Title: AI-based structure prediction empowers integrative structural analysis of human nuclear pores

Authors: Shyamal Mosalaganti[1,2,3†], Agnieszka Obarska-Kosinska[1,4†], Marc Siggel[4,5,6†], Reiya Taniguchi[1,2], [12]Beata Turoňová[1,2], Christian E. Zimmerli[1,2], Katarzyna Buczak[2‡], Florian H. Schmidt[2§], Erica Margiotta[1,2], Marie-Therese Mackmull[2¶], Wim J. H. Hagen[2], Gerhard Hummer[5,7]*, Jan Kosinski[2,4,6]*, Martin Beck[1,2]*

Affiliations:

[1] Department of Molecular Sociology, Max Planck Institute of Biophysics, Max-von-Laue-Straße 3, 60438 Frankfurt am Main, Germany.

[2] Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstraße 1, 69117, Heidelberg, Germany.

[3] Life Sciences Institute, Department of Cell & Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA.

[4] European Molecular Biology Laboratory Hamburg, Notkestraße 85, 22607, Hamburg, Germany.

[5] Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Max-von-Laue-Straße 3, 60438 Frankfurt am Main, Germany.

[6] Centre for Structural Systems Biology (CSSB), Notkestraße 85, 22607, Hamburg, Germany.

[7] Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany.

*Corresponding authors email: gerhard.hummer@biophys.mpg.de, jan.kosinski@embl.de, martin.beck@biophys.mpg.de

† These authors contributed equally to this work

‡ Present address: Proteomics Core Facility, Biozentrum, University of Basel, Spitalstrasse 41, CH-4056 Basel / Switzerland.

§ Present address: Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria.

¶ Present address: Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Zurich, Switzerland.

Abstract: Nuclear pore complexes (NPCs) mediate nucleocytoplasmic transport. Their intricate 120 MDa architecture remains incompletely understood. Here, we report a 70 MDa model of the human NPC scaffold with explicit membrane and in multiple conformational states. We combined AI-based structure prediction with in situ and in cellulo cryo-electron tomography and integrative modeling. We show that linker NUPs spatially organize the scaffold within and across subcomplexes to establish the higher-order structure. Microsecond-long molecular dynamics simulations suggest that the scaffold is not required to stabilize the inner and outer nuclear membrane fusion, but rather widens the central pore. Our work exemplifies how AI-based modeling can be integrated with in situ structural biology to understand subcellular architecture across spatial organization levels.

One-Sentence Summary: An AI-based, dynamic model of the human nuclear pore complex reveals how the protein scaffold and the nuclear envelope are coupled inside cells.

Main Text:

Nuclear Pore Complexes (NPCs) are essential for the transport between the nucleus and cytoplasm, and critical for many other cellular processes in eukaryotes (*1–4*). Analysis of the structure and dynamics of the NPC at high resolution has been a longstanding goal towards a better molecular understanding of NPC function. The respective investigations have turned out challenging because of the sheer size, and their compositional and architectural complexity. With a molecular weight of about 120 MDa, NPCs form an extensive 120 nm-wide protein scaffold of three stacked rings: two outer rings—the cytoplasmic (CR) and the nuclear ring (NR), and the inner ring (IR). Each ring comprises eight spokes that surround a 40 to 50 nm-wide transport channel (*5*, *6*). A single human NPC contains ~30 distinct nucleoporins (NUPs) in approximately 1,000 copies altogether. They arrange in multiple subcomplexes, most prominently the so-called Y-complex (*7*) arranged in head-to-tail within the outer rings (*8*). The assembly of individual subcomplexes into the higher order structure is facilitated by a yet incompletely characterized network of short linear motifs (SLiMs) embedded into flexible NUP linkers (*9–13*), which have been conceived as a molecular glue that stabilizes the scaffold. As a further complication, the assembled scaffold is embedded into the nuclear envelope (NE). Components of the NPC scaffold interact with the NE via amphipathic helices and transmembrane domains, and are believed to stabilize the fusion of the inner (INM) and the outer nuclear membrane (ONM) (*14*, *15*). Finally, the scaffold grafts FG-NUPs to the central channel, which form the permeability barrier (*16–18*) with their intrinsically disordered FG-rich domains, making them challenging to study using structural biology methods.

Due to these intricacies, the current structural models have severe shortcomings. In case of the human NPC, only 16 NUPs, accounting for ~35 MDa (30%) of the molecular weight of the complex are included in the models (*11*, *19*, *20*). Although the repertoire of atomically resolved structures of NUPs has tremendously grown (*5*, *6*) the respective structures often have gaps in their sequence coverage, while homology models used by many studies have intrinsic inaccuracies. For some NUPs, no structures or homology models are available. Also, structural models put forward for other species were either incomplete or had a limited precision (*11*, *12*, *19*, *21–23*). Moreover, the NPCs from many other species have a vastly reduced architectural complexity, which limits their usefulness for studying the human biology (*12*, *22–25*). The exact grafting sites for FG-NUPs, which are crucial for understanding the transport mechanism, remain elusive. How exactly the NPC scaffold is anchored to the membrane, how it responds to mechanical cues imposed by the nuclear envelope, and if and how it contributes to shaping the membrane, remains unknown. Finally, the models remain static snapshots that do not take conformational dynamics into account.

Here, we have combined cryo-electron tomography (cryo-ET) analysis, of the human NPC from isolated NEs and within intact cells, with artificial intelligence (AI)-based structural prediction to infer a model of >90% of the human NPC scaffold at unprecedented precision and in multiple conformations. We demonstrate that AI-based models of NUPs and their subcomplexes built using AlphaFold (*26*) and RoseTTAfold (*27*) are consistent with unreleased X-ray crystallography structures, cryo-EM maps, and complementary data. We elucidate the three-dimensional (3D) trajectory of linker NUPs, the organization of membrane-binding domains and grafting sites of most FG-NUPs in both, the constricted and dilated conformation.

Results

A 70 MDa model of the human NPC scaffold

The completeness of the previous structural models of the human NPC was limited by the resolution of the available EM maps in both the constricted and the dilated states, and the lack of atomic structures for several NUPs (*19–21*). To improve the resolution of the constricted state of the NPC, we subjected nuclear envelopes purified from HeLa cells to cryo-ET analysis, as described previously (*19*, *21*). We collected an about five times larger dataset and applied a novel geometrically-restraint classification procedure (see Methods). These improvements resulted in EM maps with resolutions of 12 Å for the CR, 12.6 Å for IR, and 23.2 Å for NR (**Fig. S1**, **Fig. S2**, **Fig. S3**). Next, we obtained an *in cellulo* cryo-ET map of dilated human NPCs in the native cellular environment within intact HeLa and HEK293 cells subjected to cryo-Focused Ion Beam (cryo-FIB) specimen thinning (**Fig. S1**). The dilated, *in cellulo* NPC exhibits the IR diameter of 54 Å in comparison to 42 Å in the constricted state, consistent with previous work in u2os (*28*), HeLa (*29*), SupT1 (*30*), and most recently, DLD-1 cells (*20*). In contrast to other species, there is no compaction along the nucleocytoplasmic axis during dilation (**Fig. S2**) (*23*, *25*). The quality of our *in cellulo* map is sufficient to discern the structural features known from the constricted state, such as a double head-to-tail arrangement of Y-complexes, IR subunits, and inter-ring connectors. In agreement with the previous cryo-ET maps of NPCs (*20*, *22*, *23*, *25*), we observe increase in the distance between the adjacent spokes within the IR.

To generate a comprehensive set of structural models of human NUPs, we used the recently published protein structure prediction software AlphaFold (*26*) and RoseTTAfold (*27*). We found that most of the NUPs can be modeled with high confidence scores (**Fig. S5**, **Table S1**). In addition, we validated the accuracy of the models by comparison to structures from accompanying publications (*31*, *32*) of human NUP358, NUP93, NUP88, NUP98, and of NUP205 and NUP188 from *Chaetomium thermophilum* (**Fig. S6**, **Table S1**). These structures were not used as input for the modeling procedure. The AI-based models also excellently fit our EM densities-with significant p-values and high cross-correlation scores (**Fig. S7A-E**). Furthermore, we used single particle cryo-EM to determine the experimental structure of human NUP155 (**Fig. S8A-C**) and validated the respective AlphaFold model. Although the model and the structure do not perfectly superpose as whole chains due to the flexibility of the protein, their local tertiary structures and side chain conformations are highly similar (global LDDT score 91.6, **Fig. S9A-E**). Notably, the loops that were not resolved in the experimentally derived structure consistently show low pLDDT scores (**Fig. S9B**), further supporting the reliability of this metric.

With full-length models at hand, we could identify the positions of NUP205 and NUP188 within the scaffold, which could not be unambiguously located in the previous hNPC cryo-ET maps (see below). The AI-predicted conformation of the N-terminal domain of NUP358 is more consistent with the observed electron density as compared to the two X-ray structures (**Fig. S7H**). The NUP358 localization is in agreement with previous analysis (*21*) and the secondary structure in the *Xenopus* EM map (*33*) **(Fig. S10**). The full-length model of ELYS, for which thus far only the N-terminal b-propeller could be placed (*21*), fits the EM map as a rigid body (**Fig. S7E**) and confirms its binding site to each of the Y-complexes in the NR. The models of NUPs in the CR agree with the secondary structure observed in the *Xenopus* cryo-EM map (**Fig. S10**).

The capacity of AI-based structure prediction tools to identify and model protein interfaces with high accuracy has recently been demonstrated (*34–36*). We therefore attempted to model NUP interfaces using the ColabFold software, a version of AlphaFold adopted for modeling protein complexes (*35*). We found that ColabFold predicted several NUP subcomplexes with inter-domain confidence scores that correlated with the accuracy of the models, while negative controls with non-specific interactions yielded low confidence models (**Fig. S11**, **Fig. S12**,

Fig. S13, **Fig. S14**). The models of these subcomplexes did not only reproduce the respective, already available X-ray structures, but also agreed with the yet unreleased X-ray structures (*31*, *32*) (**Table S1**, **Table S2**), and exhibited physical parameters similar to real interfaces (**Table S3**). Specifically, X-ray structures of *C. thermophilum* Nup205 and Nup188 in complex with Nup93 as well as Nup93 in complex with Nup35 are consistent with the human ColabFold model (**Fig. S6**, **Table S2**). These structures represent proteins in complex with the respective SLiMs and form relatively small interfaces. However, also for larger subcomplexes we obtained structural models that convincingly fit our cryo-ET maps (**Fig. S14**). For example, the structure predicted for the so-called central hub of the Y-complex was consistent with the organization seen in fungal X-ray structures and explained additional density within the cryo-ET map specific to the human NPC. The new model of the Y-complex hub includes a previously unknown interaction between NUP96 and NUP160 (**Fig. S11**). ColabFold built a model of the NUP62 complex that has high structural similarity to the fungal homolog (**Table S2**) and fits the EM map with significant p-values **(Fig. S14**) even though no structural templates were used for modeling. We were also able to obtain a trimeric model of the small arm of the Y-complex comprising NUP85, SEH1, and NUP43. The model fits the EM map with significant p-values, confirming the known structure of NUP85-SEH1 interaction (**Table S2**), and revealing how NUP43 interacts with NUP85 (**Fig. S14**). In the case of the NUP214 complex, for which no structures are available, the ColabFold model is highly consistent with the rather uniquely shaped EM density (**Fig. S14**). The interface between NUP214 and NUP88 that is biochemically validated in (*31*) has also been predicted with high-structural similarity to the equivalent interface between homologs from *C. thermophilum* and budding yeast (**Table S2**).

With the cryo-EM maps and the repertoire of structural models of individual NUPs and their subcomplexes, we built a nearly complete model of the human NPC scaffold (**Fig. 1A, Supplementary Text, Materials and Methods**). The individual components are detailed in **Table S4**. We used the previous model (*19*, *21*) as a reference for modeling the scaffold of the constricted state and replaced all previously fitted domains with human AlphaFold and ColabFold models. We then added the remaining newly modeled subunits by systematic fitting to the EM map and refinement using Assembline (*37*) (**Fig. S14**, **Fig. S15**). In addition to fitting the models, we have added several disordered linkers that connect spatially separated domains and SLiMs within the NPC (**Fig. S16**, **Fig. S17**, **Fig. S18**,). We then built the model of the dilated state by fitting the constricted NPC model into the dilated NPC map and refining the fits using Assembline. The resulting models (**Fig. 1**) include 25 out of ~30 human NUPs (**Fig. S19**, **Table S4**). The protein regions explicitly included in the models account for 70 MDa of the molecular weight of the NPC (>90% of the scaffold molecular weight) in comparison to 16 NUPs and 35 Mda (46% of the scaffold) of the previous model, and largely account for the EM density observed in the constricted and dilated states.

This model yields novel insights into the organization of the human NPC (**Fig. 1**). Within the IR, NUP188 localizes to the outer, while NUP205 localizes to the inner subcomplexes, consistent with previous analysis in other species (*12*, *22*, *23*). We furthermore localized two copies of NUP205 in the CR and one in the NR (*33*), thus resolving previous ambiguities (*11*, *19*, *21*). Two previously undetected copies of NUP93 bridge the inner and outer Y-complexes both in the CR and NR, with an inherent C2 symmetry. This observation is consistent with biochemical experiments that initially identified interactors of NUP93 in the outer rings (*38*). The copy of NUP93 in the CR is located underneath the NUP358 complex, further corroborating a role of NUP358 in stabilizing the higher order structure (*21*). Yet another copy of NUP93 that is unique to the CR bridges the inner Y-complexes from two consecutive spokes. This is consistent with an additional copy of NUP205 in the CR as compared to the NR, since

NUP93 and NUP205 heterodimerize through a SLiM within the extended N-terminus of NUP93 (see below, **Fig. S10** (*31*, *32*)). The AI-based model of the NUP214 subcomplex interacts with NUP85 pointing towards the central channel, likely to optimally position the associated helicase crucial for mRNA export.

Linker NUPs fulfill dedicated roles of spatial organization within the higher-order assembly

Since the exact 3D trajectory of the linkers through the NPC scaffold was unknown, it remained difficult to understand their precise structural role beyond conceptualization as molecular glue. In our model, AI-based models of human NUP-SLiM subcomplexes allowed us to map the anchor points of the linkers to the scaffold. The AI-based models correctly recapitulated SLiM interactions known from X-ray structures but also revealed previously unknown human NUP-SLiM interactions. In comparison to the X-ray structures, the AI-based models more extensively covered the structured domains, thus reducing the length of the linkers and restricting their possible conformational freedom within our model.

