex.

The composite Y complex model in the NPC assembly

While our new composite model represents one snapshot of the Y complex conformation, we tested how well it could be positioned into the recently published cryo-electron tomography (cryo-ET) reconstruction from *S. cerevisiae*[4]. The cryo-ET map is resolved to ~25 Å, so placement of nups is only tentative without secondary structure definition. Our model of the Y complex fit quite well into the density for both the cytoplasmic and nucleoplasmic rings of the NPC (Fig. 7). The structure of Nup84-$Nup133_{CTD}$ reported here is flatter and straighter than previous models suggested[25,32] (Fig. 6), resulting in a better fit for our new composite model. The $Nup133_{NTD}$ is only approximately positioned relative to the $Nup133_{CTD}$, as the connection between the two is highly flexible. However, $Nup133_{NTD}$ is positioned with its ALPS motif adjacent to the NE. The fit for the nucleoplasmic ring was slightly better, suggesting some conformational differences between the Y complexes on the cytoplasmic and nucleoplasmic faces of the NPC (Fig. 7b, see Materials and Methods). The difference between the cytoplasmic

and nucleoplasmic Y complex conformation is more striking in docking attempts using our previous composite model (Fig. 7). While the fitting correlation score and positioning for the cytoplasmic ring was comparable between both models, the previous model fit much more poorly to the nucleoplasmic ring (see Materials and Methods). The overall shape and dimensions of this new composite model are more consistent with this cryo-ET reconstruction.

Discussion

Here we describe two new structures, Nup84-Nup133$_{CTD}$ and Nup133$_{NTD}$, that complete the structural inventory for the *S. cerevisiae* Y complex. This new model pushes forward our molecular understanding of the greater NPC structure, which will ultimately give us a better molecular understanding of the many functions of the NPC. Complementary to our 'bottom-up' structural approach studying nup subassemblies, studies on the massive assembly of the NPC from a multitude of species by cryo-ET yielded density maps at a range of resolution. However, no cryo-ET map has been solved to high enough resolution to discern secondary structures, which would ultimately allow for unambiguous placement of nups[49]. In yeast, there are two recent cryo-ET reconstructions. One analyzed NPCs after detergent extraction from the NE, with a local resolution of about ~7 nm in the cytoplasmic and nucleoplasmic rings, precluding any reasonable attempt at docking the Y complex[32]. A second study describes a cryo-ET reconstruction from intact yeast nuclei to ~2.5 nm resolution[4]. Both of these *in situ* docking studies have docked single nups or small nup assemblies, as the previous Y complex models did not fit well as a rigid body. However, our new composite model fit quite well into the higher resolution cryo-ET map (Fig. 7)[4]. There still remains room for improvement in both the resolution of the cryo-ET map to secondary structure resolution and the accuracy of the model for the Y complex. Now that we have a complete Y complex and multiple structures for most of the Nups, we can delineate the flexible elements within the structure. We can also now see the conformational differences in the Y complex on different faces of the NPC (Fig. 7). These

differences were also noted in a cryo-ET study on *Xenopus laevis* NPCs, suggesting this difference as a conserved feature of the assembly[3]. Future studies can employ techniques such as molecular dynamics, moving the molecule at known hinge points to fit into the map, rather than individual nups that look highly similar at >10 Å resolution, to ultimately get a high-resolution assembly model. Overall, the structures presented here provide additional elements for better constructing composite NPC structures combining X-ray crystallographic and electron microscopy methods.

The observation of the great flexibility of the Y complex has led many to speculate its importance in NPC biology. One key contributor to this flexibility is the ACE1 fold in multiple Y complex nups. The structurally related Sec16 and Sec31 of the COPII vesicle coat both need to coat dynamically shaped, highly curved membranes through a vesicle's life cycle[50]. While the Y complex arrangement within the NPC differs from that of the COPII vesicle coat, the conservation of the ACE1 fold and its inherent flexibility could be critical for NPC assembly and integrity. NPC heterogeneity in both size and constituents has been suggested by a number of studies, and therefore it is likely the core scaffold of the NPC would need to accommodate these variances [33,51–54]. The NPC and Y complex may also have to be flexible in order to adapt to a variety of cargoes. The NPC can facilitate the transport of surprisingly large cargoes, up to ~39 nm in diameter, such as intact Hepatitis B viral capsids[55,56]. Conceptually, there is no need for the nucleoplasmic and cytoplasmic rings to adopt to cargo traveling the central channel, since the ring opening is much larger than even the biggest cargoes. However, integral inner nuclear membrane proteins pass through the NPC presumably via peripheral channels adjacent to the NE[57], unless they are inserted post-translationally[58]. Some of these proteins, such as Heh1 and Heh2 in *S. cerevisiae,* have significantly large nucleoplasmic domains that would need to pass near to the scaffold of the NPC as they translocate[59,60]. As the main scaffolding component of both the cytoplasmic and nucleoplasmic rings, it is still an unsolved problem if and how the cyto-

and nucleoplasmic rings are involved in this process. It is easy to imagine that the observed Y complex flexibility plays a part in this translocation process.

Another potential for required flexibility could be in the assembly of new NPCs. The Y complex is one of the first complexes present during NPC assembly[61,62]. Knowing that Nup133 can anchor into the nuclear envelope with its ALPS motif, it is not unlikely that the Y complex ring of a newly assembling NPC would need to change shape as the inner and outer nuclear membranes transition through the extreme curvatures required for NE fusion during the building of a new NPC[37]. This would be especially critical in assembly into a pre-formed NE during interphase or a closed-mitotic system, such as in *S. cerevisiae*.

The conformation of the yeast Nup133 ALPS motif also raises many questions as to its relevance *in vivo*. Various deletions of Nup133, including both a complete deletion and deletions to portions of its N terminus result in decreased fitness in yeast and a 'clustering' phenotype of the NPC on the nuclear envelope[63–65]. The mechanism and effects on the NPC structure and assembly during clustering is still unknown. Deletion of only the DA_{34} loop has only been done in humans and was found to be essential for interphase assembly[37]. Careful dissection of the role of the Nup133 ALPS motif could answer questions on the conservation of NPC assembly across yeast and human. Confirming the presence of the ALPS motif on Nup133 puts the $Nup133_{NTD}$ in immediate proximity to the membrane. Again, this β-propeller domain has been separately docked *in situ*, but its orientation was unclear[30,66]. With knowing the relative position of Nup133, one can begin to examine the head-to-tail interaction of neighboring Y complexes in greater detail.

Materials and Methods

Construct generation

$Nup133_{NTD}$ (aa55-481), $Nup133_{CTD}$ (aa521-1157), and Nup84 were cloned from *S. cerevisiae*. $Nup133_{NTD}$ was N-terminally fused with a human rhinovirus 3C (3C) protease

cleavable 6xHis tag and cloned into an ampicillin resistant, T7-promoter-based bacterial expression vector. We cloned the heterodimeric complex of Nup84 and $Nup133_{CTD}$ into a T7-promoter-based bicistronic bacterial expression vector with ampicillin resistance. Nup84 was N-terminally fused with a 3C cleavable 10xHis-8xArg-SUMO tag. $Nup133_{CTD}$ was C-terminally fused with an 8xArg tag. The $Nup133_{NTD}$ ALPS mutant was generated by replacing residues 252-270 with GGGGSGGGS by inverse PCR. Upon VHH selection by phage display and ELISA, VHH-SAN4, 5, 8, and 9, were sub-cloned for expression. Each VHH sequence was N-terminally fused with a 14xHis bdSUMO tag[67] and cloned into a T7-promoter-based bacterial expression vector with ampicillin resistance. The Nup145C-Sec13fusion-Nup841-424 construct was previously described in [17].

Protein Expression and purification

The Nup84-$Nup133_{CTD}$ vector was transformed into *Escherichia coli* LOBSTR-RIL(DE3) (Kerafast)[68] cells and protein production was induced with 0.2 mM IPTG at 18 °C for 12-14 h. $Nup133_{NTD}$ was expressed separately under identical conditions. Cells were collected by centrifugation at 6,000g, resuspended in lysis buffer (50 mM potassium phosphate (pH 8.0), 500 mM NaCl, 30 mM imidazole, 3 mM β-mercaptoethanol (βME), 1 mM PMSF) and lysed using a high-pressure cell homogenizer (Microfluidics LM20). The lysate was cleared by centrifugation at 12,500g for 25 min. The soluble fraction was incubated with Ni Sepharose 6 Fast Flow beads (GE Healthcare) for 30 min on ice. After washing the beads with lysis buffer, the protein was eluted (250 mM imidazole, pH 8.0, 150 mM NaCl, 3 mM βME).

Nup84-$Nup133_{CTD}$ was dialyzed into 10 mM HEPES/NaOH pH 8.0, 100 mM NaCl, 0.1 mM EDTA, and 1 mM DTT and loaded directly onto a 5 ml HiTrap SP Sepharose fast flow column (GE Healthcare). Bound Nup84-Nup133 was eluted from the column using a 15 column volume (CV) gradient of 10 mM HEPES/NaOH pH 8.0, 1 M NaCl, 0.1 mM EDTA, and 1 mM DTT. The pooled eluate was loaded onto a Superdex 200 26/60 gel filtration column equilibrated

in 10 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT. The Nup84-Nup133 peak was pooled and concentrated.

$Nup133_{NTD}$ was dialyzed into 10 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT. $Nup133_{NTD}$ was then concentrated and loaded on a Superdex 75 16/60 gel filtration column equilibrated in the same buffer. The $Nup133_{NTD}$ peak was pooled and concentrated. The $Nup133_{NTD}$ ALPS mutant was prepared identically.

All VHH constructs were transformed, grown, harvested, and lysed as above. After lysis, the soluble fraction was incubated with Ni Sepharose 6 Fast Flow beads (GE Healthcare) for 30min on ice. The beads were then washed with lysis buffer and transferred to low imidazole buffer (50 mM potassium phosphate pH 8.0, 500 mM NaCl, 10 mM imidazole, 3 mM βME) along with 10 µg SEN-p protease and incubated 2 hrs at 4 °C. Then the flow through containing the cleaved VHHs was collected, along with a 2 CV wash with low imidazole buffer. Cut tags and uncut protein was eluted as above. Each VHH was concentrated and loaded onto a Superdex 75 16/60 gel filtration column equilibrated in 10 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT. Each VHH peak was pooled and concentrated.

VHH Library and M13 phage generation

Alpaca immunization and library generation was done as previously described[69]. The animal was immunized against recombinantly expressed full-length Y complex (Nup120-Nup85-Seh1-Nup145C-Sec13-Nup84-Nup133). The library was then grown to mid-log phase in 100ml SOC with 50 µg/ml ampicillin. Then, the culture was infected with 100µl 10^{14} PFU/ml VCSM13 helper phage. Following 2hr incubation at 37°C, the cells were harvested by centrifugation and re-suspended in 100ml 2YT, 0.1% glucose, 50µg/ml kanamycin and 50 µg/ml ampicillin. Cultures were incubated overnight at 30°C, then centrifuged for 20min at 7,700 × g, followed by

phage precipitation from the resulting supernatant with 1% PEG-6000, 500mM NaCl at 4°C, and re-suspended in PBS.

Selection of VHHs by Phage Display

VHHs were selected by panning against Nup84 and $Nup133_{NTD}$. 100 µg recombinant Nup84 and $Nup133_{NTD}$ was biotinylated by coupling Chromalink NHS-biotin reagent (Solulink) to primary amines for 90 min in 100 mM sodium phosphate pH 7.4, 150 mM NaCl. Unreacted biotin was removed using a Zeba desalting column (Thermo Fisher). Biotin incorporation was monitored using absorbance at 354 nM. 100 µl MyOne Streptavidin-T1 Dynabeads (Life Technologies) were blocked in 2% (w/v) bovine serum albumin (Sigma) in PBS for 2 h at 37 °C. 20 µg biotinlyated Nup84 and $Nup133_{NTD}$ in PBS was added to the blocked beads and incubated for 30 min at 25°C with agitation. The beads were then washed 3 times in PBS and 200 µl of 10^{14} PFU ml^{-1} M13 phage displaying the VHH library were added in 2% BSA in PBS for 1 hour at room temperature. The beads were then washed 15 times with PBS, 0.1% Tween-20 (PBST). Phage was eluted by the addition of *E. coli* ER2738 (NEB) for 15 min at 37°C, followed by elution with 200 mM glycine, pH 2.2, for 10 min at 25°C. The eluate was neutralized with 1M Tris/HCl pH 9.1, pooled with the *E. coli* culture, and plated onto 2YT agar plates supplemented with 2% glucose, 5 µg/ml tetracycline and 10 µg/ml ampicillin, and grown overnight at 37°C. A second round of panning was performed with the following modifications: 2 µg of biotinylated Nup84 and $Nup133_{NTD}$ were used as bait, and incubated with 2 µl 10^{14} PFU/ml M13 phage displaying the first-round VHH library for 15 min at 37°C, followed by 15 washes in PBST.

ELISA

Following two rounds of phage panning, 96 colonies were isolated in 96-well round-bottom plates and grown to mid-log phase at 37°C in 200 µl 2YT, 10 µg/ml ampicillin, 5 µg/ml

tetracycline, induced with 3 mM IPTG and grown overnight at 30°C. Plates were centrifuged at 12,000g for 10 min, and 100 µl of supernatant was mixed with an equal volume of 5% (w/v) nonfat dry milk in PBS. This mixture was added to an ELISA plate coated with 1 µg/ml Nup84-133 or $Nup133_{NTD}$. Following four washes with 1% Tween-20 in PBS, anti-llama-HRP antibody (Bethyl) was added at a 1:10,000 dilution in 5% (w/v) nonfat dry milk in PBS for 1 hour at 25°C. The plate was developed with fast kinetic TMB (Sigma) and quenched with 1M HCl. Absorbance at 450nm was determined in a plate reader (Spectramax; Molecular Devices).

Protein Crystallization

Initial crystals of Nup84-$Nup133_{CTD}$-VHH-SAN8 were obtained at 18 °C in 2 days as part of the Morpheus *HT* suite (Molecular Dimensions) in a 96-well sitting drop tray with a reservoir containing 12% (v/v) ethylene glycol, 6% (w/v) PEG 8,000, 0.1 M imidazole titrated with MES pH 6.5, 30 mM sodium nitrate, 30 mM sodium phosphate dibasic, 30 mM ammonium sulfate. Sitting drops of 2 µl protein at 8 mg/ml and 1 µl of precipitant (5-6% PEG 8,000, 30-32.5% ethylene glycol, 0.1M imidazole titrated with MES pH 6.5, 15mM sodium nitrate, 15mM sodium phosphate dibasic, 15mM ammonium sulfate) produced crystals in 3 days, which continued to grow over the following week. Nup84-$Nup133_{CTD}$+VHH-SAN8/9 crystallized under identical conditions. Crystals were harvested without additional cryo-protectant. Derivative crystals were obtained by applying 0.2ul of 0.1M $[TeW_6O_{24}]^{6-}$ (MiTeGen) to drops containing Nup84-$Nup133_{CTD}$-VHH-SAN8 crystals. Crystals were harvested after 1 hour and cryo-cooled in liquid nitrogen.

Initial crystals of $Nup133_{NTD}$-VHH-SAN4 were obtained at 18°C in 14 days as part of the Index*HT* suite (Hampton Research) in a 96-well sitting drop tray with a reservoir containing 25% PEG 3,350, 0.1M Bis-Tris/HCl pH 5.5, 0.2M ammonium sulfate. Hanging drops of 1 µl protein at 7mg/ml and 1 µl precipitant (23-24% PEG 3,350, 0.2M Ammonium sulfate, 0.1M Bis-Tris/HCl pH

5.5) produced diffraction-quality crystals of rod-shaped clusters in 12 days to 1 month. Crystals were transferred into a cryo-protectant solution containing the crystallization condition with 12% PEG 8,000 and cryo-cooled in liquid nitrogen.