To generate a connectivity map of the NUP linkers **(Fig. 2**), we used a multi-step procedure. Firstly, we calculated all geometrically possible connections. Next, we eliminated linker combinations that were too distant, caused steric clashes or were combinatorically impossible. Finally, we used Assembline to model the remaining linkers in explicit atomic representation for both constricted and dilated states (**Fig. 2**,**Fig. S16**, **Fig. S17**, **Fig. S18**, **Supplementary Text, Materials and Methods**).

The resulting connectivity map (**Fig. 2A**) reveals that the NUP35 linker regions bridge between neighboring spokes of the IR. In our model, the NUP35 dimer is positioned into previously unassigned EM density between spokes (**Fig. S14**) and each of the two copies reaches out with its SLiMs to NUP155 and NUP93 of the adjacent spokes (**Fig. 2A, Fig. S16**). The NUP35 dimer, which is critical during early NPC biogenesis (*39*), thus functions as an architectural organizer for the IR membrane coat in a horizontal direction along the membrane plane.

In contrast, the connectivity map demonstrates that the linkers at the N-terminus of NUP93 copies that connect anchor points at NUP205 or NUP188, and NUP62 complex in the IR, cannot reach across spokes, and thus connect subunits within a single IR sub-complex inside of the same spoke (**Fig. S17**). Thereby, the two outer copies bind to NUP188 while the two inner copies bind to NUP205. Thus, NUP93 acts as an architectural organizer within, but not across spokes.

In the CR and NR, the linkage between NUP93 and NUP205 is geometrically possible (**Fig. S18**), emphasizing the similar architectural design of the respective complexes. Thereby, the NUP93 SLiM that binds the NUP62 complex could also facilitate linkage to the homologous NUP214 complex, although the corresponding structural information is still missing. The duplication of NUP205 and NUP93 in the CR is suggestive of yet another copy of the NUP214 complex that is not well resolved in the cryo-EM map of the constricted state, and thus remains to be further investigated. This analysis is consistent with the biochemical analysis and structural modeling in the accompanying publication (*32*). In conclusion, the individual linker NUPs specialize in dedicated spatial organization functions responsible for distinct aspects of assembly and maintenance of the NPC scaffold architecture.

A transmembrane interaction hub organizes the interface between outer and inner rings

Several types of structural motifs associate the NPC scaffold with the membrane. The spatial distribution of the amphipathic helices and membrane-binding loops harbored by NUPs 160, 155, and 133 across the scaffold has been previously revealed (*19*). The analysis of the protein linkers in our new model allowed the mapping of an approximate location of the amphipathic

helices of NUP35. In addition to these motifs, the human NPC further contains three transmembrane NUPs, the precise location of which remains unknown.

Among the three human transmembrane proteins, NDC1 is the only one that is conserved across eukaryotes (*40*). NDC1 is known to interact with the poorly characterized scaffold NUP ALADIN (*41*, *42*). We confirmed this interaction using proximity labeling mass spectrometry with BirA-tagged ALADIN and identified NDC1 and NUP35 as the most prominently enriched interactors **(Fig. S20**). NDC1 is predicted to comprise six transmembrane helices followed by a cytosolic domain containing mainly a-helices, while ALADIN is predicted to have a b-propeller fold. Structures of both NDC1 and ALADIN, however, remain unknown. Using AlphaFold/Colabfold, we could model the structures both as monomers and a heterodimeric complex with high-confidence scores **(Fig. S5, Fig. S11**). Systematic fitting of the hetero-dimeric models to the EM map unambiguously identified two locations within the IR (**Fig. S14**). The EM density was not used as a restraint for modeling but matches the structure of the model and is consistent with the only patches of density spanning the bilayer (**Fig. 3, Fig. S14, Supplementary text**), therefore further validating the model. The two locations are C2-symmetrically equivalent across the nuclear envelope plane, thus assigning two copies of ALADIN and NDC1 per spoke, corroborating experimentally determined stoichiometry (*43*). The identified locations are close to the membrane-binding N-terminal domains of NUP155 and amphipathic helices of NUP35, which is also consistent with our proximity labeling data (**Fig. S11**) and previous functional analysis (*44*). The proximity labeling data prominently identified Gle1, which was previously shown to interact with NUP155 (*45*). Using AlphaFold we predicted an interaction between the C-terminus of NUP155 and the N-terminal SLiM in Gle1 with high confidence scores (**Fig. S20B**). We also predicted an ALADIN-binding SLiM in NUP35, located between the amphipathic helix and NUP155-binding SLiM (**Fig. 3, Fig. S11**). We therefore propose that together with NUP155, NUP35, ALADIN and NDC1 form a transmembrane interaction hub that anchors the inner membrane coat of the IR and orients the NUP155 connectors towards the outer rings. The central position of ALADIN within the NPC might explain functional consequences of mutations in ALADIN that are implicated in the AAA syndrome (*46–48*) and is consistent with the absence of ALADIN in fungi, which lack the NUP155 connectors (*24*).

We next examined the structure of NUP210, which contains a single-pass transmembrane helix and is the only NUP that primarily resides in the NE lumen. This NUP is composed of multiple immunoglobin-like domains and is thought to form a ring around the NPC within the NE lumen (*49*), but the structure of the ring has only been modeled in fungi. We used RoseTTAfold to model full-length NUP210 and obtained an elongated model with clearly defined interfaces between consecutive domains. Already as a rigid body, this model fitted the density sufficiently well to allow tracing NUP210 monomers in the cryo-EM map of the *Xenopus laevis* NPC, which is superbly resolved in the luminal region (*50*). Thus, we could assign eight copies of NUP210 per spoke (**Fig. S21**). Modeling of individual NUP210 fragments and inter-NUP210 interactions with AlphaFold/ColabFold (**Fig. S21**, **Table S1**) led to an overall model that explained the entire density of the luminal ring of the human and Xenopus cryo-EM maps, including the C-terminal transmembrane helix. The helix is long enough to span the NE and reaches the IR in the vicinity of the NDC1/ALADIN/NUP155/NUP35 transmembrane interaction hub. This location is consistent with known interactions of NUP210 homologs (*51*) and our proximity labeling data (**Fig. S20**). The model of the NUP210 ring also matches the luminal density visible in our *in cellulo* EM map, allowing us to model the ring in the context of both constricted and dilated NPC (Fig. 1, **Fig. S21**). To further confirm the NUP210 assignment in human cells, we deleted *NUP210* in HEK293 cells using CRISPR/Cas9 and analyzed the structure of the NPCs *in cellulo* using cryo-ET. The resulting map indeed showed

a lack of the luminal ring density (**Fig. S4**). The NPC scaffold in the resulting map appears overall unchanged, including its diameter, suggesting the NUP210 is not required for faithful NPC assembly.

Taken together, our model includes all known membrane-binding domains except for the cell-type specifically expressed POM121. The precise location of which within the NPC remains unknown, and neither AlphaFold or RoseTTAfold could build structural models with high confidence. The resulting membrane association map reveals that the membrane-binding b-propellers of the Y-complex (NUP160 and NUP133) and the IR (NUP155) are distributed as multiple pairs over the entire scaffold, whereby they follow a well-defined pattern. They form an overall Z-shaped outline within an individual spoke (**Fig. 3**). The NDC1/ALADIN/NUP155/NUP35 membrane-binding hub is situated at the interface of the IR with the outer rings and is distinct from the additional NUP155 pair at the NE symmetry plane. The membrane-binding motifs arrange in similar clusters in both the constricted and dilated state. Their relative arrangement is not changing uniformly during dilation, rather the relative distances within the spokes remain constant while the spacing of the spokes increases (**Fig. 3**).

The NPC scaffold prevents membrane constriction in the absence of membrane tension

It has been a longstanding view that the scaffold architecture of the NPC has evolved distinct membrane-binding motifs to stabilize the membrane at the fusion of the INM and ONM (*52*, *53*). To test the contribution of the scaffold architecture to membrane curvature, we used molecular dynamics (MD) simulations using a coarse-grained Martini force field (*54*, *55*). We first simulated a double-membrane pore without proteins with an initial pore diameter and membrane spacing as seen in the constricted NPC cryo-ET map. We found that the pore constricts during 1 μs simulations and stabilizes once the radii of the INM/ONM and the NE hole are the same, such that the mean curvature nearly vanishes (**Fig. 4A, Fig. S23, Movie1**). This is in line with Helfrich membrane elastic theory, which predicts a catenoid-like pore shape with equal radii of curvature at the pore center as the lowest energy structure, and an energetic cost of about 200 k_BT to widen the pore (**Supplementary Text**). Interestingly, the opening of the relaxed double-membrane pore is considerably smaller than even the most constricted NPC conformation. The NPC scaffold thus keeps the pore wider than it would be without the scaffold.

These findings would predict that the nuclear membranes push against the NPC scaffold even in the most constricted state, which is in agreement with experimental data (*23*). To examine the effects of this tension on the NPC, we generated an NPC scaffold model with explicit membrane and water as a solvent, and ran 1 μs MD simulations (**Fig. 4B, Fig. S23, Fig. S24, Fig. S25, Fig. S26, Fig. S27, Movies 2-3**). We omitted the luminal NUP210 and focused our analysis on the architecturally important NR, CR and IR. In these simulations, we found that the membrane pore wrapped tightly around the IR plane, adopting an octagonal shape (**Fig. 4C**). Similarly tight wrapping and octagonal shapes have been seen in the previous EM analyses of NPCs (*21*, *56*, *57*). We also observed that the diameter of the NPC scaffold constricted by about 9% (**Fig. 4B**). We attribute this tightening primarily to mechanical tension in the pore widened beyond the catenoid shape. First, we observed similar contraction in simulations with rescaled protein-protein interactions (**Fig. S27**). Second, by applying lateral tension on the double membrane, we could maintain the pore width or widen it (**Fig. 4B, Fig. S23B**). At even higher tension, the membrane spontaneously detached from the NPC scaffold (**Fig. S28**). Taken together, our data support a model in which the role of the NPC scaffold is not to stabilize the membrane fusion *per se*, but rather to widen the diameter of the membrane hole without necessitating a wider envelope.

Discussion

We have built a 70MDa model of the human NPC scaffold in the constricted state (smaller diameter) as adopted in purified nuclear envelopes and in the dilated state as adopted in cells, whereby recent work in fungi has identified constricted NPCs inside of cells under specific physiological conditions (*23*). Our model includes multiple previously unassigned domains and proteins, resolves long-standing ambiguities in alternative NUP assignments, lays out a connectivity map of the protein linkers across the NPC scaffold, maps out the membrane-anchoring motifs, and provides a high-quality basis for further investigations of NPC dynamics and function. Our analysis demonstrates that coarse-grained dynamic model is of sufficient quality for molecular simulations, which in future could quantitively and predictively describe how the NPC interplays with the nuclear membrane and how it responds to mechanical challenges. The model also provides a more accurate starting point for simulations of nucleocytoplasmic transport by providing the native constraints on the diameter and a more precise mapping of the positions where the FG tails emanate from the scaffold (**Fig. S22**).

How an intricate structure consisting of ~1000 components can be faithfully assembled in the crowded cellular environment is a very intriguing question. Our connectivity map captures the 3D trajectory of linker NUPs through the assembled scaffold. Taken together with previous analysis of NPC assembly (*9–13*, *19*), it suggests that the linker NUPs facilitate dedicated spatial organization functions. The connections of NUP93 within individual IR complexes and to the NUP214-complex suggest a role in ensuring isostoichiometric assembly. This finding is consistent with the recent analysis of early NPC biogenesis, suggesting that NUP93 associates isostoichiometrically with the NUP62-complex already during translation in the cytosol (*58*). Thus, the stoichiometric assembly of the NUP62 subcomplex together with NUP205/188-NUP93 heterodimer is likely pre-assembled away from sites of NPC biogenesis, explaining the importance of the linker for intra-subcomplex interactions. How the spokes form a C2 symmetric interface at the NE plane remains to be addressed.

In the IR membrane coat, multiple interactions converge into an intriguing transmembrane interaction hub. We propose that its core is formed by ALADIN-NDC1 heterodimer at the interface between the outer and inner rings. This transmembrane interaction hub is likely a spatial organizer for two proximate copies of NUP155 within the same spoke that point towards the outer rings and IR, respectively. ALADIN-NDC1 likely further associates with NUP210, which arches between spokes in the NE lumen. The hub also binds NUP35, which connects to NUP155 copies of neighboring spokes, thus facilitating the horizontal, cylindrical oligomerization. Since NUP35 associates with NUP155 early during NPC assembly (*39*), its dimerizing domain appears critical to scaffold its flexible linkers towards neighboring spokes within IR membrane coat.

The often-emphasized notion that ‘NPCs fuse the INM and ONM’ or that they ‘stabilize the fusion of the INM and ONM’ is not necessarily supported by our analysis. Our simulations suggest that the membrane fusion topology *per se* is stable under certain conditions, relaxing toward a catenoid shape with zero membrane bending energy. Indeed, some species maintain the fusion topology in the absence of NPCs, e.g. during semi-closed mitosis in *Drosophila melanogaster* (*59*). Our analysis instead suggests that NPCs stabilize a pore that is wider than in the relaxed, tensionless double membrane hole. This notion agrees with the ultrastructural analysis of post-mitotic NPC assembly, which has revealed that NE holes are formed at small diameters and dilate once NPC subcomplexes are recruited (*60*). These data argue that the membrane shape defines the outline of the NPC scaffold and not *vice versa*.

We use AI-based structure prediction programs AlphaFold and RoseTTAfold to model all atomic structures that were used for fitting to the EM maps. Although X-ray and cryo-EM structures were used for validation, no experimental atomic structures were directly

incorporated into the model. Predicted atomic structures traditionally exhibited various inaccuracies, limiting their usage for detailed near-atomic model building in low-resolution EM maps. However, AlphaFold and RoseTTAfold have recently demonstrated unprecedented accuracy in predicting structures of monomeric proteins (*26*, *27*, *61–66*) and complexes (*34*, *36*, *62*, *67*). They accurately assess their confidence at the level of individual residues and inter-domain contacts (*26*, *27*, *62*). Indeed, we could successfully validate our models by comparing them to unpublished crystal structures, cryo-EM maps, and biochemical data. The resulting model of the NPC scaffold is almost complete and exhibits near-atomic level precision at several interfaces. The model also contains several peripheral NUPs, e.g. parts of the NUP214 and NUP358 complexes. Projection of the locally estimated accuracy into an asymmetric unit of the NPC reveals that the structured regions are generally modeled with good confidence, while linkers and peripheral loops are less well defined (**Fig. S29**). Although the entire EM density for those peripheral NUPs is unlikely to be resolved in the near future due to their flexibility, the complete model of the human NPC could be in reach by integrating data from complementary techniques that can address flexible proteins, such as super-resolution microscopy, FRET, and site-specific labeling (*18*).