Initial crystals of $Nup133_{NTD}$-VHH-SAN5 were obtained at 18°C in 1 day as part of the Protein Complex suite (Qiagen) in a 96-well sitting drop tray with a reservoir containing 15% PEG MME 2,000, 0.1M potassium chloride, 0.1M Tris/HCl pH 8. Hanging drops of 1 µl protein at 5 mg/ml and 1 µl precipitant (11% PEG MME 2,000, 0.2M potassium chloride, 0.1M Tris/HCl pH 8) produced diffraction-quality diamond-shaped crystals in 1-3 days. Crystals were transferred into a cryo-protectant solution containing the crystallization condition with 12% PEG 200 and cryo-cooled in liquid nitrogen.

Data Collection and Structure Determination

Data collection was performed at the Advanced Photon Source end station 24-IDC. All data processing steps were carried out with programs provided through SBgrid[70]. Data reduction was performed using HKL2000[71] and XDS[72]. Statistical parameters of data collection and refinement are given in Table 1. Structure figures were created in PyMOL (Schrödinger LLC).

For Nup84-Nup133-VHH-SAN8, the structure was solved with single anomalous scattering using a $[TeW_6O_{24}]^{6-}$ (MiTeGen) derivative dataset. One cluster site was found with SHELXC/D/E from the CCP4 suite[73]. *AutoSol* in the PHENIX suite was used to refine the site and generate an initial map[74]. The map was iteratively refined using *phenix.refine* after placement of poly-alanine helices in *Coot*[75]. Once the density had improved, known fragment of Nup133 heel (PDB: 3KFO) was placed via *Dock in Map* in *Phenix*. After one round of refinement, a known fragment of Nup84 (PDB: 3JRO) was placed in the map via molecular replacement. The remaining parts of the proteins were manually built, using the homologous human structure as a guide (PDB: 3I4R). The asymmetric unit contains one Nup84-Nup133-VHH-SAN8 complex. *Phenix.rosetta_refine*[76] was employed once all of the density for Nup84

and Nup133 was occupied. Only rigid body refinement was done on the previously solved parts of the molecule, along with applying secondary structure restraints throughout the refinement process. Near the end of refinement, TLS parameters were used. Since the complementarity determining region (CDR) loops were unclear in the map, a generic nanobody with its CDR loops removed was used as a rigid body (PDB: 1BZQ)[77]. The model was built without side chains, as no clear density for them can be observed in the electron density map. The sequence has been applied to the model based on the docked, previously solved fragments and alignment with the homologous human structure of Nup133 (PDB: 3I4R).

The structure of $Nup133_{NTD}$-VHH-SAN5 was solved by molecular replacement (MR) using *Phaser-MR* in PHENIX[74]. A two-part MR solution was obtained by sequentially searching with models of $Nup133_{NTD}$ and VHH-SAN5. For $Nup133_{NTD}$, we used the structure of *Vanderwaltozyma polyspora* $Nup133_{NTD}$ (PDB: 4Q9T, 45% sequence identity). For VHH-SAN5, we used a nanobody structure with its CDR loops removed (PDB: 1BZQ). The asymmetric unit contains six copies of $Nup133_{NTD}$+VHH-SAN5. The structure of $Nup133_{NTD}$-VHH-SAN4 was solved by MR using *Phaser-MR* in PHENIX, just as $Nup133_{NTD}$-VHH-SAN5. However, we used our previously solved $Nup133_{NTD}$ as the search model. The asymmetric unit contains one copy of $Nup133_{NTD}$+VHH-SAN4. All manual model building steps were carried out with Coot and *phenix.refine* was used for iterative refinement. Near the end of refinement for $Nup133_{NTD}$+VHH-SAN5, TLS parameters were used.

Liposome preparation

To prepare liposomes, a thin-film of yeast polar lipid extract (YPL, Avanti Polar Lipids) at 10 mg/ml in chloroform was dried in a glass vial under nitrogen gas, washed twice with anhydrous pentane (Sigma-Aldrich) and resuspended overnight at 10 mg/ml in 10 mM Tris/HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT. Following resuspension, three freeze-thaw cycles were completed. To obtain unilamellar liposomes, extrusion (Avanti Mini-extruder) was

519 completed with 21 passes through either a 0.2 µM, 0.1 µM, or 0.05 µM polycarbonate
520 membrane (Avanti Polar Lipids). Liposomes were stored at 4°C and used within a week.
521
522 **Liposome Flotation Assay**
523 Liposomes preparations and protein purification were completed as detailed above. A
524 total of 1µg protein was incubated with 500 µg yeast polar lipids in 50 µl of floatation buffer (100
525 mM potassium acetate, 2 mM sodium acetate, 20 mM HEPES/NaOH pH 7.4, 2 mM DTT) for 30
526 mins at room temperature. After incubation, the 50 µl liposome sample was mixed with 644µl
527 48% Nycodenz (5-(N-2,3-dihydroxypropylacetamido)-2,4,6-tri-iodo-N,N-bis(2,3-
528 dihydroxypropyl)isophthalamide, Accurate Chemical) solution in flotation buffer and applied to
529 the bottom of a 11 x 34 mm ultracentrifugation thick-wall tube. A 1.524 ml layer of 40%
530 nycodenz solution was layered above, followed by a 592 µl layer of flotation buffer. Samples
531 underwent ultracentrifugation in an SW55Ti rotor for 2hrs at 4°C, 240,000 x g in an Optima Max
532 Ultracentrifuge (Beckman). Following ultracentrifugation, four fractions were taken, top (832 µl),
533 middle 1 (566 µl), middle 2 (566 µl), and bottom (844 µl). Protein was extracted from the
534 fractions by methanol chloroform precipitation. Analysis of the assay was completed by a 12%
535 SDS-PAGE visualized with silver stain (Thermo Fisher Scientific).
536
537 **Negative Stain Electron Microscopy**
538 $Nup133_{NTD}$-liposome mixtures at 1 mg/ml were loaded on glow-discharged (EMS 100,
539 Electron Microscopy Sciences) continuous carbon film grids (CF200-Cu, Electron Microscopy
540 Sciences). After 45 sec of adsorption on grids, the samples were blotted with Whatman filter
541 paper and the specimen on the grid was immediately stained with 2% w/v uranyl acetate for
542 30 sec. The specimen was blotted, stained once more, re-blotted, and air dried. Electron
543 micrographs were recorded on a FEI Tecnai Spirit BioTwin microscope (FEI) operated at 80 keV
544 and equipped with a tungsten filament and an AMT XR16 CCD detector.

545

546 **Fitting the composite Y complex model into the cryo-ET map of the *S. cerevisiae* NPC**

547 A global fitting approach was performed using our composite model of the Y complex
548 into the cryo-ET map (EMD-10198) from *S. cerevisiae*[4]. Density corresponding to the
549 cytoplasmic and nucleoplasmic rings alone were conservatively defined as subregions in
550 Chimera to eliminate positioning in the nuclear envelope or inner ring, respectively. Docking was
551 done with Fitmap in Chimera, with 10,000 placements with an apparent resolution of 30 Å for
552 the Y complex. Scoring was done as correlation around zero. The highest scoring solution for
553 the cytoplasmic placement had a correlation score of 0.68 (Fig. 7a). The highest scoring
554 solution for the nucleoplasmic placement had a correlation score of 0.73 (Fig. 7b). Fitting of our
555 previous model of the Y complex[25] had a correlation score of 0.66 in the cytoplasmic ring and
556 0.61 in the nucleoplasmic ring (Fig. 7).

557

558 **PDB accession codes**

559 Coordinates and structure factors have been deposited in the Protein Data Bank under
560 accession codes PDB 6X02 (Nup84-Nup133$_{CTD}$-VHH-SAN8), 6X03 (Nup84-Nup133$_{CTD}$-VHH-
561 SAN8/9), 6X04 (Nup133$_{NTD}$-VHH-SAN5), 6X05 (Nup133$_{NTD}$-VHH-SAN4).