Thanks to *in situ* and *in cellulo* cryo-ET and powerful AI-based prediction (*26*, *27*), intricate structures such as the NPC can be now modeled. Not all subunit or domain combinations that we attempted to model with AI-based structure prediction led to structural models that were consistent with complementary data, emphasizing that experimental structure determination will still be required in the future for cases in which *a priori* knowledge remains sparse. However, even if AI-based modeling does not yield high confidence results, the models can still serve as tools for hypothesis generation and subsequent experimental validation.

Materials and Methods

Mammalian cell cultivation and subcellular fractionation

Modified human embryonic kidney cells 293 (HEK Flp-In™ T-Rex™) 293 Cell Line, Life Technologies) designed for rapid generation of stably transfected cell lines with a tetracycline-inducible expression system were used as parental cells. The NUP210 CRISPR-knockout line (HEK NUP210D) has been previously described (*68*). In general, all cells were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 5g/L glucose and 10% heat inactivated fetal bovine serum (FBS, Sigma-Aldrich). HeLa Kyoto cell line was maintained in Dulbecco's modified Eagle medium (DMEM) medium containing 1 g/L glucose supplemented with 2 mM L-glutamine. Cells grown close to confluency (~90%) were trypsinized with 0.25% trypsin containing EDTA (Life Technologies) and passaged for further growth.

For the preparation of nuclear envelopes, HeLa cells were cultured and subjected to subcellular fraction as described before (*8*, *43*).

Grid preparation

Grids with HeLa nuclear envelopes were prepared exactly as described in (*21*). For the *in cellulo* work, Au200 R2/1 SiO_2 grids (Quantifoil Micro Tools GmbH) were glow discharged on both sides and sterilized under UV light. In a 6-well cell culture dish either 250,000 cells/well (for HeLa) or 400'000 cells/well (HEK293) were pipetted onto the grids pre-wetted with DMEM medium. Cells were left to settle and attach to the grids for 4h at 37°C in 5% CO2. Subsequently the HeLa or HEK293 grids were plunge frozen with a Leica EM GP plunger with set chamber environment to 99% humidity and 37°C. Grids were blotted from the backside for 2 sec and plunged in liquid ethane-propane mix (37% and 63%) at ~ -195°C. HEK NUP210Δ grids were washed once with PBS containing 8% dextran (35.45kDa) and blotted for 3 sec prior to plunge freezing in -186°C liquid ethane.

Cryo-FIB milling and Data acquisition

Plunge-frozen sample grids were FIB-milled on an Aquilos FIB-SEM (ThermoFisher Scientific) as described before (*22*, *23*). In brief, samples were coated with inorganic platinum (Pt-sputtering). Subsequently, a protective layer of organometallic platinum was deposited for ~ 20 sec using the gas injection system. Cells were then stepwise milled at a 20° angle to a final thickness of around 200-250 nm using decreasing ion-beam currents of 1 nA to 50 pA. A final round of Pt-sputtering was applied before unloading the sample.

Cryo-electron tomography and subtomogram averaging of the human NPC from nuclear envelopes

Tilt series were collected with SerialEM as described by Kosinski et al. (*19*). The angular coverage of the tilts spanned from -60 to +60 degrees. 10 8K x 8K frames, per tilt, were collected in the super-resolution mode on a K2 direct electron detector (Gatan Inc.) equipped with an BioQuantum Imaging Filter (GIF). An average total dose of 120 e-/Å2 per tomogram was used. 516 new tilt-series were collected and combined with 101 tilt-series reported by Kosinski et al. (*19*) leading to a total of 617 tilt-series. Incomplete tilt-series (missing more than 7 tilts in either direction or terminated due to autofocusing error because of the edge of the grid bar) were discarded. CTF was determined using CTFFind4 (*69*). Tilt-series with large discrepancies in the two defocus values estimated by CTFFind4 were also removed. This resulted in a total of 554 tilt-series. Tilt-series were manually aligned by tracking gold fiducials in IMOD (*70*). Tilt-series were filtered according to accumulated exposure based on parameters described by Kosinski et al. (*19*). Tilt-series were reconstructed with 3D CTF correction using NovaCTF (*71*).

7711 subtomograms containing individual NPCs were extracted, which correspond to 61688 asymmetric units. The pixel size at the specimen level was 3.37 Å. Tomograms were binned by Fourier cropping 2x (bin2), 4x (bin4) and 8x (bin8) and subtomograms were extracted at each level of binning corresponding to a pixel size of 6.74, 13.48 and 26.96 respectively. Subtomogram averaging was performed on a whole pore level with bin8 dataset. Subsequently, asymmetric units were extracted from the aligned pores and averaged as described in Kosinski et al. (*19*). Subunits with the center outside of tomogram boundaries were excluded from further processing. The cytoplasmic (CR), inner (IR) and nuclear (NR) ring were processed separately, i.e. the positions of subunit centers were moved to be in the center of each ring, resulting in three different set of subtomograms. Rings with the center outside of tomogram boundaries were excluded from further processing. The subtomograms were iteratively aligned first on bin4, then on bin2 level and the final alignment was refined on bin1 level. The complete subtomogram averaging and alignment was performed using novaSTA (*72*), the masks necessary for the alignment were created in Dynamo (*73*) and Relion (*74*).

After the bin4 alignment, the quality of each subtomogram was assessed using geometrically-restraint classification that is based on the expected geometrical shape of a complete ring. More precisely, the subunits corresponding to one NPC ring should still be part of a ring after the alignment. For each ring and each subunit, the angular distance of its normal vector to all other normal vectors of subunits within the same ring was computed. Subsequently, the distances were averaged (for each NPC separately) and for each subunit the deviation from the normal vector from the average was computed. The same was done for so-called in-plane vectors, i.e. vectors describing the direction from the NPC center to the center of a subunit. The expected/ideal angular distance for the normal vectors is zero while for the in-plane vectors it is 45°. This analysis was performed only on the rings with at least three subunits that were retained from the initial sub-tomogram averaging runs. The rings with less than three subunits were removed from further processing. The computed deviations were used to identify poorly

aligned ring subunits in order to remove them. For CR and SR all subunits with the normal vector deviation greater than 30° were removed. The deviation of in-plane angles from expected 45° greater than 15° or CR and 10° for SR was used as threshold for additional removal of ring subunits. The threshold values were determined empirically. For CR and SR, the number of subunits left for processing after geometrical cleaning were 31774 and 35281 respectively. The geometrical-restrain classification was added to the publicly available novaSTA package (*72*). The final subunit cleaning to remove poor quality subunits was performed on the final average using the constrained cross-correlation (CCC) value which was computed between each subtomogram and the reference during the last iteration of alignment. Subtomograms with the worst CCC values were subsequently removed in batches of 1000 as long as the resolution improved. The number of subunits/subtomograms contributing towards the final structure of the CR and SR ring were 21604 and 30000, respectively.

In contrast to CR and IR, adding additional tilt-series followed by geometric cleaning procedure on NR did not yield any significant improvement in comparison to the dataset reported in Kosinski et.al (*19*). Thus, the original map of the NR was used for the presented analysis. The map was created using steps described in Kosinski et al. (*19*) and the total number of particles contributing to the final average were 11112.

Cryo-electron tomography and subtomogram averaging of human NPC *in cellulo*

Data acquisition was performed on a Titan Krios G2 (for HeLa) or G4 (for HEK) (ThermoFisher), operating at 300kV and equipped with Gatan K2 Summit direct electron detector and energy filter as described before (*23*). In brief, tilt series were acquired in dose-fractionation mode at 4k x 4k resolution with a nominal pixel size of 3.37 (HeLa) or 3.45 Å (HEK) using a automated dose-symmetric acquisition scheme (*75*) starting at a given pre-tilt corresponding to the tilt of the FIB-milled lamellae (typically +/- 13°). Tilt series were acquired with a tilt increment of 3° and a tilt range interval of -50°/+50°, and a total dose per tomogram of 120-150 e/Å^2.

Tilt series preprocessing and tomogram reconstruction was performed as described previously (*22*, *23*). Subtomogram averaging was performed as described before (*22*, *23*) . In brief, for the HeLa control dataset 53 NPCs were extracted from 13 tomograms. For the HEK dataset, 30 control and 43 NUP210Δ NPCs were extracted form 8 control and 14 NUP210Δ tomograms respectively. Whole pores were aligned using bin8 and bin4 subtomograms with imposed 8-fold symmetry. Upon convergence 280, 150 and 222 subunits were extracted from the control HeLa, control HEK and HEK NUP210Δ dataset respectively and the cytoplasmic, inner and nuclear ring (CR, IR and NR) subunits were further refined independently using bin4 subtomograms. The individual ring subunits were refined without splitting the data into independent half sets to a final resolution of <54 Å (NR) and <48 Å (CR and IR) as estimated by Fourier Shell Correlation (FSC) using the 0.5 criterion for the HEK datasets. For the HeLa dataset gold-standard criteria was used to calculate the FSC, which resulted in final resolution of 45 Å (CR and IR) and 53 Å for the NR.

Structure determination of human NUP155

The gene encoding human NUP155 (UniProt ID: O75694) was synthesized by GeneArt (Life Technologies) and cloned into a modified pFastBac vector, with a His_6 tag and an enhanced GFP tag followed by an HRV 3C protease site at the N terminus, and a Strep tag at the C terminus. The predicted membrane binding loop (residues 260–273) was deleted to improve the protein stability. The resulting construct was expressed in Sf21 insect cells using the Bac-to-Bac baculovirus expression system (Thermo Fisher Scientific). Sf21 cells were cultured in Sf900III medium (Gibco) at 27°C, and infected at a density of 1–2 × 10^6 cells ml^{-1}. After 48

h of incubation, cells were collected by centrifugation (3,000g, 10 min, 27°C) and the pellets were stored at −80°C until purification.

For the purification, frozen cell pellet from 100 ml culture was resuspended in 10 ml of buffer containing 20 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM DTT, and 0.1 mM phenylmethylsulfonyl fluoride, and disrupted by sonication for 5 min using Branson sonifier 250. After removing cell debris by centrifugation (3,000g, 10 min, 4°C), the supernatant was mixed with 500 µl of Strep-Tactin Sepharose resin (IBA Lifesciences) and incubated at 4°C for 30 min. The resin was washed with 8 ml of buffer containing 20 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM DTT, and the bound sample was eluted with the same buffer supplemented with 5 mM biotin. The eluted fractions were concentrated with an Amicon Ultra 0.5 ml centrifugal filter (100 kDa molecular weight cut-off, Millipore), mixed with Turbo-3C protease (Sigma-Aldrich), and incubated at 4°C overnight. The sample was then ultracentrifuged (71,680g, 15 min, 4°C), and loaded onto Superose 6 Increase 3.2/300 equilibrated with 20 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM DTT. The peak fractions were aliquoted and stored at −80°C until use.

For the preparation of EM grids, the protein concentration was adjusted to 0.4 mg ml^{-1} in 20 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM DTT, and 0.001% dodecyl maltoside. The diluted sample was then applied onto a freshly glow-discharged UltrAuFoil R0.6/1.0 gold grids (300 mesh, Quantifoil), blotted for 6 s at 4°C in 100% humidity, and plunge-frozen in liquid ethane using Vitrobot Mark IV. Cryo-EM data were collected on a Titan Krios G4 microscope (Thermo Scientific) operated at 300 kV, equipped with a E-CFEG, a Falon 4 direct electron detector (Thermo Scientific), and a Selectris X energy filter (Thermo Scientific) operated with a slit width of 10 eV. Automated data acquisition was performed with the EPU software at a nominal magnification of ×165,000, corresponding to a pixel size of 0.730 Å/pix. Movies were acquired at a dose rate of 4.95 electrons per pixel per second, and a total dose of 50 electrons per $Å^2$, resulting in EER movies consisting of 1,407 frames. In total, 6,430 movies were acquired with a defocus range of −1.0 to −2.0 µm. Acquired images were first processed in cryoSPARC (*76*), and selected particles were further processed in RELION-3.1 using csparc2star.py script in UCSF pyem (*77*) for transfer of particles. To analyze the conformational flexibility, multi-body refinement (*78*) was performed on the consensus map. The details of data processing are summarized in **Fig. S8**.

The model of human NUP155 was manually built in *Coot* (*79*), using the crystal structure of Nup170, a homologue of NUP155, from *Chaetomium thermophilum* (PDB ID: 5HAX) as a starting model. Secondary structure prediction from PSIPRED (*80*) and multiple sequence alignment were used to facilitate the model building. The model was iteratively refined using phenix.real_space_refine (*81*). Figures were prepared with UCSF Chimera (*82*), UCSF ChimeraX (*83*), and CueMol (http://www.cuemol.org/).

Proximity labeling using BioID

BioID analysis of ALADIN was done as previously described (*84*). In brief, ALADIN was BirA tagged and over-expressed in Hek293 Flp-In Trex cells. Quantitative mass spectrometry was done in four biological replicates and in comparison to control cells expressing BirA tagged NLS-NES-Dendra that resides within the central channel.

Structural modeling of NUPs and NPC subcomplexes

The structures of all individual NUPs and selected subcomplexes were modeled using AlphaFold (*26*) or downloaded from AlphaFold Database (*61*). The models of monomeric proteins (NUP155, NUP133, NUP107, NUP93, NUP205, NUP188, NUP160, NUP358 and ELYS), were download from from AlphaFold Database (*61*). To model subcomplexes or their

parts around the interfaces (NUP62-NUP54-NUP58, NUP205-NUP93, NUP188-NUP93, NUP155-NUP35, NUP93-NUP35, NDC1-ALADIN, NUP35 homo-dimer, NUP85-SEH1-NUP43, NUP160-NUP96-SEC13, NUP160-NUP37, NUP133-NUP107, NUP96-NUP107, and NUP160-NUP96-NUP85, NUP214-NUP62-NUP88, NUP88-NUP98), we used the AlphaFold version modified for modeling complexes, available through ColabFold (*35*) with all parameters set to default except max_recycles parameter set to between 12 and 48 depending on the subcomplex. For NUP210, we first built the initial full-length using RoseTTAfold (*27*), as AlphaFold did not provide a full-length model fitting well into the EM density map. After fitting the model into the EM maps as a rigid body (see below), we used AlphaFold to model successive monomeric and homodimeric fragments of NUP210, superposed them onto the fitted RoseTTAfold model, and refined the fits. The quality of the AlphaFold models was first assessed by the scores provided by the authors—the predicted local-distance difference test (pLDDT), which predicts the local accuracy, and Predicted Aligned Error, which assesses the packing between domains and protein chains. In addition, we validated the models by comparing to structures not used for modeling, yet unpublished structures provided in the accompanying paper, fits to the cryo-ET maps, and previously published biochemical data (**Fig. S5**, **Fig. S6**, **Fig. S7**, **Fig. S10**, **Fig. S11**, **Fig. S14**).

Systematic fitting of atomic structures to cryo-ET maps

We used the previously published procedure for systematic fitting (*8*, *19*, *21*, *22*, *25*, *37*, *85*) to both locate the atomic structures in the cryo-ET maps and validate the AlphaFold models. Prior to fitting, all the high-resolution structures were filtered to 10-15 Å. The resulting simulated model maps were subsequently fitted into individual ring segments of cryo-ET maps by global fitting as implemented in UCSF Chimera (*82*) using scripts within Assembline (*37*). The maps used for fitting did not include nuclear envelope density in order to eliminate the possibility of fits significantly overlapping with the membrane. All fitting runs were performed using 100,000 random initial placements with the requirement of at least 30-60% (depending on the size of the structure) of the simulated model map to be covered by the cryo-ET density envelope defined at a low threshold. For each fitted model, this procedure resulted in around 1000-20,000 fits with non-redundant conformations upon clustering. The cross-correlation about the mean (cam score, equivalent to Pearson correlation) score from UCSF Chimera (*82*) was used as a fitting metric for each atomic structure, similarly to our previously published works. The statistical significance of every fitted model was evaluated as a p-value derived from the cam scores. The calculation of p-values was performed by first transforming the cross-correlation scores to z-scores (Fisher's z-transform) and centering, from which subsequently two-sided p-values were computed using standard deviation derived from an empirical null distribution (based on all obtained non-redundant fits and fitted using fdrtool (*86*) R-package). Finally, the p-values were corrected for multiple testing with Benjamini-Hochberg procedure (*87*).

Modeling of the human NPC scaffold

To assemble the models of the entire NPC scaffold based on the constricted and dilated cryo-ET maps we used our integrative modeling software Assembline (*37*), which is based on Integrative Modeling Platform (IMP) (*88*) version 2.15 and Python Modeling Interface (PMI) (*89*). First, we built the model of the constricted NPC owing to its higher resolution. The AlphaFold models of NUP domains and subcomplexes already present in our previous human NPC models (*19*, *21*), were placed in the map by superposing them onto the published models. The remaining domains and subcomplexes added in this work (NUP358, NUP35, NUP93 in the outer rings, NDC1 and ALADIN, NUP214 complex), were placed using systematic fitting (as above) and global optimization procedure of Assembline. In addition to using models of subcomplexes as rigid bodies for fitting, several inter-subunit interfaces were restrained by

elastic distance network derived from ColabFold models overlapping with and bridging already fitted models. During the refinement, the structures were used as rigid-bodies and simultaneously represented at two resolutions: in Cα-only representation and a coarse-grained representation, in which each 10-residue stretch was converted to a bead. The 10-residue bead representation was used for all restraints to increase computational efficiency except for the domain connectivity restraints, for which the Cα-only representation was used. The flexible protein linkers between the domains were added as chains of one-residue beads. The entire structure was optimized using the refinement step of Assembline to optimize the fit to the map, minimize steric clashes, and ensure connectivity of the protein linkers. The scoring function for the refinement comprised the EM fit restraint, clash score (SoftSpherePairScore of IMP), connectivity distance between domains neighboring in sequence, a term preventing overlap of the protein mass with the nuclear envelope, a restraint promoting the membrane-binding loops of NUP133, NUP160, NUP155 to interact with the envelope, implemented using MapDistanceTransform of IMP (predicted by similarity to known or predicted ALPS motifs *X. laevis* and *S. cerevisiae* homologs (*6*, *21*, *90*), and elastic network restraints derived from the subcomplexes modeled with AlphaFold/ColabFold. The final atomic structures were generated based on the refinement models by back-mapping the coarse-grained representation to the original AlphaFold atomic models. The conformation of the linkers was further optimized using Modeller (*91*) and Isolde (*92*). The stereochemistry of the final model was optimized using steepest descent minimization in GROMACS (*93*).

The model of the dilated NPC was built by fitting the asymmetric units of the individual cytoplasmic, inner and nuclear rings of the constricted NPC model to the dilated cryo-ET maps and refining the fits with Assembline. The refinement procedure was performed as above.

To calculate the percentage of the molecular weight of the full NPC and the NPC scaffold covered by the new and the old models, we defined the full NPC as composed of the following 32 NUPs, with the stoichiometry indicated in the parentheses: NUP160 (32), NUP96 (32), NUP85 (32), SEH1 (32), SEC13 (32), NUP107 (32), NUP133 (32), NUP358 (40), NUP43 (32), ELYS (16), NUP37 (32), NUP188 (16), NUP205 (40), NUP155 (48), NUP93 (56), NUP35 (32), NUP62 (48), NUP54 (32), NUP58 (32), NUP88 (16), NUP214 (16), NUP98 (48), NDC1 (16), NUP210 (64), and ALADIN (16), POM121 (32), TPR (32), NUP153 (32), NUP50 (16), CG1 (8), DDX19 (16), GLE1 (8). The scaffold NPC was defined as composed of 25 NUPs: NUP160, NUP96, NUP85, SEH1, SEC13, NUP107, NUP133, NUP358, NUP43, ELYS, NUP37, NUP188, NUP205, NUP155, NUP93, NUP35, NUP62, NUP54, NUP58, NUP88, NUP214, NUP98, NDC1, NUP210, and ALADIN. The stoichiometry for the scaffold was the same as for the full NPC with exception of NUP214 complex for which only one copy was counted, as the second copy is not clearly visible in the EM density. Note that for some nucleoporins like NUP98 or POM121, the exact stoichiometry is still uncertain. The coiled-coil domains of the peripheral NUPs of the NUP214 complex and the a-solenoid domain of NUP358 were included in the scaffold. The FG regions were excluded. These definitions resulted in the molecular weight of 119 MDa for the full NPC and 76 MDa for the scaffold. The scaffold diameters were described by two distances between the opposite spokes: the membrane-to-membrane distance and the distance between ferredoxin-like domains of NUP54 at the residue 220.

Figures were produced using UCSF ChimeraX (*83*).

Molecular dynamics (MD) simulations

We performed MD simulations of half-toroidal membrane pores in isolation and including the hNPC scaffold. In the following, we describe the setup of the simulation models, the relevant MD parameters, and the analysis of the MD trajectories.

Membrane model. First, a coarse-grained POPC lipid bilayer patch with the size of 30×30 nm^2 was generated using *insane.py* (*94*, *95*). The bilayer was placed in a periodic simulation box, solvated on both sides, energy minimized, and simulated for 100 ns using standard MD parameters, as noted below.

Then, half-toroidal membrane pores were constructed with the BUMpy software (*94*) using this initial flat bilayer as membrane input. The following command line flags were used when running bumpy: -s double_bilayer_cylinder -z 10 -g l_cylinder:10 r_cylinder:430 r_junction:120 l_flat:1400 (see bumpy documentation). The resulting membranes coincided reasonably with the cryo-ET density of the double membrane pore and allowed us to place the membrane anchoring motifs of the NPC model into the membrane.

Two carbon nanotube porins (CNTPs) were inserted into the membrane in the corners of the simulation box distant from the NPC (see, e.g., Fig. S19A). The CNTPs with a length of 3.6 nm and a diameter of 14.7 nm enable water transfer in and out of the otherwise disconnected luminal volume, as is required for membrane-mechanical equilibration. Without CNTPs, the luminal volume would be effectively fixed and, as a result, changes in the membrane shape during MD simulations without NPC scaffold would induce artifactual membrane buckling. The CNTPs were built according to previous work (*94–96*). The outermost carbon rings at either CNTP end consisted of polar SNda beads for stable membrane embedding (*94*). To stiffen the wide CNTPs, the improper dihedral force constant was increased to 1000 kJ mol^{-1} rad^{-2}. The CNTP parameters were otherwise set as previously reported (*96*, *97*)**.** The code to generate CNTP models and the parameters for simulations are available at: https://github.com/bio-phys/cnt-martini (*96*). The CNTPs were embedded in the flat patch of the NPC membranes away from the NPC. Lipids within 8 Å of the CNTPs or inside their circumference were removed.

NPC scaffold model. The MD simulation model of the NPC included the entire scaffold (see **Table S4** and **Fig. S19** for a summary of the hNPC simulation model) except the disordered FG-NUP C- and N-terminal tails. For simplicity and to limit the system size, we also excluded the NUP210 glycoprotein in the nuclear envelope lumen. Otherwise, the models were complete as described above.

Each protein chain was coarse-grained individually using martinize.py as follows. All chain termini were uncharged and otherwise default protonation states were used. Secondary structure restraints were assigned according to DSSP (*98*). The tertiary structure of each protein chain was maintained by an elastic network using the recommended default settings with a cutoff R_c = 0.9 nm and a force constant k = 500 kJ mol^{-1} nm^{-2}. For each protein chain, the ElNeDyn2.2 protein force field was used in conjunction with the Martini 2.2 force field (*54*, *55*). Simulations were performed with the default protein-protein interaction (α = 1.0; results shown in Supplementary Materials), and with protein-protein interactions scaled relative to protein-solvent interactions with α = 0.7 (*99*) (results shown in the main text) to correct for the effect of reportedly overestimated non-bonded interactions (*100*). This procedure used the *martinize.py* script and was wrapped in custom python code to automatically generate the structures for each protein chain with the aforementioned parameters. To enable easier handling of the large number of protein chains, each protein chain was assigned a unique segid. Importantly, with this MD simulation model all protein-protein interactions between distinct chains could dissociate and new interactions could form in principle, and the structure of linker regions could relax.

All individually coarse-grained protein chains were then merged into one PDB structure file. The resulting coarse-grained NPC scaffold model was centered within the half-toroidal

membrane pore model containing the CNTPs described above. Any lipids within 8 Å of any bead of the scaffold proteins in the initial assembly were removed.

Solvation. All systems were solvated with coarse-grained water containing 10% anti-freeze WF particles and Na^+ ions to neutralize the system using standard GROMACS tools. All systems simulated in this study are listed in **Table S5**.

MD simulations. All molecular dynamics simulations were performed using the GROMACS software package and the coarse-grained Martini force field v2.2 (*54*, *93*, *101*)**.** Each system was first steepest-descent energy minimized using a soft-core potential to remove steric clashes in the initial model. The systems were then equilibrated in an NPT ensemble with semisotropic pressure coupling first for 2.5 ns with a 5 fs timestep and then for 100 ns with a 15 fs timestep with position restraints on the protein backbone beads with a force constant of 1000 kJ mol^{-1} nm^{-2}, maintaining a temperature of 310 K and pressure of 1 bar using the Berendsen barostat and velocity rescaling thermostat (*102*, *103*). Characteristic coupling times of 12 and 1 ps were used, respectively. During production simulations the Parrinello-Rahman barostat was used (*104*)**.**

The Verlet neighbor search algorithm was used to update the neighbor list, with the length and update frequency being automatically determined. Lennard-Jones and Coulomb forces were cut off at 1.1 nm with the potential shifted to zero using the Verlet-shift potential modifier. A 15 fs timestep was used in all production simulations. Production simulations were performed for 1 μs each.

Membrane tension. To apply lateral tension on the double-membrane structure, an anisotropic pressure tensor was used with an out-of-plane pressure of $P_\perp = 2\Delta P/3$ and an in-plane pressure of $P_\parallel = p - \Delta P/3$, with $p = 1$ bar. This results in a traceless lateral strain $S = diag(-\Delta P/3, -\Delta P/3, 2\Delta P/3)$ where $\Delta P \equiv P_\perp - P_\parallel$.The resulting tension on the double-membrane system is $\sigma = (P_\perp - P_\parallel)L_z = \Delta P L_z$with L_z the box height. To allow for gradual equilibration under tension, ΔP was increased in steps of 1 bar until reaching the target value (see Table S1).

Analysis of MD simulations. Images and movies were generated using VMD (*105*) and time series were analyzed using the MDAnalysis library (*106*)**.**

To monitor conformational changes, we calculated the root-mean-square distance (RMSD) from the starting structure using the qcprot RMSD alignment algorithm implemented in MDAnalysis (*106*). The RMSD was calculated every 1.5 ns for the backbone (BB) beads with respect to the rigid-body aligned initial structure. In addition to the individual protein chains, we analyzed in this way the b-propeller present in the three nucleoporins NUP133 (residues 1-480), NUP155 (residues 1-500), and NUP160 (residues 1-500); as well as the respective alpha solenoid domain NUP133 (residues 500-end), and NUP155, and NUP160 (residues 507-end); and each of the eight spokes as a whole. In the RMSD analysis, averages and standard deviations were calculated across the eight spokes or across equivalent protein copies in the NPC scaffold, respectively.

During the MD simulations, the diameter of the NPC membrane pore was determined by least-square fitting the center and radius of a circle in the xy-plane to the membrane center (C4A and C4B lipid beads). The fit was performed at the narrowest region of the half-toroidal membrane pore.

Possible limitations. We note that the time scale currently accessible to MD simulations is too short to fully recapitulate the complete NPC dilation and constriction processes, including the large-scale NPC structural rearrangements. We also note that the elastic network on proteins of the Martini model restricts internal conformational changes, which might be required for

larger-scale NPC dilation. The coarse-grained interaction model may also weaken some protein-protein interactions and strengthen others. Finally, we expect that the missing FG mesh in the MD model contributes to the compaction of the NPC scaffold seen in the MD simulations, acting on top of the mechanical tension in the widened double-membrane pore (Supplementary Text).

References and Notes

1. A. Ibarra, M. W. Hetzer, Nuclear pore proteins and the control of genome functions. *Genes Dev.* **29** (2015), pp. 337–349.

2. M. Raices, M. A. D'Angelo, Nuclear pore complexes and regulation of gene expression. *Curr. Opin. Cell Biol.* **46** (2017), pp. 26–32.

3. C. Strambio-De-Castillia, M. Niepel, M. P. Rout, The nuclear pore complex: Bridging nuclear transport and gene regulation. *Nat. Rev. Mol. Cell Biol.* **11** (2010), pp. 490–501.

4. P. De Magistris, W. Antonin, The Dynamic Nature of the Nuclear Envelope. *Curr. Biol.* **28** (2018), pp. R487–R497.

5. B. Hampoelz, A. Andres-Pons, P. Kastritis, M. Beck, Structure and Assembly of the Nuclear Pore Complex. *Annu. Rev. Biophys.* **48**, 515–536 (2019).

6. D. H. Lin, A. Hoelz, The structure of the nuclear pore complex (An Update). *Annu. Rev. Biochem.* **88**, 725–783 (2019).

7. S. Siniossoglou, M. Lutzmann, H. Santos-Rosa, K. Leonard, S. Mueller, U. Aebi, E. Hurt, Structure and assembly of the Nup84p complex. *J. Cell Biol.* **149**, 41–53 (2000).