562

563 **Acknowledgements**

564 We thank Alexander Ulrich for generating initial expression constructs and preliminary
565 crystals for the Nup84-133 complex. Research was supported by the NIH under grant number
566 R01GM77537 (T.U.S.) and T32GM007287 (S.A.N.) The X-ray data collection was conducted at
567 the APS NE-CAT beamlines, supported by award GM103403 from the NIH. Use of the APS is
568 supported by the US Department of Energy, Office of Basic Energy Sciences, under contract
569 no. DE-AC02-06CH11357.

570

Author Contributions

S.A.N. and T.U.S. designed the study. S.A.N. performed the experiments. D.L.T. assisted with the liposome interaction experiments. S.A.N. and T.U.S analyzed and interpreted the data, and wrote the manuscript with input from D.L.T.

Competing Interests

The authors declare no competing financial interests.

References

1. Strambio-De-Castillia, C., Niepel, M. & Rout, M. P. The nuclear pore complex: bridging nuclear transport and gene regulation. *Nat Rev Mol Cell Bio* **11**, 490 (2010).

2. Appen, A. von *et al.* In situ structural analysis of the human nuclear pore complex. *Nature* **526**, 140–3 (2015).

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

4. Allegretti, M. *et al.* In cell architecture of the nuclear pore complex and snapshots of its turnover. *Biorxiv* 2020.02.04.933820 (2020) doi:10.1101/2020.02.04.933820.

5. Knockenhauer, K. E. & Schwartz, T. U. The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell* **164**, 1162–1171 (2016).

6. Schwartz, T. U. The Structure Inventory of the Nuclear Pore Complex. *J Mol Biol* **428**, 1986–2000 (2016).

7. Beck, M. & Hurt, E. The nuclear pore complex: understanding its function through structural insight. *Nat Rev Mol Cell Bio* **18**, 73–89 (2016).

8. Lin, D. H. & Hoelz, A. The Structure of the Nuclear Pore Complex (An Update). *Annu Rev Biochem* **88**, 725–783 (2019).

9. Harel, A. *et al.* Removal of a Single Pore Subcomplex Results in Vertebrate Nuclei Devoid of Nuclear Pores. *Mol Cell* **11**, 853–864 (2003).

10. Walther, T. C. *et al.* The Conserved Nup107-160 Complex Is Critical for Nuclear Pore Complex Assembly. *Cell* **113**, 195–206 (2003).

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

12. Leksa, N. C., Brohawn, S. G. & Schwartz, T. U. The Structure of the Scaffold Nucleoporin Nup120 Reveals a New and Unexpected Domain Architecture. *Structure* **17**, 1082–1091 (2009).

13. Brohawn, S. G., Leksa, N. C., Spear, E. D., Rajashankar, K. R. & Schwartz, T. U. Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. *Sci New York N Y* **322**, 1369–73 (2008).

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

15. Debler, E. W. *et al.* A Fence-like Coat for the Nuclear Pore Membrane. *Mol Cell* **32**, 815–826 (2008).

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

17. Brohawn, S. G. & Schwartz, T. U. Molecular architecture of the Nup84-Nup145C-Sec13 edge element in the nuclear pore complex lattice. *Nat Struct Mol Biol* **16**, 1173–7 (2009).

18. Nagy, V. *et al.* Structure of a trimeric nucleoporin complex reveals alternate oligomerization states. *P Natl Acad Sci Usa* **106**, 17693–8 (2009).

19. Hsia, K.-C., Stavropoulos, P., Blobel, G. & Hoelz, A. Architecture of a Coat for the Nuclear Pore Membrane. *Cell* **131**, 1313–1326 (2007).

20. Boehmer, T., Jeudy, S., Berke, I. C. & Schwartz, T. U. 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).

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

22. Whittle, J. R. R. & Schwartz, T. U. Architectural Nucleoporins Nup157/170 and Nup133 Are Structurally Related and Descend from a Second Ancestral Element. *J Biol Chem* **284**, 28442–28452 (2009).

23. Sampathkumar, P. *et al.* Structure of the C-terminal domain of Saccharomyces cerevisiae Nup133, a component of the nuclear pore complex. *Proteins* **79**, 1672–7 (2011).

24. Kim, S. J. *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).

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

26. Stuwe, T. *et al.* Architecture of the nuclear pore complex coat. *Science* **347**, 1148–1152 (2015).

27. Bui, K. H. *et al.* Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. *Cell* **155**, 1233–1243 (2013).

28. Kampmann, M. & Blobel, G. Three-dimensional structure and flexibility of a membrane-coating module of the nuclear pore complex. *Nat Struct Mol Biol* **16**, 782–788 (2009).

29. Brohawn, S. G., Partridge, J. R., Whittle, J. R. R. & Schwartz, T. U. The Nuclear Pore Complex Has Entered the Atomic Age. *Structure* **17**, 1156–1168 (2009).

643 30. Lin, D. H. *et al.* Architecture of the symmetric core of the nuclear pore. *Science* **352**,
644 aaf1015 (2016).

645 31. Kosinski, J. *et al.* Molecular architecture of the inner ring scaffold of the human nuclear pore
646 complex. *Science* **352**, 363–365 (2016).

647 32. Kim, S. J. *et al.* Integrative structure and functional anatomy of a nuclear pore complex.
648 *Nature* **555**, 475–482 (2018).

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

651 34. Rajoo, S., Vallotton, P., Onischenko, E. & Weis, K. Stoichiometry and compositional
652 plasticity of the yeast nuclear pore complex revealed by quantitative fluorescence microscopy.
653 *Proc National Acad Sci* **115**, E3969–E3977 (2018).

654 35. Rout, M. P. *et al.* The Yeast Nuclear Pore Complex. *J Cell Biology* **148**, 635–652 (2000).

655 36. Drin, G. *et al.* A general amphipathic α-helical motif for sensing membrane curvature. *Nat*
656 *Struct Mol Biol* **14**, 138–146 (2007).

657 37. Doucet, C. M., Talamas, J. A. & Hetzer, M. W. Cell cycle-dependent differences in nuclear
658 pore complex assembly in metazoa. *Cell* **141**, 1030–41 (2010).

659 38. Schooley, A., Vollmer, B. & Antonin, W. Building a nuclear envelope at the end of mitosis:
660 coordinating membrane reorganization, nuclear pore complex assembly, and chromatin de-
661 condensation. *Chromosoma* **121**, 539–54 (2012).

662 39. Ingram, J. R., Schmidt, F. I. & Ploegh, H. L. Exploiting Nanobodies' Singular Traits. *Annu*
663 *Rev Immunol* **36**, 695–715 (2018).

664 40. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize
665 evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344-50 (2016).

666 41. Shapovalov, M. V. & Dunbrack, R. L. A smoothed backbone-dependent rotamer library for
667 proteins derived from adaptive kernel density estimates and regressions. *Struct Lond Engl 1993*
668 **19**, 844–58 (2011).

669 42. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web
670 portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845–858 (2015).

671 43. Dokudovskaya, S. *et al.* A conserved coatomer-related complex containing Sec13 and Seh1
672 dynamically associates with the vacuole in Saccharomyces cerevisiae. *Mol Cell Proteom Mcp*
673 **10**, M110.006478 (2011).

674 44. Algret, R. *et al.* Molecular architecture and function of the SEA complex, a modulator of the
675 TORC1 pathway. *Mol Cell Proteom Mcp* **13**, 2855–70 (2014).

676 45. Xu, C. & Min, J. Structure and function of WD40 domain proteins. *Protein Cell* **2**, 202–14
677 (2011).

678 46. Höfer, C. T., Herrmann, A. & Müller, P. Use of liposomes for studying interactions of soluble
679 proteins with cellular membranes. *Methods Mol Biology Clifton N J* **606**, 69–82 (2010).

680 47. Vollmer, B. *et al.* Dimerization and direct membrane interaction of Nup53 contribute to
681 nuclear pore complex assembly. *Embo J* **31**, 4072–84 (2012).

682 48. Mészáros, N. *et al.* Nuclear Pore Basket Proteins Are Tethered to the Nuclear Envelope and
683 Can Regulate Membrane Curvature. *Dev Cell* **33**, 285–298 (2015).

684 49. Cassidy, C. K., Himes, B. A., Luthey-Schulten, Z. & Zhang, P. CryoEM-based hybrid
685 modeling approaches for structure determination. *Curr Opin Microbiol* **43**, 14–23 (2017).

686 50. Zanetti, G. *et al.* The structure of the COPII transport-vesicle coat assembled on
687 membranes. *Elife* **2**, e00951 (2013).

688 51. Hinshaw, J. E. & Milligan, R. A. Nuclear pore complexes exceeding eightfold rotational
689 symmetry. *J Struct Biol* **141**, 259–268 (2003).

690 52. Löschberger, A., Franke, C., Krohne, G., Linde, S. van de & Sauer, M. Correlative super-
691 resolution fluorescence and electron microscopy of the nuclear pore complex with molecular
692 resolution. *J Cell Sci* **127**, 4351–5 (2014).

693 53. Stanley, G. J., Fassati, A. & Hoogenboom, B. W. Atomic force microscopy reveals structural
694 variability amongst nuclear pore complexes. *Life Sci Alliance* **1**, e201800142 (2018).

695 54. Beck, M., Lučić, V., Förster, F., Baumeister, W. & Medalia, O. Snapshots of nuclear pore
696 complexes in action captured by cryo-electron tomography. *Nature* **449**, 611 (2007).

697 55. Panté, N. & Kann, M. Nuclear pore complex is able to transport macromolecules with
698 diameters of about 39 nm. *Mol Biol Cell* **13**, 425–34 (2002).

699 56. Sage, V. L. & Mouland, A. J. Viral subversion of the nuclear pore complex. *Viruses* **5**, 2019–
700 42 (2013).

701 57. Laba, J. K., Steen, A. & Veenhoff, L. M. Traffic to the inner membrane of the nuclear
702 envelope. *Curr Opin Cell Biol* **28**, 36–45 (2014).

703 58. Ungricht, R. & Kutay, U. Establishment of NE asymmetry—targeting of membrane proteins
704 to the inner nuclear membrane. *Curr Opin Cell Biol* **34**, 135–41 (2015).

705 59. King, M. C., Lusk, C. & Blobel, G. Karyopherin-mediated import of integral inner nuclear
706 membrane proteins. *Nature* **442**, 1003–1007 (2006).

707 60. Meinema, A. C. *et al.* Long unfolded linkers facilitate membrane protein import through the
708 nuclear pore complex. *Sci New York N Y* **333**, 90–3 (2011).

709 61. Otsuka, S. *et al.* Nuclear pore assembly proceeds by an inside-out extrusion of the nuclear
710 envelope. *Elife* **5**, e19071 (2016).

711 62. Vollmer, B. *et al.* Nup153 Recruits the Nup107-160 Complex to the Inner Nuclear Membrane
712 for Interphasic Nuclear Pore Complex Assembly. *Dev Cell* **33**, 717–28 (2015).

713 63. Fernandez-Martinez, J. *et al.* Structure–function mapping of a heptameric module in the
714 nuclear pore complex. *J Cell Biology* **196**, 419–434 (2012).

715 64. Pemberton, L. F., Rout, M. P. & Blobel, G. Disruption of the nucleoporin gene NUP133
716 results in clustering of nuclear pore complexes. *Proc National Acad Sci* **92**, 1187–1191 (1995).

65. Heath, C. V. *et al.* Nuclear pore complex clustering and nuclear accumulation of poly(A)+ RNA associated with mutation of the Saccharomyces cerevisiae RAT2/NUP120 gene. *J Cell Biology* **131**, 1677–1697 (1995).

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

67. Frey, S. & Görlich, D. A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J Chromatogr A* **1337**, 95–105 (2014).

68. Andersen, K. R., Leksa, N. C. & Schwartz, T. U. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins* **81**, 1857–61 (2013).

69. Ingram, J. R. *et al.* Allosteric activation of apicomplexan calcium-dependent protein kinases. *P Natl Acad Sci Usa* **112**, E4975-84 (2015).

70. Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, e01456 (2013).

71. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–26 (1997).

72. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biological Crystallogr* **66**, 133–144 (2010).

73. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biological Crystallogr* **67**, 235–42 (2011).

74. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biological Crystallogr* **66**, 213–21 (2010).

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

76. DiMaio, F. *et al.* Improved low-resolution crystallographic refinement with Phenix and Rosetta. *Nat Methods* **10**, 1102–4 (2013).

77. Decanniere, K. *et al.* A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* **7**, 361–370 (1999).

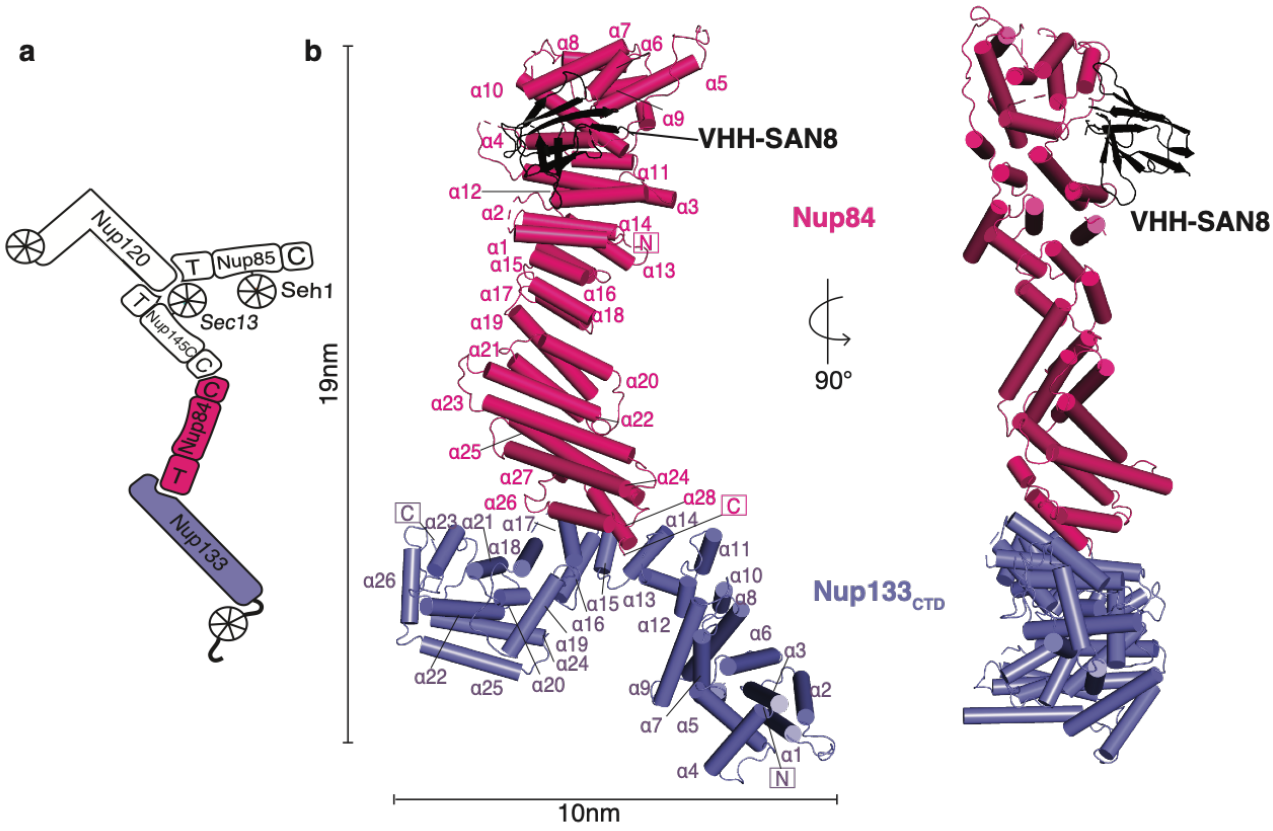

Figure 1.