8. K. H. Bui, A. von Appen, A. L. DiGuilio, A. Ori, L. Sparks, M.-T. Mackmull, T. Bock, W. Hagen, A. Andrés-Pons, J. S. Glavy, M. Beck, Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. *Cell*. **155**, 1233–1243 (2013).

9. M. Beck, E. Hurt, The nuclear pore complex: Understanding its function through structural insight. *Nat. Rev. Mol. Cell Biol.* **18**, 73–89 (2017).

10. 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**, 1–10 (2015).

11. D. H. Lin, T. Stuwe, S. Schilbach, E. J. Rundlet, T. Perriches, G. Mobbs, Y. Fan, K. Thierbach, F. M. Huber, L. N. Collins, A. M. Davenport, Y. E. Jeon, A. Hoelz, Architecture of the symmetric core of the nuclear pore. *Science (80-.).* **352** (2016), doi:10.1126/science.aaf1015.

12. S. J. Kim, J. Fernandez-Martinez, I. Nudelman, Y. Shi, W. Zhang, B. Raveh, T. Herricks, B. D. Slaughter, J. A. Hogan, P. Upla, I. E. Chemmama, R. Pellarin, I. Echeverria, M. Shivaraju, A. S. Chaudhury, J. Wang, R. Williams, J. R. Unruh, C. H. Greenberg, E. Y. Jacobs, Z. Yu, M. J. de la Cruz, R. Mironska, D. L. Stokes, J. D. Aitchison, M. F. Jarrold, J. L. Gerton, S. J. Ludtke, C. W. Akey, B. T. Chait, A. Sali, M. P. Rout, Integrative structure and functional anatomy of a nuclear pore complex. *Nature*. **555**, 475–482 (2018).

13. T. Stuwe, C. J. Bley, K. Thierbach, S. Petrovic, S. Schilbach, D. J. Mayo, T. Perriches, E. J. Rundlet, Y. E. Jeon, L. N. Collins, F. M. Huber, D. H. Lin, M. Paduch, A. Koide, V. Lu, J. Fischer, E. Hurt, S. Koide, A. A. Kossiakoff, A. Hoelz, Architecture of the nuclear pore inner ring complex. *Science*. **350**, 56 (2015).

14. S. G. Brohawn, T. U. Schwartz, Molecular architecture of the Nup84-Nup145C-Sec13 edge element in the nuclear pore complex lattice. *Nat. Struct. Mol. Biol.* **16**, 1173–1177 (2009).

15. S. O. Obado, M. Brillantes, K. Uryu, W. Zhang, N. E. Ketaren, B. T. Chait, M. C. Field, M. P. Rout, Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex. *PLoS Biol.* **14**, e1002365 (2016).

16. A. G. Floch, B. Palancade, V. Doye, Fifty years of nuclear pores and nucleocytoplasmic transport studies: Multiple tools revealing complex rules. *Methods Cell Biol.* **122**, 1–40 (2014).

17. D. Görlich, U. Kutay, Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15** (1999), pp. 607–660.

18. E. A. Lemke, The Multiple Faces of Disordered Nucleoporins. *J. Mol. Biol.* **428** (2016), pp. 2011–2024.

19. J. Kosinski, S. Mosalaganti, A. von Appen, R. Teimer, A. L. DiGuilio, W. Wan, K. H. Bui, W. J. H. Hagen, J. A. G. Briggs, J. S. Glavy, E. Hurt, M. Beck, Molecular architecture of the inner ring scaffold of the human nuclear pore complex. *Science*. **352**, 363–5 (2016).

20. A. P. Schuller, M. Wojtynek, D. Mankus, M. Tatli, R. Kronenberg-Tenga, S. G. Regmi, P. V. Dip, A. K. R. Lytton-Jean, E. J. Brignole, M. Dasso, K. Weis, O. Medalia, T. U. Schwartz, The cellular

environment shapes the nuclear pore complex architecture. *Nature*. **598**, 667–671 (2021).

21. A. Von Appen, J. Kosinski, L. Sparks, A. Ori, A. L. A. L. DiGuilio, B. Vollmer, M.-T. M.-T. M.-T. M.-T. Mackmull, N. Banterle, L. Parca, P. Kastritis, K. Buczak, S. Mosalaganti, W. Hagen, A. Andres-Pons, E. A. E. A. E. A. Lemke, P. Bork, W. Antonin, J. S. Glavy, K. H. K. H. K. H. Bui, M. Beck, In situ structural analysis of the human nuclear pore complex. *Nature*. **526**, 140–143 (2015).

22. M. Allegretti, C. E. Zimmerli, V. Rantos, F. Wilfling, P. Ronchi, H. K. H. Fung, C. W. Lee, W. Hagen, B. Turoňová, K. Karius, M. Börmel, X. Zhang, C. W. Müller, Y. Schwab, J. Mahamid, B. Pfander, J. Kosinski, M. Beck, In-cell architecture of the nuclear pore and snapshots of its turnover. *Nature*. **586**, 796–800 (2020).

23. C. E. Zimmerli, M. Allegretti, V. Rantos, S. K. Goetz, A. Obarska-Kosinska, I. Zagoriy, A. Halavatyi, G. Hummer, J. Mahamid, J. Kosinski, M. Beck, Nuclear pores dilate and constrict in cellulo. *Science*. **374** (2021), doi:10.1126/science.abd9776.

24. M. Beck, S. Mosalaganti, J. Kosinski, From the resolution revolution to evolution: structural insights into the evolutionary relationships between vesicle coats and the nuclear pore. *Curr. Opin. Struct. Biol.* **52** (2018), pp. 32–40.

25. S. Mosalaganti, J. Kosinski, S. Albert, M. Schaffer, D. Strenkert, P. A. Salomé, S. S. Merchant, J. M. Plitzko, W. Baumeister, B. D. Engel, M. Beck, In situ architecture of the algal nuclear pore complex. *Nat. Commun.* **9**, 2361 (2018).

26. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature*. **596**, 583–589 (2021).

27. M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. Dustin Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. Van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. Christopher Garcia, N. V. Grishin, P. D. Adams, R. J. Read, D. Baker, Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. **373**, 871–876 (2021).

28. T. Maimon, N. Elad, I. Dahan, O. Medalia, The Human Nuclear Pore Complex as Revealed by Cryo-Electron Tomography. *Structure*. **20**, 998–1006 (2012).

29. J. Mahamid, S. Pfeffer, M. Schaffer, E. Villa, R. Danev, L. K. Cuellar, F. Förster, A. A. Hyman, J. M. Plitzko, W. Baumeister, Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science*. **351**, 969–972 (2016).

30. V. Zila, E. Margiotta, B. Turoňová, T. G. Müller, C. E. Zimmerli, S. Mattei, M. Allegretti, K. Börner, J. Rada, B. Müller, M. Lusic, H. G. Kräusslich, M. Beck, Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell*. **184**, 1032-1046.e18 (2021).

31. Bley et al., Architecture of the cytoplasmic face of the nuclear pore. *bioRxiv* (2021), doi:10.1101/2021.10.26.465790.

32. Petrovic et al., Architecture of the linker-scaffold in the nuclear pore. *bioRxiv* (2021), doi:10.1101/2021.10.26.465796.

33. G. Huang, Y. Zhang, X. Zhu, C. Zeng, Q. Wang, Q. Zhou, Q. Tao, M. Liu, J. Lei, C. Yan, Y. Shi, Structure of the cytoplasmic ring of the Xenopus laevis nuclear pore complex by cryo-electron microscopy single particle analysis. *Cell Res. 2020 306*. **30**, 520–531 (2020).

34. R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper, D. Hassabis, *bioRxiv*, in press, doi:10.1101/2021.10.04.463034.

35. M. Mirdita, S. Ovchinnikov, M. Steinegger, *bioRxiv*, in press, doi:10.1101/2021.08.15.456425.

36. I. Humphreys, J. Pei, M. Baek, A. Krishnakumar, I. Anishchenko, S. Ovchinnikov, J. Zhang, T. J. Ness, S. Banjade, S. R. Bagde, V. G. Stancheva, X. H. Li, K. Liu, Z. Zheng, D. J. Barrero, U. Roy, J. Kuper, I.

S. Fernández, B. Szakal, D. Branzei, J. Rizo, C. Kisker, E. C. Greene, S. Biggins, S. Keeney, E. A. Miller, J. C. Fromme, T. L. Hendrickson, Q. Cong, D. Baker, Computed structures of core eukaryotic protein complexes. *Science*. **374** (2021), doi:10.1126/science.abm4805.

37. V. Rantos, K. Karius, J. Kosinski, Integrative structural modeling of macromolecular complexes using Assembline. *Nat. Protoc.* **17**, 152–176 (2022).

38. N. Schrader, P. Stelter, D. Flemming, R. Kunze, E. Hurt, I. R. Vetter, Structural basis of the nic96 subcomplex organization in the nuclear pore channel. *Mol. Cell*. **29**, 46–55 (2008).

39. B. Vollmer, A. Schooley, R. Sachdev, N. Eisenhardt, A. M. Schneider, C. Sieverding, J. Madlung, U. Gerken, B. Macek, W. Antonin, Dimerization and direct membrane interaction of Nup53 contribute to nuclear pore complex assembly. *EMBO J.* **31**, 4072–4084 (2012).

40. J. Mansfeld, S. Güttinger, L. A. Hawryluk-Gara, N. Panté, M. Mall, V. Galy, U. Haselmann, P. Mühlhäusser, R. W. Wozniak, I. W. Mattaj, U. Kutay, W. Antonin, The Conserved Transmembrane Nucleoporin NDC1 Is Required for Nuclear Pore Complex Assembly in Vertebrate Cells. *Mol. Cell*. **22**, 93–103 (2006).

41. Y. Yamazumi, A. Kamiya, A. Nishida, A. Nishihara, S. ichiro Iemura, T. Natsume, T. Akiyama, The transmembrane nucleoporin NDC1 is required for targeting of ALADIN to nuclear pore complexes. *Biochem. Biophys. Res. Commun.* **389**, 100–104 (2009).

42. B. Kind, K. Koehler, M. Lorenz, A. Huebner, The nuclear pore complex protein ALADIN is anchored via NDC1 but not via POM121 and GP210 in the nuclear envelope. *Biochem. Biophys. Res. Commun.* **390**, 205–210 (2009).

43. A. Ori, N. Banterle, M. Iskar, A. Andres-Pons, C. Escher, H. Khanh Bui, L. Sparks, V. Solis-Mezarino, O. Rinner, P. Bork, E. A. Lemke, M. Beck, Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol Syst Biol*. **9**, 648 (2013).

44. 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).

45. D. H. Lin, A. R. Correia, S. W. Cai, F. M. Huber, C. A. Jette, A. Hoelz, Structural and functional analysis of mRNA export regulation by the nuclear pore complex. *Nat. Commun.* **9**, 1–19 (2018).

46. J. M. Cronshaw, M. J. Matunis, The nuclear pore complex protein ALADIN is mislocalized in triple A syndrome. *Proc. Natl. Acad. Sci. USA*. **100**, 5823–5827 (2003).

47. A. Huebner, A. M. Kaindl, K. P. Knobeloch, H. Petzold, P. Mann, K. Koehler, The triple A syndrome is due to mutations in ALADIN, a novel member of the nuclear pore complex. *Endocr. Res.* **30**, 891–899 (2004).

48. A. R. Cho, K. J. Yang, Y. Bae, Y. B. Young, E. Kim, H. Lee, K. K. Jeong, W. Park, H. Rhim, Y. C. Soo, T. Imanaka, S. Moon, J. Yoon, K. Y. Sungjoo, Tissue-specific expression and subcellular localization of ALADIN, the absence of which causes human triple A syndrome. *Exp. Mol. Med.* **41**, 381–386 (2009).

49. P. Upla, S. J. Kim, P. Sampathkumar, K. Dutta, S. M. Cahill, I. E. Chemmama, R. Williams, J. B. Bonanno, W. J. Rice, D. L. Stokes, D. Cowburn, S. C. Almo, A. Sali, M. P. Rout, J. Fernandez-Martinez, Molecular Architecture of the Major Membrane Ring Component of the Nuclear Pore Complex. *Structure*. **25**, 434–445 (2017).

50. Y. Zhang, S. Li, C. Zeng, G. Huang, X. Zhu, Q. Wang, K. Wang, Q. Zhou, C. Yan, W. Zhang, G. Yang, M. Liu, Q. Tao, J. Lei, Y. Shi, Molecular architecture of the luminal ring of the Xenopus laevis nuclear pore complex. *Cell Res.* **30**, 532–540 (2020).

51. 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).

52. S. G. Brohawn, T. U. Schwartz, Molecular architecture of the Nup84-Nup145C-Sec13 edge element in the nuclear pore complex lattice. *Nat. Struct. Mol. Biol.* **16**, 1173–1177 (2009).

53. D. Devos, S. Dokudovskaya, F. Alber, R. Williams, B. T. Chait, A. Sali, M. P. Rout, Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol*. **2**, e380 (2004).

54. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, The MARTINI force field:

Coarse grained model for biomolecular simulations. *J. Phys. Chem. B*. **111**, 7812–7824 (2007).

55. X. Periole, M. Cavalli, S.-J. Marrink, M. A. Ceruso, Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput*. **5**, 1–7 (2009).

56. J. G. Gall, Octagonal Nuclear Pores. *J. Cell Biol.* **32**, 391–399 (1967).

57. C. Wolf, M. R. K. Mofrad, On the Octagonal Structure of the Nuclear Pore Complex: Insights from Coarse-Grained Models. *Biophys. J.* **95**, 2073–2085 (2008).

58. M. Seidel, A. Becker, F. Pereira, J. J. M. Landry, N. T. D. de Azevedo, C. M. Fusco, E. Kaindl, N. Romanov, J. Baumbach, J. D. Langer, E. M. Schuman, K. R. Patil, G. Hummer, V. Benes, M. Beck, Co-translational assembly orchestrates competing biogenesis pathways. *Nat. Commun.* **13**, 1224 (2022).

59. J. P. Stafstrom, L. A. Staehelin, Dynamics of the nuclear envelope and of nuclear pore complexes during mitosis in the Drosophila embryo. *Eur. J. Cell Biol.* **34** (1984).

60. S. Otsuka, A. M. Steyer, M. Schorb, J. K. Hériché, M. J. Hossain, S. Sethi, M. Kueblbeck, Y. Schwab, M. Beck, J. Ellenberg, Postmitotic nuclear pore assembly proceeds by radial dilation of small membrane openings. *Nat. Struct. Mol. Biol.* **25**, 21–28 (2018).

61. K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Žídek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, S. Velankar, G. J. Kleywegt, A. Bateman, R. Evans, A. Pritzel, M. Figurnov, O. Ronneberger, R. Bates, S. A. A. A. Kohl, A. Potapenko, A. J. Ballard, B. Romera-Paredes, S. Nikolov, R. Jain, E. Clancy, D. Reiman, S. Petersen, A. W. Senior, K. Kavukcuoglu, E. Birney, P. Kohli, J. Jumper, D. Hassabis, Highly accurate protein structure prediction for the human proteome. *Nature*. **596**, 590–596 (2021).

62. M. Akdel, D. E. V Pires, E. P. Pardo, J. Jänes, A. O. Zalevsky, B. Mészáros, P. Bryant, L. L. Good, R. A. Laskowski, G. Pozzati, A. Shenoy, W. Zhu, P. Kundrotas, V. R. Serra, C. H. M. Rodrigues, A. S. Dunham, D. Burke, N. Borkakoti, S. Velankar, A. Frost, K. Lindorff-Larsen, A. Valencia, S. Ovchinnikov, J. Durairaj, D. B. Ascher, J. M. Thornton, N. E. Davey, A. Stein, A. Elofsson, T. I. Croll, P. Beltrao, A structural biology community assessment of AlphaFold 2 applications. *bioRxiv* (2021), doi:10.1101/2021.09.26.461876.

63. G. Masrati, M. Landau, N. Ben-Tal, A. Lupas, M. Kosloff, J. Kosinski, Integrative Structural Biology in the Era of Accurate Structure Prediction. *J. Mol. Biol.* **433**, 167127 (2021).

64. J. Pereira, A. J. Simpkin, M. D. Hartmann, D. J. Rigden, R. M. Keegan, A. N. Lupas, High-accuracy protein structure prediction in CASP14. *Proteins Struct. Funct. Bioinforma.* **89**, 1687–1699 (2021).

65. V. Ruiz-Serra, C. Pontes, E. Milanetti, A. Kryshtafovych, R. Lepore, A. Valencia, Assessing the accuracy of contact and distance predictions in CASP14. *Proteins Struct. Funct. Bioinforma.* **89**, 1888–1900 (2021).

66. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Applying and improving AlphaFold at CASP14. *Proteins Struct. Funct. Bioinforma.* **89**, 1711–1721 (2021).

67. P. Bryant, G. Pozzati, A. Elofsson, *Nat. Commun.*, in press, doi:10.1038/s41467-022-28865-w.

68. K. Buczak, Spatial proteomics: from tissue organization to protein function. *Dr. thesis* (2019), doi:10.11588/HEIDOK.00025909.

69. A. Rohou, N. Grigorieff, S. LaValle, J. Mitchell, I. Lazić, U. Lücken, H. Rullgård, O. Öktem, B. Rieger, P. Beltran-Alvarez, D. Castaño-Diez, W. Chen, D. Devos, M. Güell, T. Norambuena, I. Racke, V. Rybin, A. Schmidt, E. Yus, R. Aebersold, R. Herrmann, B. Böttcher, A. Frangakis, R. Russell, L. Serrano, P. Bork, A. Gavin, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

70. D. N. Mastronarde, S. R. Held, Automated tilt series alignment and tomographic reconstruction in IMOD. *J. Struct. Biol.* **197**, 102–113 (2017).

71. B. Turoňová, F. K. M. Schur, W. Wan, J. A. G. Briggs, Efficient 3D-CTF correction for cryo-electron

tomography using NovaCTF improves subtomogram averaging resolution to 3.4 Å. *J. Struct. Biol.* **199**, 187–195 (2017).

72. B. Turonova, turonova/novaSTA: Advanced particle analysis (2022), , doi:10.5281/zenodo.5921012.

73. D. Castaño-Díez, M. Kudryashev, M. Arheit, H. Stahlberg, Dynamo: A flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *J. Struct. Biol.* **178**, 139–151 (2012).

74. S. H. W. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

75. W. J. H. Hagen, W. Wan, J. A. G. Briggs, Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J. Struct. Biol.* **197**, 191–198 (2016).

76. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*. **14**, 290–296 (2017).

77. D. Asarnow, E. Palovcak, Y. Cheng, asarnow/pyem: UCSF pyem v0.5 (2019) (available at https://zenodo.org/record/3576630).

78. T. Nakane, D. Kimanius, E. Lindahl, S. H. H. W. Scheres, Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *Elife*. **7** (2018), doi:10.7554/eLife.36861.

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

80. L. J. McGuffin, K. Bryson, D. T. Jones, The PSIPRED protein structure prediction server. *Bioinformatics*. **16**, 404–405 (2000).

81. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 531–544 (2018).

82. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. **25**, 1605–1612 (2004).

83. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

84. M.-T. Mackmull, B. Klaus, I. Heinze, M. Chokkalingam, A. Beyer, R. B. Russell, A. Ori, M. Beck, Landscape of nuclear transport receptor cargo specificity. *Mol. Syst. Biol.* **13**, 962 (2017).

85. M. I. Dauden, J. Kosinski, O. Kolaj-Robin, A. Desfosses, A. Ori, C. Faux, N. A. Hoffmann, O. F. Onuma, K. D. Breunig, M. Beck, C. Sachse, B. Séraphin, S. Glatt, C. W. Müller, Architecture of the yeast Elongator complex. *EMBO Rep.* (2016), doi:10.15252/embr.201643353.

86. K. Strimmer, fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics*. **24**, 1461–1462 (2008).

87. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B*. **57**, 289–300 (1995).

88. B. Webb, S. Viswanath, M. Bonomi, R. Pellarin, C. H. Greenberg, D. Saltzberg, A. Sali, Integrative structure modeling with the Integrative Modeling Platform. *Protein Sci.* **27**, 245–258 (2018).

89. D. Saltzberg, C. H. Greenberg, S. Viswanath, I. Chemmama, B. Webb, R. Pellarin, I. Echeverria, A. Sali, in *Methods in Molecular Biology* (2019), vol. 2022.

90. B. Vollmer, M. Lorenz, D. Moreno-Andrés, M. Bodenhöfer, P. De Magistris, S. A. Astrinidis, A. Schooley, M. Flötenmeyer, S. Leptihn, W. Antonin, Nup153 Recruits the Nup107-160 Complex to the Inner Nuclear Membrane for Interphasic Nuclear Pore Complex Assembly. *Dev. Cell*. **33**, 717–728 (2015).

91. B. Webb, A. Sali, Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* (2016), doi:10.1002/cpbi.3.

92. T. I. Croll, ISOLDE : a physically realistic environment for model building into low-resolution electron-

density maps. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 519–530 (2018).

93. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. **1**–**2**, 19–25 (2015).

94. K. J. Boyd, E. R. May, BUMPy: A Model-Independent Tool for Constructing Lipid Bilayers of Varying Curvature and Composition. *J. Chem. Theory Comput.* **14**, 6642–6652 (2018).

95. T. A. Wassenaar, H. I. Ingólfsson, R. A. Böckmann, D. P. Tieleman, S. J. Marrink, Computational Lipidomics with insane : A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).

96. M. Vögele, J. Köfinger, G. Hummer, Molecular dynamics simulations of carbon nanotube porins in lipid bilayers. *Faraday Discuss.* **209**, 341–358 (2018).

97. R. M. Bhaskara, S. M. Linker, M. Vögele, J. Köfinger, G. Hummer, Carbon Nanotubes Mediate Fusion of Lipid Vesicles. *ACS Nano*. **11**, 1273–1280 (2017).

98. K. W, S. C, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. **22**, 2577–2637 (1983).

99. Z. Benayad, S. Von Bülow, L. S. Stelzl, G. Hummer, Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. *J. Chem. Theory Comput.* **17**, 525–537 (2021).

100. M. Javanainen, H. Martinez-Seara, I. Vattulainen, Excessive aggregation of membrane proteins in the Martini model. *PLoS One*. **12**, e0187936 (2017).

101. L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, S.-J. Marrink, The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).

102. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

103. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101 (2007).

104. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

105. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

106. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **32**, 2319 (2011).

107. P. Georgiades, V. J. Allan, G. D. Wright, P. G. Woodman, P. Udommai, M. A. Chung, T. A. Waigh, The flexibility and dynamics of the tubules in the endoplasmic reticulum. *Sci. Reports 2017 71*. **7**, 1–10 (2017).

108. B. Różycki, E. Boura, J. H. Hurley, G. Hummer, Membrane-Elasticity Model of Coatless Vesicle Budding Induced by ESCRT Complexes. *PLoS Comput. Biol.* **8**, e1002736 (2012).

109. S. Amlacher, P. Sarges, D. Flemming, V. van Noort, R. Kunze, D. P. P. Devos, M. Arumugam, P. Bork, E. Hurt, V. van Noort, R. Kunze, D. P. P. Devos, M. Arumugam, P. Bork, E. Hurt, V. van Noort, R. Kunze, D. P. P. Devos, M. Arumugam, P. Bork, E. Hurt, Insight into Structure and Assembly of the Nuclear Pore Complex by Utilizing the Genome of a Eukaryotic Thermophile. *Cell*. **146**, 277–289 (2011).

110. P. De Magistris, M. Tatarek-Nossol, M. Dewor, W. Antonin, A self-inhibitory interaction within Nup155 and membrane binding are required for nuclear pore complex formation. *J. Cell Sci.* **131** (2018), doi:10.1242/jcs.208538.

111. L. Tai, Y. Zhu, H. Ren, X. Huang, C. Zhang, F. Sun, 8 Å structure of the outer rings of the Xenopus laevis nuclear pore complex obtained by cryo-EM and AI. *Protein Cell* (2022), doi:10.1007/s13238-021-00895-y.

112. R. W. Wozniak, E. Bartnik, G. Blobel, Primary structure analysis of an integral membrane glycoprotein of the nuclear pore. *J. Cell Biol.* **108**, 2083–2092 (1989).

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

114. P. Grandi, T. Dang, N. Pané, A. Shevchenko, M. Mann, D. Forbes, E. Hurt, 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 (1997).

115. G. Huang, X. Zhan, C. Zeng, X. Zhu, K. Liang, Y. Zhao, P. Wang, Q. Wang, Q. Zhou, Q. Tao, M. Liu, J. Lei, C. Yan, Y. Shi, Cryo-EM structure of the nuclear ring from Xenopus laevis nuclear pore complex. *Cell Res.* (2022), doi:10.1038/s41422-021-00610-w.

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

117. B. A. Rasala, A. V Orjalo, Z. X. Shen, S. Briggs, D. J. Forbes, ELYS is a dual nucleoporin/kinetochore protein required for nuclear pore assembly and proper cell division. *Proc. Natl. Acad. Sci. USA*. **103**, 17801–17806 (2006).

118. Y. Zhang, J. Skolnick, TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005).

119. S. Mukherjee, Y. Zhang, MM-align: A quick algorithm for aligning multiple-chain protein complex structures using iterative dynamic programming. *Nucleic Acids Res.* **37** (2009), doi:10.1093/nar/gkp318.

120. V. Mariani, M. Biasini, A. Barbato, T. Schwede, IDDT: A local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*. **29**, 2722–2728 (2013).

121. M. Gao, J. Skolnick, B. Rost, IAlign: a method for the structural comparison of protein-protein interfaces. *Bioinformatics*. **27**, 2259–2265 (2011).

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

123. E. W. Debler, Y. Ma, H. S. Seo, K. C. Hsia, T. R. Noriega, G. Blobel, A. Hoelz, A Fence-like Coat for the Nuclear Pore Membrane. *Mol. Cell*. **32**, 815–826 (2008).

124. X. Liu, J. M. Mitchell, R. W. Wozniak, G. Blobel, J. Fan, Structural evolution of the membrane-coating module of the nuclear pore complex. *Proc Natl Acad Sci U S A*. **109**, 16498–16503 (2012).

125. K. Yoshida, H. S. Seo, E. W. Debler, G. Blobel, A. Hoelz, 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 (2011).

126. 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**, 425–31 (2015).

127. T. Pupko, R. E. Bell, I. Mayrose, F. Glaser, N. Ben-Tal, Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. *Bioinformatics*. **18**, S71–S77 (2002).

128. E. Krissinel, K. Henrick, Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797 (2007).

129. S. Malhotra, A. P. Joseph, J. Thiyagalingam, M. Topf, Assessment of protein–protein interfaces in cryo-EM derived assemblies. *Nat. Commun.* **12** (2021), doi:10.1038/s41467-021-23692-x.

130. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67** (2011), pp. 235–242.

Acknowledgments: We acknowledge support from the Electron Microscopy Core Facility (EMCF) and IT services of EMBL Heidelberg. We thank Sonja Welsch at the Central Electron Microscopy Facility of the Max Planck Institute of Biophysics for technical expertise. We thank Thomas Hoffman and Renato Alves for help with the AlphaFold installation.

Funding:

MB acknowledges funding by EMBL, the Max Planck Society, and the European Research Council (ComplexAssembly 724349).

JK acknowledges funding from the Federal Ministry of Education and Research of Germany (FKZ 031L0100).

M.S. and G.H.'s work on computer simulations was supported by the Max Planck Society.

M.S. was supported by the EMBL Interdisciplinary Postdoc Programme under Marie Curie COFUND actions.

M.S. and G.H. were supported by the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz (LOEWE) DynaMem program of the State of Hessen.

Author contributions: SM prepared the samples, collected, and analyzed the HeLa envelope data. SM, WH, and EM collected the HeLa envelope and HeLa in cellulo data. SM and BT performed the cryo-ET analysis. CEZ prepared, collected, and analyzed the HEK control and NUP210D data. RT determined the structure of human NUP155. MTM performed the BioID experiments. FS and KB prepared the HEK NUP210D cell line. AOK performed modeling and prepared figures, MS performed the MD simulations, GH performed the membrane elastic theory analysis, JK and AOK prepared software, SM, GH, JK and MB conceptualized the study, supervised the project, and wrote the manuscript.

CRediT roles:

Conceptualization: SM, AOK, GH, JK, MB

Methodology: SM, AOK, BT, MS, JK

Investigation: SM, AOK, MS, RT, BT, CEZ, KB, FS, EM, MTM, WH, GH, JK, MB

Visualization: SM, AOK, MS, RT, CEZ, MTM, JK

Funding acquisition: MB

Project administration: GH, JK, MB

Supervision: SM, GH, JK, MB

Writing – original draft: SM, MS, GH, MB, JK

Writing – review & editing: SM, AOK, MS, GH, MB, JK

Competing interests: Authors declare that they have no competing interests.