Structure of the *Saccharomyces cerevisiae* Nup84-Nup133$_{CTD}$-VHH-SAN8 complex.
(a) Schematic of the *S. cerevisiae* Y complex. Regions included in the structure are colored (Pink – Nup84, Purple – Nup133). Elements of ACE1 fold proteins are indicated: T – tail and C – crown adjacent to the central trunk element. (b) Structure of Nup84-Nup133$_{CTD}$-VHH-SAN8, with Nup84 in pink, Nup133$_{CTD}$ in purple, and VHH-SAN8 in black. N and C termini are indicated and helices are numbered. Right panel shows the side view (90° rotation) of the complex.

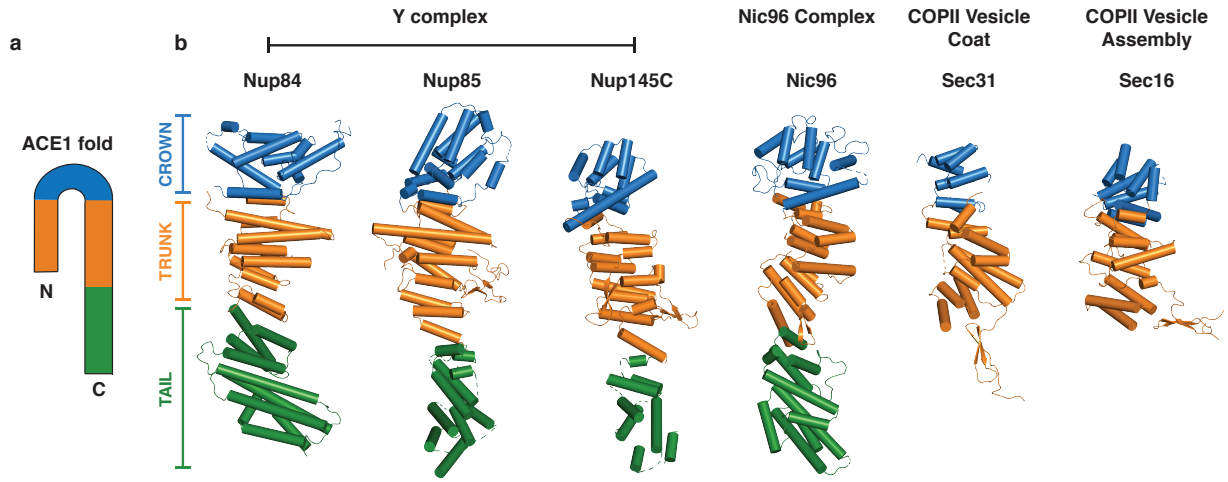

Figure 2.

ACE1 domain proteins of the Nuclear Pore and COPII vesicle coat.

(a) ACE1 fold topology schematic. (b) ACE1 fold proteins. Domains indicated by color: Crown (blue), trunk, (orange), tail (green). PDB codes are as follows: Nup84 (this study), Nup85 (4XMM), Nup145C (4XMM), Nic96 (2QX5), Sec31 (2PM6), Sec16 (3MZK).

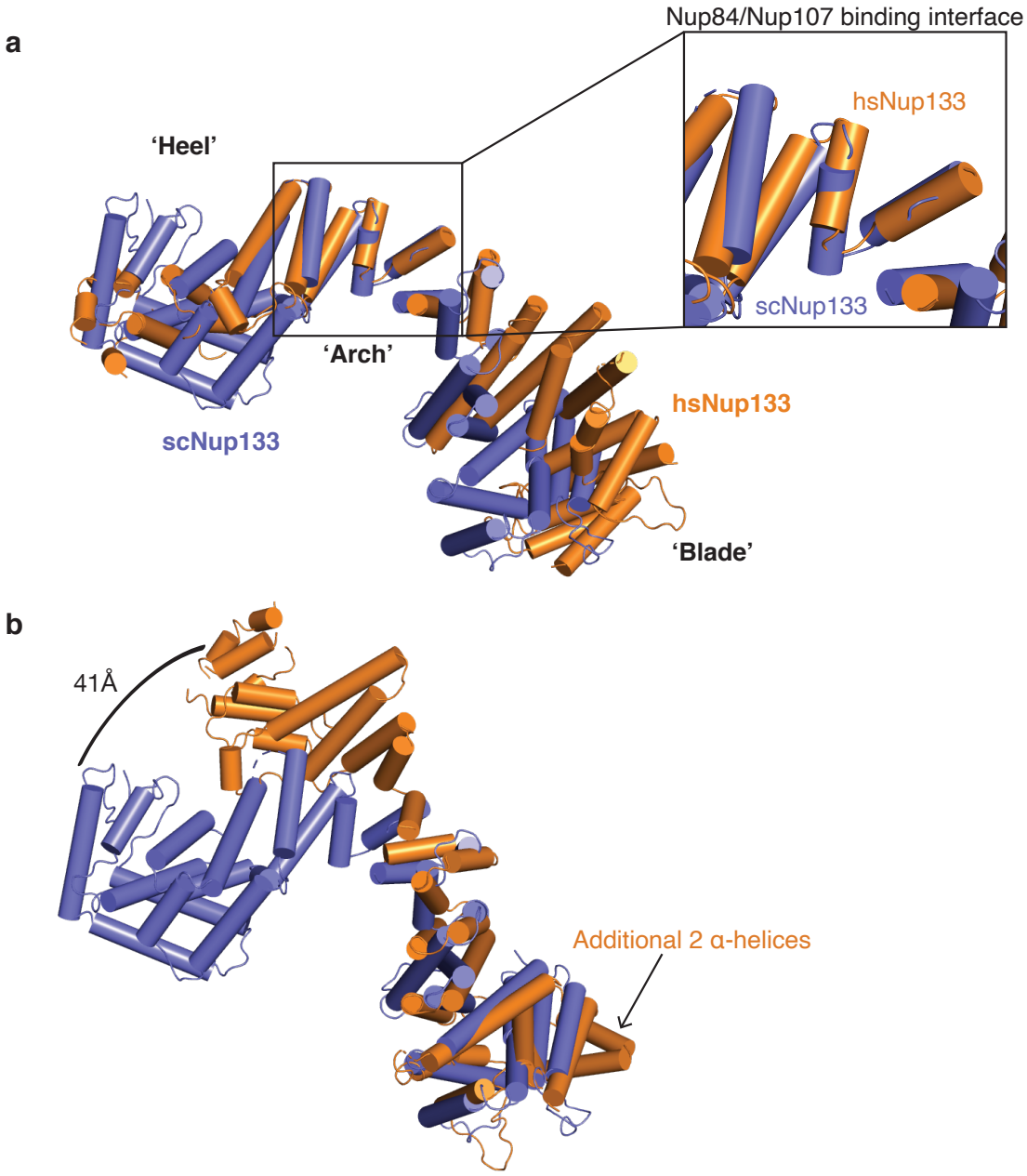

Figure 3.

Comparison between scNup133$_{CTD}$ and hsNup133$_{CTD}$.
(a) Superposition of hsNup133 (orange, PDB: 3I4R) and scNup133 (purple), using the 'arch' element for matching. Inset zooms in on the 'arch' where scNup84/hsNup107 bind. (b) Same superposition as in (a), but using the 'blade' element for matching. Arrow indicates additional helices in the hsNup133 blade domain. This superposition emphasizes the large, 41Å conformational shift between the two heel domains.