Data and materials availability:

EM maps associated with the manuscript have been deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD-14243, EMD-14321, 14322, 14325, 14326, 14327, 14328, and 14330, respectively. The models have been deposited into the Protein Data Bank (PDB) under accession codes 7R1Y, 7R5K and 7R5J respectively. The predicted local quality scores (pLDDT) have been reported in the B-factor property of the deposited models. Software code for subtomogram averaging and geometric cleaning is available at Zenodo (zenodo.5921012) (72).

Supplementary Materials

Supplementary Text

Figs. S1 to S29

Tables S1 to S6

References 107 to 130

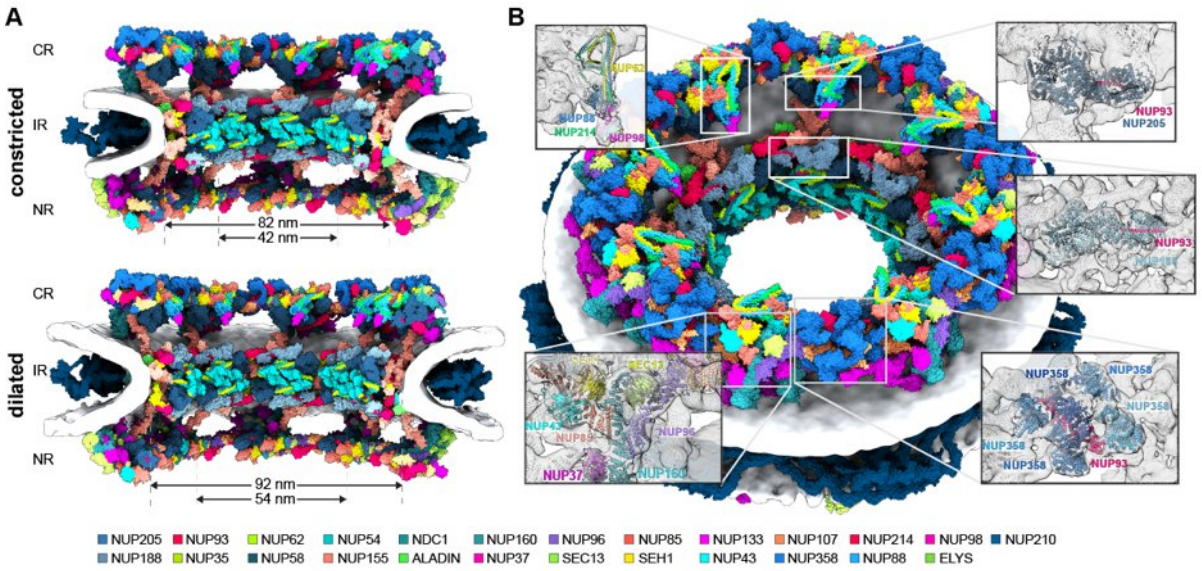


Fig. 1 Scaffold architecture of the human NPC. (A) The near-complete model of the human NPC scaffold is shown for the constricted and dilated states as cut-away views. High-resolution models are color-coded as indicated in the color bar. The nuclear envelope is shown as a gray isosurface. **(B)** Same as **(A)** but shown from the cytoplasmic side for the constricted NPC. The insets show individual features of the CR and IR enlarged with secondary structures displayed as cartoons and superimposed with the isosurface-rendered cryo-ET map of the human NPC (gray).

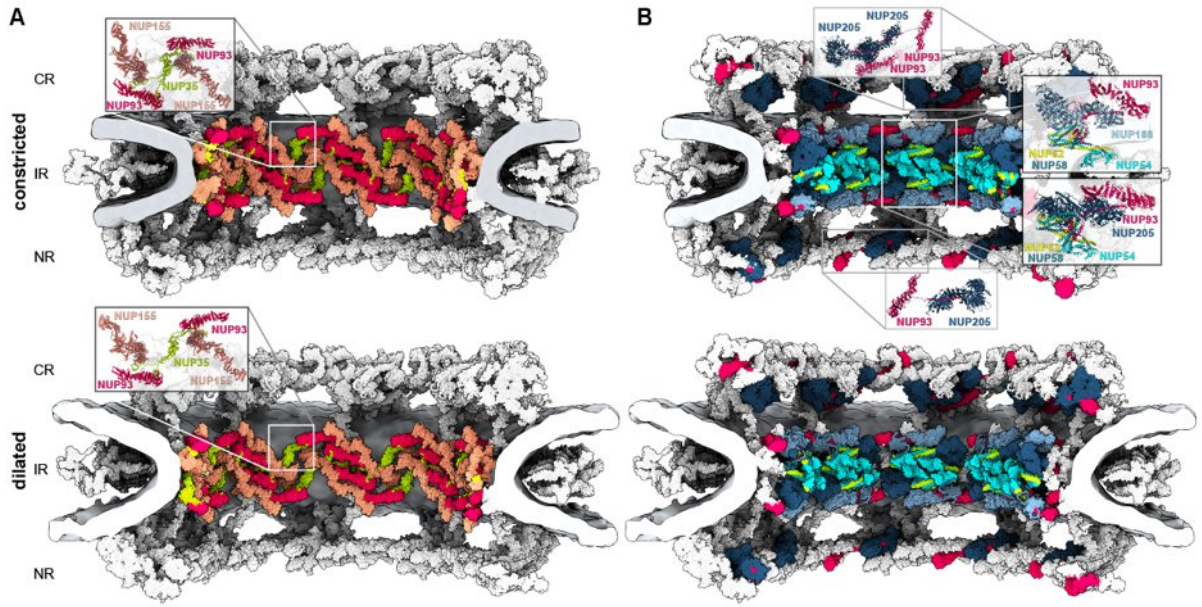


Fig. 2 The connectivity of protein linkers within the human NPC. (A) The NUP35 dimer interconnects adjacent spokes across different subcomplex species, thus facilitating cylindrical assembly of the IR in both constricted (top) and dilated (bottom) states. The NUP205, NUP188, NUP62 complex, and the N-terminus (aa. 1-170) of NUP93 are hidden from view to expose NUP35. **(B)** The N-terminus of NUP93 isostoichiometrically connects the subunits NUP205,

NUP188, NUP62, NUP58 and NUP54 within the same subcomplex species. Insets show NUP93 connectivity**,** highlighting its interaction with two copies of NUP205 in the CR, two copies each with NUP205 and 188 in the IR and a single copy of NUP205 in the NR. The respective subunits are shown color-coded as in Fig. 1, while all other subunits and the nuclear membranes are shown in gray.

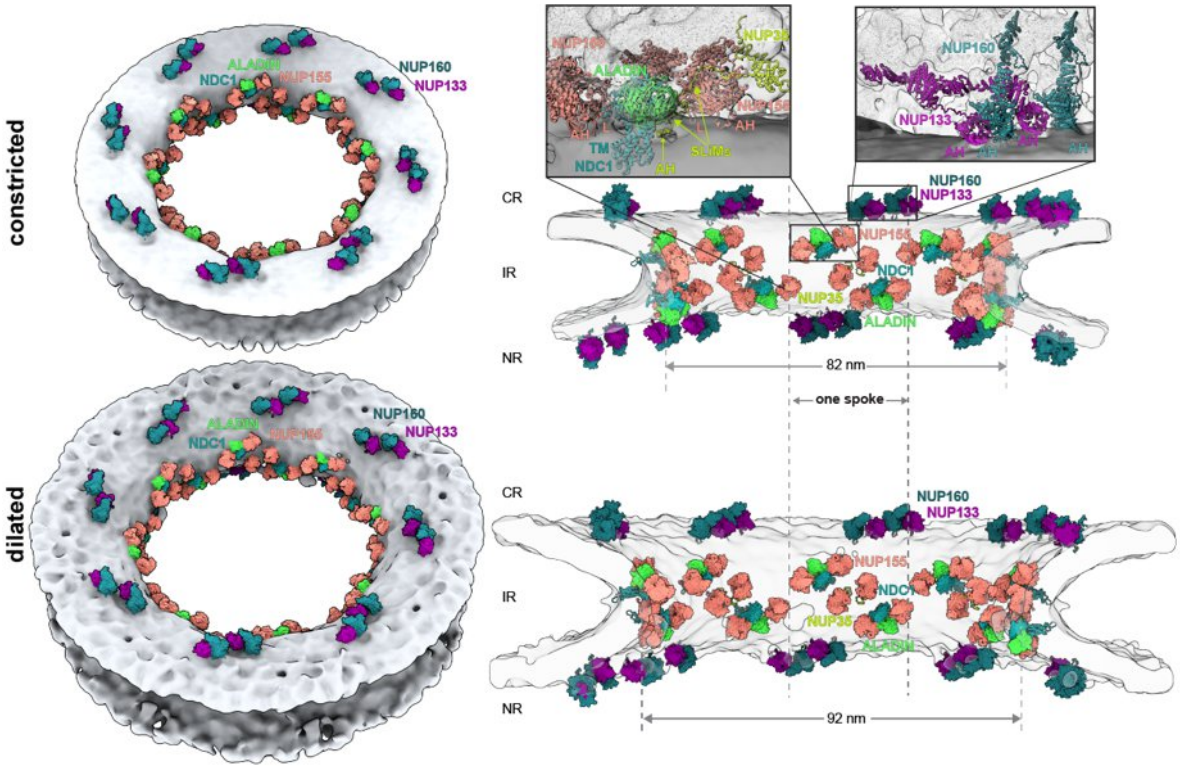


Fig. 3 The membrane-anchoring motifs of the human NPC are distributed over the entire scaffold. The membrane-binding β-propellers of the Y-complex and IR complex are shown color-coded and arranged as pairs of the respective inner and outer copies. ALADIN and NDC1 form a transmembrane interaction hub with the inner and connector copies of NUP155, which is shown enlarged in the inset in the cut-away side view (right). The nuclear membranes are shown as a gray isosurface.

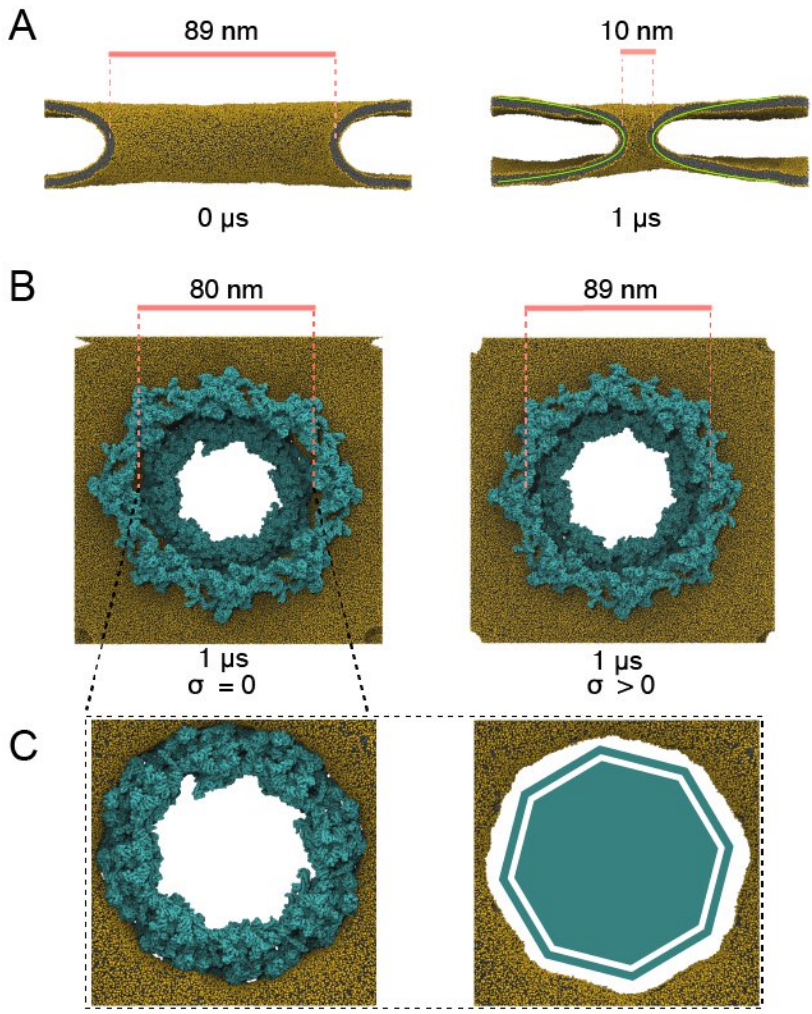


Fig. 4 Dynamics of the NPC from molecular simulations. (A) An isolated half-toroidal double-membrane pore shaped initially as in the tomographic structure of the constricted NPC (left) tightens over the course of 1 µs of MD (right) toward the catenoid-like shape (green) predicted by membrane elastic theory. Shown are cuts along the axis of the double-membrane pore with lipid headgroups and tails in gold and gray, respectively. The solvent is not shown. (B) The NPC (cyan) widens by ~10% in response to lateral membrane tension (right; ΔP=2 bar) compared to a zero-tension simulation (left; ΔP=0). Shown are snapshots of the relaxed structures after 1 µs of MD. (C) The membrane fits tightly around the NPC inner ring (cyan, left; ΔP=0) and forms an octagonally-shaped pore (right, NPC not shown).

Movie 1 MD simulation trajectory of an isolated half-toroidal double membrane shaped initially as in the tomographic structure of the constricted NPC. The pore tightens within 1.2 μs of MD (see also Fig. 4B and Fig. S23A for the diameter time trace). A top view of lipids is shown. Solvent is omitted for clarity.

Movie 2 MD simulation of the NPC (cyan) covering approximately 1.2 μs with a=1.0 (see also Fig. 4B and Fig. S23A for the diameter time trace). A top view of the NPC with membrane is shown. Solvent is omitted for clarity.

Movie 3 MD simulation of the NPC (cyan) covering approximately 1.2 μs with 10 a=0.7 (see also Fig. 4B and Fig. S23A for the diameter time trace). A top view of the NPC with membrane is shown. Solvent is omitted for clarity.

Science
AAAS

Supplementary Materials for

AI-based structure prediction empowers integrative structural analysis of human nuclear pores

Authors: Shyamal Mosalaganti, Agnieszka Obarska-Kosinska, Marc Siggel, Reiya Taniguchi, Beata Turoňová, Christian E. Zimmerli, Katarzyna Buczak, Florian H. Schmidt, Erica Margiotta, Marie-Therese Mackmull, Wim J. H. Hagen, Gerhard Hummer, Jan Kosinski, Martin Beck

Correspondence to: Gerhard Hummer, gerhard.hummer@biophys.mpg.de; Jan Kosinski, jan.kosinski@embl.de; Martin Beck, martin.beck@biophys.mpg.de

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Supplementary Materials

Supplementary Text

Pore shape from Helfrich membrane elastic theory

According to Helfrich membrane elastic theory, a double-membrane pore will relax to a catenoid shape, $r(z) = R\cosh(z/R)$, where r and z are radial and axial cylindrical coordinates, and $R$is the pore radius. For a catenoid, the mean curvature is zero and the bending energy vanishes. In MD simulations of a double-membrane pore under periodic boundary conditions, we found the membrane shape to approach the catenoid in the region of the pore (**Fig. 4**A). Further contraction in the simulations is limited by the distortions in a membrane of finite thickness and by membrane-membrane repulsion.