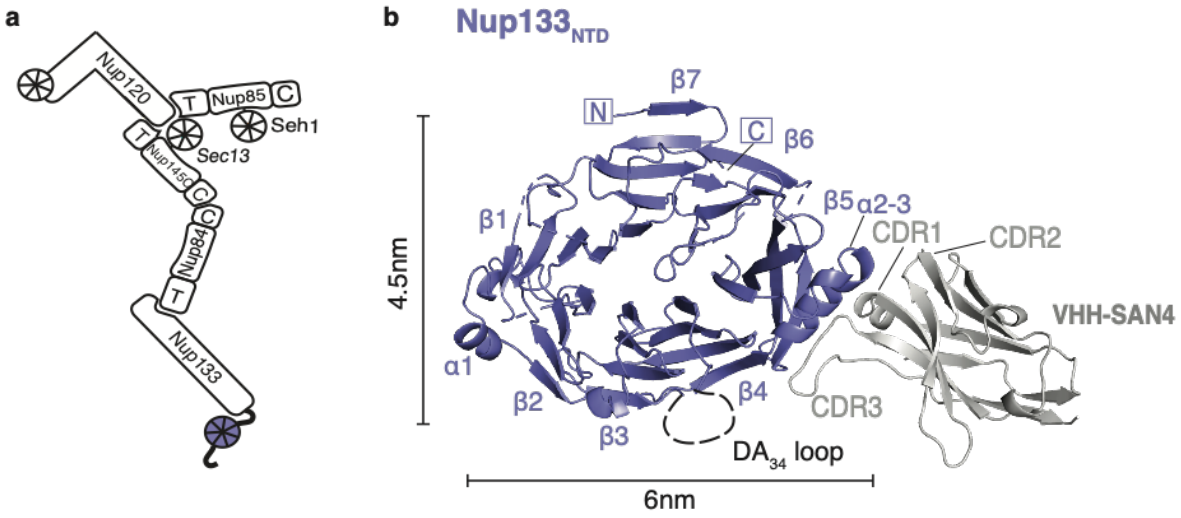

Figure 4.

Structure of $Nup133_{NTD}$-VHH-SAN4 complex.

(a) Schematic of the Y complex. The region included in the structure is colored (purple – Nup133). Elements of ACE1 fold proteins are indicated: T – tail and C – crown adjacent to the central trunk element. (b) $Nup133_{NTD}$-VHH-SAN4, with $Nup133_{NTD}$ in purple and VHH-SAN4 in light gray. N and C termini are indicated, and β-sheets and α-helices are labeled on $Nup133_{NTD}$. Complementarity determining region (CDR) loops are labeled on VHH-SAN4.

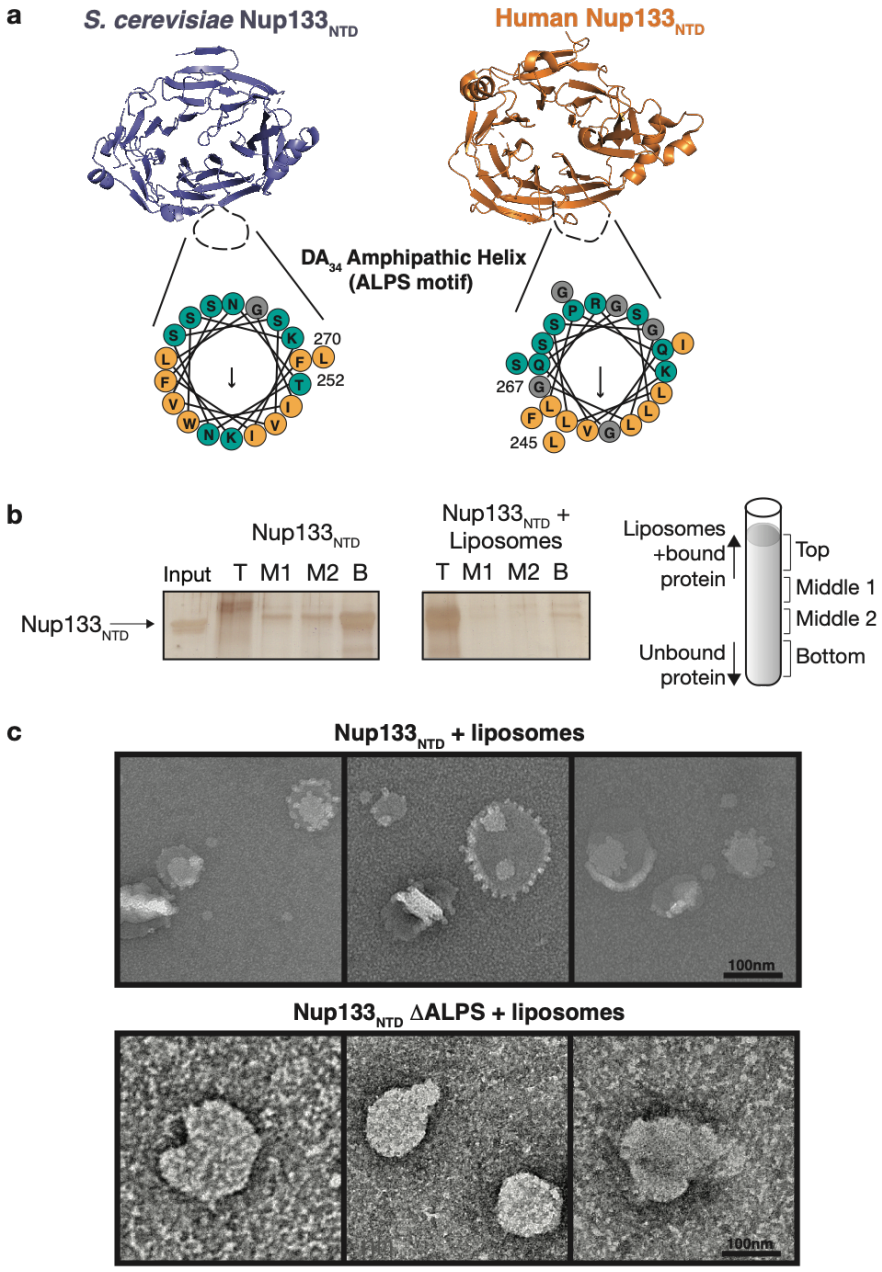

Figure 5.

Nup133$_{NTD}$ has an ALPS motif.

(a) Structures of scNup133$_{NTD}$ (purple) and hsNup133$_{NTD}$ (orange). The missing DA$_{34}$ loop in the structures are indicated by a black dotted line. Helical wheel diagram for each loop is shown with polar residues (teal), non-polar (orange), and glycine (gray) colored. Arrow indicates predicted hydrophobic face and length of arrow scales with strength of the hydrophobic moment. (b) Liposome floatation assay with Nup133$_{NTD}$. Panels show silver stain SDS-PAGE fractions isolated from the gradients. A cartoon of fractions is shown on the right. Left-most panel is with Nup133$_{NTD}$ alone (no liposomes) and middle panel is Nup133$_{NTD}$ with liposomes. (c) Negative stain electron microscopy images of liposomes pre-incubated with Nup133$_{NTD}$ (upper) and Nup133$_{NTD}$ ΔALPS (lower). Scale bar at the bottom left of each panel.

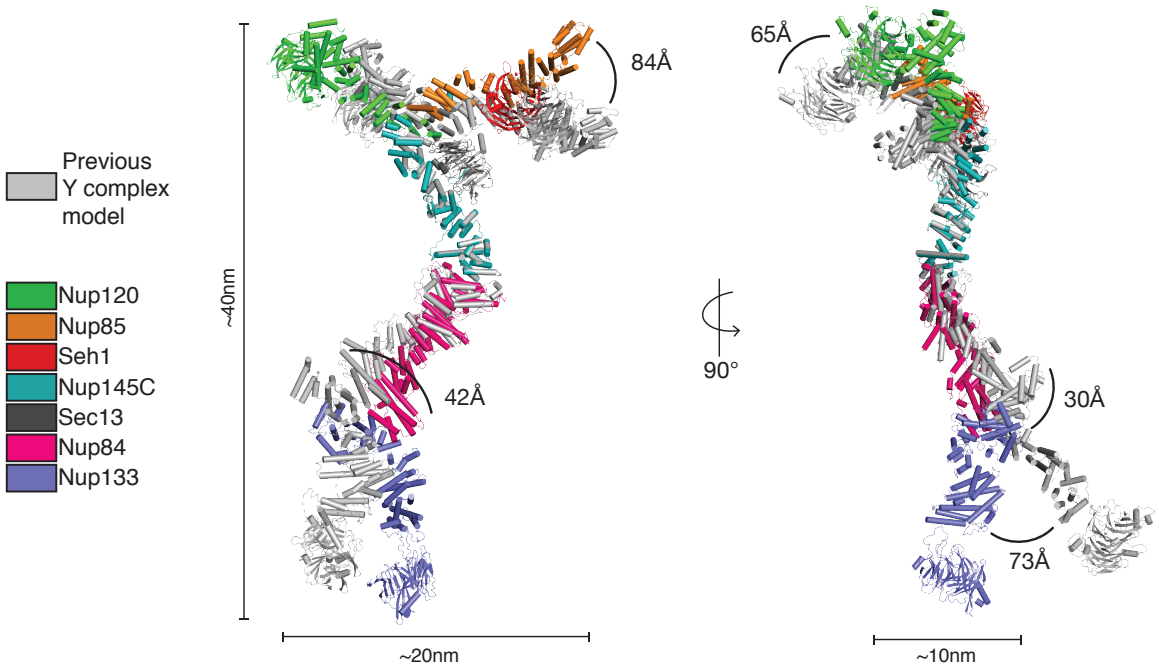

Figure 6.

New composite model of the Y complex.

New composite model of the *S. cerevisiae* Y complex super-positioned with the previous *S. cerevisiae* composite model[25], using Nup145C for matching. Colors of each nup are indicated in the key on the left. Overall dimensions of the Y complex are indicated. Conformational differences between the models are shown by arcs between the same nups, with the distance of each listed.

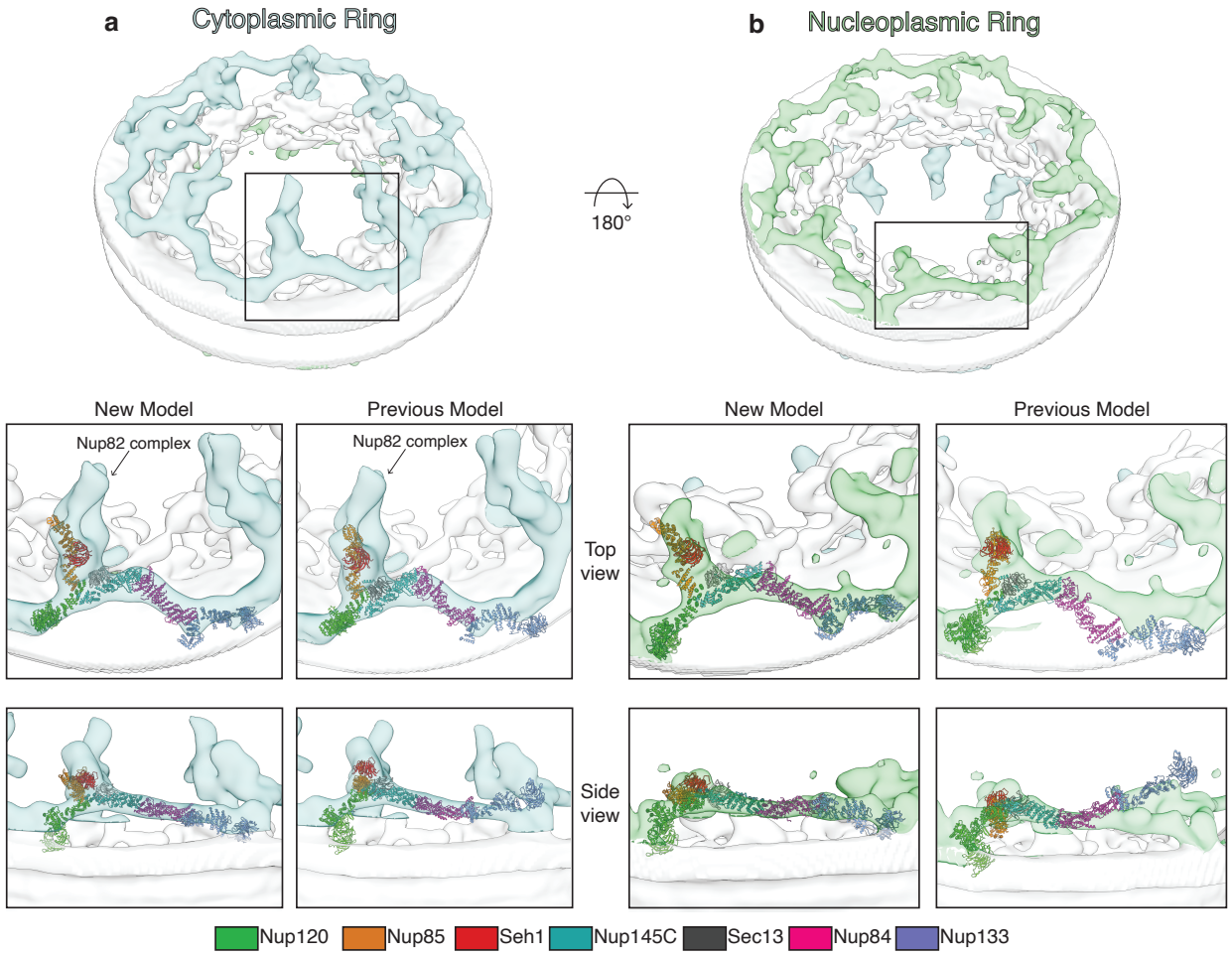

Figure 7.

Fitting of the composite *S. cerevisiae* Y complex into the cryo-ET map of the NPC. Cryo-ET map from *S. cerevisiae* (EMD-10198) with the density corresponding to the (a) cytoplasmic ring of the NPC colored in light blue and (b) the nucleoplasmic ring colored in light green. The top-scoring global fit for both the old (left panel) and new (right panel) composite *S. cerevisiae* Y complex models shown in both a top and side view. Y complex nups colored as in legend below.

806

Table 1. Data collection and refinement statistics

Protein	Nup84, $Nup133_{521-1157}$, VHH-SAN8 (PDB 6X02)	Nup84, $Nup133_{521-1157}$, VHH-SAN8 TeW derivative	Nup84, $Nup133_{521-1157}$, VHH-SAN8, VHH-SAN9 (PDB 6X03)	$Nup133_{55-481}$, VHH-SAN4 (PDB 6X05)	$Nup133_{55-481}$, VHH-SAN5 (PDB 6X04)
Organism	*S. cerevisiae, V. pacos*				
Data collection					
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$
Cell Dimensions					
a, b, c (Å)	72.3, 287.7, 297.4	72.5, 283.6, 295.7	72.3, 295.2, 295.6	93.4, 154.7, 41.4	85.7, 204.7, 205.2
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	103.4-6.4 (6.63-6.40)[a]	147.8-7.8 (8.72-7.80)	147.8-7.4 (7.66-7.40)	59.6-2.1 (2.18-2.10)	49.7-2.7 (2.80-2.70)
$R_{p.i.m.}$	2.2 (81.3)	4.2 (42.0)	2.7 (61.7)	5.2 (48.0)	3.5 (68.4)
I/σ	55.3 (1.0)	13.8 (2.1)	46.6 (1.1)	19.3 (1.5)	29.1 (1.5)
$CC_{1/2}$	1.00 (0.58)	1.00 (0.89)	0.97 (0.54)	1.00 (0.51)	1.00 (0.57)
Completeness (%)	99.8 (100.0)	99.9 (100.0)	99.9 (100.0)	99.9 (99.8)	99.6 (97.3)
Redundancy	24.1 (25.5)	12.3 (12.9)	12.0 (12.7)	11.9 (10.3)	8.6 (8.3)
Refinement					
Resolution (Å)	103.4-6.4		147.8-7.4	59.6-2.1	49.7-2.7
No. Reflections	13,453		9,375	35,967	100,268
R_{work}/R_{Free}	31.4 / 32.4		33.5 / 35.3	18.3 / 22.3	23.6 / 25.9
No. Atoms					
Protein	6,655		7,106	3,890	20,184
Water	0		0	285	17
***B* factors (Å²)**					
Protein	574.6		668.3	43.0	88.5
Water	-		-	43.0	46.9
r.m.s. deviations					
Bond length (Å)	0.003		0.003	0.007	0.003
Bond angles (°)	0.84		0.81	0.87	0.57

807

808 [a]Values in parenthesis are for highest-resolution shell (10% of the data). One crystal was used
809 for each dataset.