In the tomographic structures of the NPC, the double-membrane pore is substantially wider than expected for a catenoid. In the constricted state, the pore diameter is about $D = 2R \approx 90$ nm, and the distance between ONM and INM about $H \approx 30$ nm. We estimated the energy to deform the membrane to a diameter-to-width ratio of $x = D/H \approx 3$ using membrane elastic theory. We approximate the shape of the pore neck by a nodoid, which in cylindrical coordinates is given by

$$r(t) = a\left(\sqrt{\beta + cos^2 t} - cos\, t\right)$$

$$z(t) = a\left(sin\, t - \int_0^t \frac{cos^2 u}{\sqrt{\beta + cos^2 u}}\right)$$

for $-\pi/2 \le t \le \pi/2$, $a > 0$, and $\beta > 0$ with a pore radius of $R = r(0) = a\left(\sqrt{\beta + 1} - 1\right)$. As a surface of constant mean curvature $H = 1/(2a)$, the nodoid is a minimum energy surface in membrane elastic theory for a non-zero spontaneous curvature or, equivalently, for a given area difference, as would be expected for an asymmetric membrane with protein anchors placed in the outer leaflet. The bending energy for the nodoid half-toroidal pore is

$$E_b = \frac{2\kappa A}{(2a)^2} = 2\pi\kappa\left[\sqrt{\beta}\left(2E(-1/\beta) - K(-1/\beta)\right) - 2\right]$$

with κ the membrane bending rigidity, and $K(u) = \int_0^{\pi/2} dt / \sqrt{1 - u\, sin^2 t}$ and $E(u) = \int_0^{\pi/2} dt \sqrt{1 - u\, sin^2 t}$ the complete elliptic integrals of the first and second kind. The ratio of pore diameter to inter-membrane spacing is

$$x = \frac{D}{H} = \frac{r(0)}{z(\pi/2)} = \frac{1 + \beta - \sqrt{1 + \beta}}{\sqrt{1 + \beta} - (1 + \beta)\, E\left[\frac{1}{1 + \beta}\right] + \beta K\left[\frac{1}{1 + \beta}\right]}$$

For large $x = D/H$ratios, the bending energy grows linearly as $E_b \sim \kappa(\pi^2 x - 10.45)$. For the experimentally observed ratios of $x \approx 3$, the bending energy of the nodoid-shaped pore is about $E_b \approx 20.5\kappa$. Even for the comparably low bending rigidity of $\kappa \approx 25$ kJ/mol reported for the nuclear envelope (*107*), this amounts to a bending energy of about 500 kJ/mol or about 200 k_BT. For reference, we also performed a variational minimization of the membrane bending energy with respect to the function $r(z)$ defining an axially symmetric pore shape in cylindrical coordinates. In addition to constraining the ratio x, we imposed infinite and zero slope at reduced arc lengths $\tau_0 = 0$ and $\tau_1 = 1$, respectively. For the minimization, we used a

Fourier representation with 20 modes (*108*). The minimized bending energy is about $E_b \approx 18.2\kappa$, only about 10% lower than that for the nodoid. The energetic cost of widening the pore diameter to about three times the spacing between ONM and INM is thus substantial and requires compensation by protein-protein and protein-membrane interactions of the NPC that stabilize the pore shape.

Consistency of NUP assignments with complementary experimental data

The NUP SLiM interactions, within NUP93-NUP62-complex, NUP93-NUP205/188, NUP35-NUP155, NUP35-NUP93 have been biochemically well characterized or structurally analyzed in various species (*10–12*, *44*, *109*, *110*). ColabFold detects them in human proteins which is consistent with the accompanying publication (*32*). The helical secondary structure of the NUP93 SLiM binding to NUP205 in the CR and NR is observed in Xenopus cryo-EM maps (**Fig. S10** (*33*, *111*). Their geometry is analyzed in Figures (**Fig. S16**, **Fig. S17**, **Fig. S18**).

The assignment of NUP35 at the interface of adjacent spokes within IR is consistent with the observed electron optical density (**Fig. S14**). It matches the biochemically determined interaction network of NUP35 with NUP93, NUP155 (*12*, *32*, *39*, *109*). The proximity to membrane is verified by the proximity labeling data (**Fig. S20**) that identifies the transmembrane NUP NDC1 as a prominent hit. The C-terminal amphipathic helix of NUP35 (*39*) is in reach to the membrane.

The positioning of NUP210 (**Fig. S21**) is consistent with its topology; it is the only nucleoporin extends into the luminal space between the nuclear membranes (*112*).The model is consistent with the very unique elongated shape and domain organization as chain of multiple IgG folds (*49*), the electron optical density observed in Xenopus cryo-EM maps (**Fig. S10** (*50*)) and known interactions of NUP210 homologs with NDC1 (*51*). The model further is validated using the cryo-EM map obtained from knock out conditions in which the respective density is missing (**Fig. S3**).

The assignments of additional copies of NUP93 in the CR and NR are consistent with its secondary structure as observed in Xenopus cryo-EM maps (**Fig. S10**; (*33*, *111*) it's known interactions with NUP205 (*113*) and established components of the CR and NR (*21*, *32*, *114*) as well as the respective linker connectivity (**Fig. S18**).

The assignment of NDC1/ALADIN is consistent with the interaction network of these proteins with Nup53, NDC1 and IR components that has been established biochemically (*41*, *42*, *43*) and the MS-based proximity labeling data (**Fig. S19 A**) that places both proteins adjacent to NUP35 and NUP155 in the IR. This data further identifies nucleoporin GLE1 as enriched, that has been shown to interact with NUP155 (*45*) and this interaction is accurately captured by AlphaFold (**Fig. S19B**). The NDC1/ALADIN model fits into the cryo-EM map significantly (**Fig. S14**). NDC1 is the only multiple pass transmembrane nucleoporin and its position consistent with the only electron optical density spanning the membrane that is observed at 16 positions within the IR in cryo-EM maps (*21*). A very similar assignment has been recently proposed (*115*).

The assignment of the N-terminal domain of NUP358 is consistent with previous localization experiments that placed it into this region of the CR (*21*) and secondary structure observed in Xenopus cryo-EM maps (**Fig. S10** (*33*)). The experimental structure of the N-terminus of NUP358 does not perfectly superpose with the AlphaFold (**Fig. S5**), possibly due to crystal contacts. Nevertheless, the AlphaFold model fits the cryo-EM map better, as quantified in **Fig. S7**H. At last, this assignment is consistent with structural and biochemical work reported in the accompanying paper by the Hoelz laboratory (*33).*

The positioning of the NUP214 complex is consistent with previous localization experiments that placed it into this region of the CR (8), with the observed electron density (**Fig. S14**) and secondary structure observed in Xenopus cryo-EM maps (**Fig. S10** (*33*)). The predicted interfaces NUP214 and NUP88 that is biochemically validated in (*31*) and structurally similar to the fungal X-ray structures (PDB IDs: 5CWW and 3PBP).

The assignment of ELYS is consistent with its localization in the NR (*21*), its known interactions with NUP160 (*116*, *117*), and secondary structure observed in Xenopus cryo-EM maps (**Fig. S10**, (*111*)).

Figs. S1 to S29

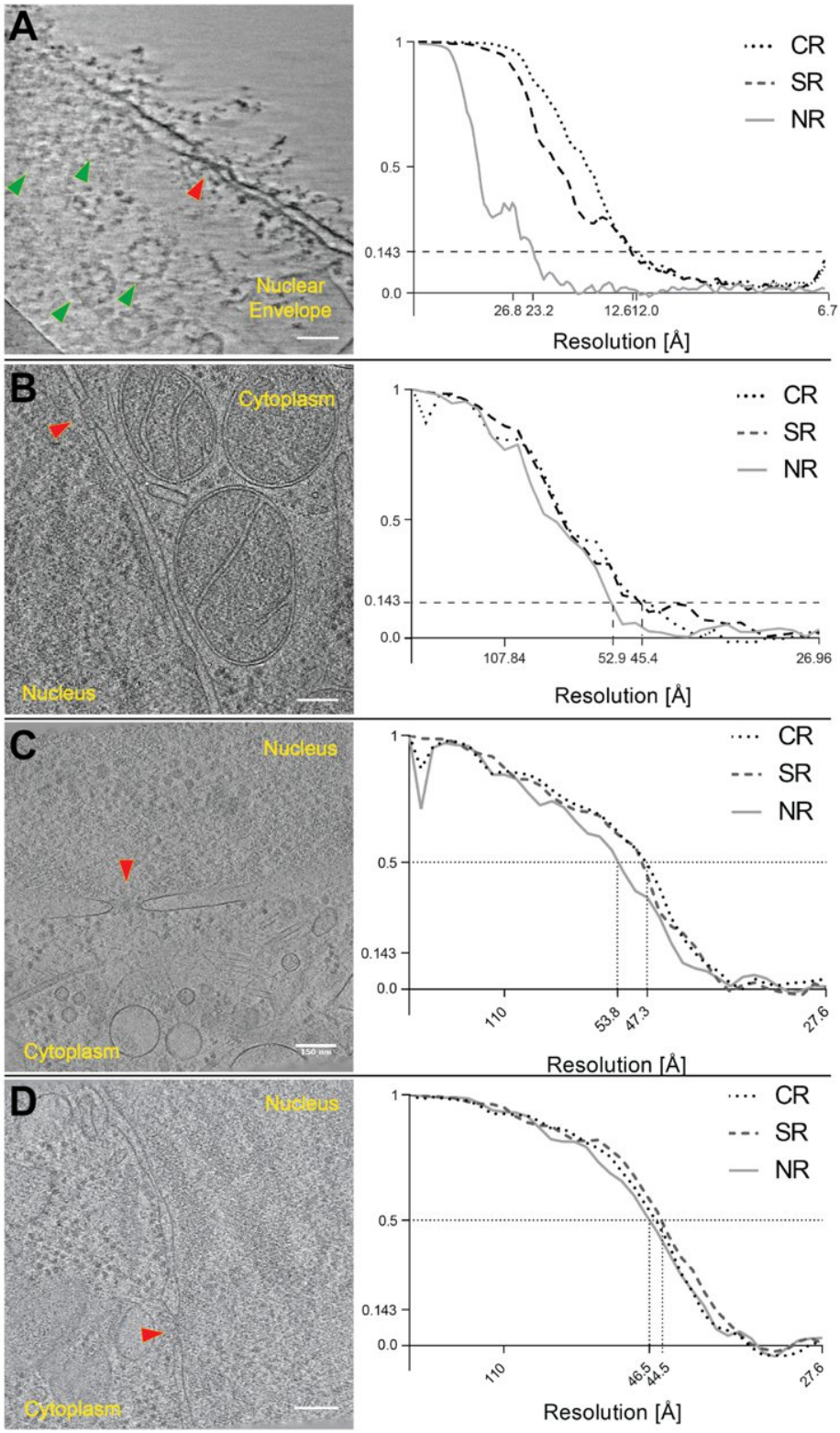


Fig. S1 Slices through tomograms and corresponding Fourier Shell correlation plots. Slices through representative tomograms of isolated nuclear envelopes **(A)**, cryo-FIB milled HeLa cells **(B)**, cryo-FIB milled HEK cells **(C)** or cryo-FIB milled HEK NUP210Δ cells **(D)**. NPCs are highlighted with red arrowheads (side view) or green arrowheads (top view, seen only in case of nuclear envelopes). Fourier Shell Correlation plots of the corresponding NPC averages of the individual rings is shown on the right. Scale bar: 100 nm (A), 150 nm (B-D).

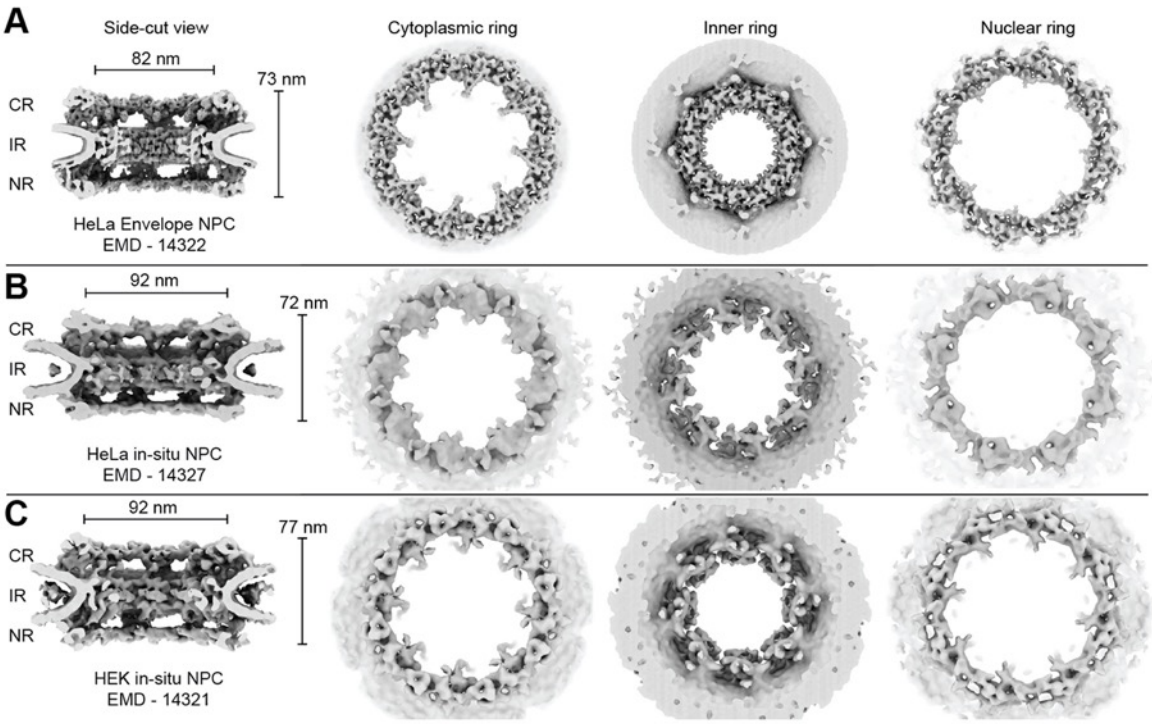


Fig. S2 Structure of human NPCs from isolated NEs or *in cellulo.* Isosurface views of the cut-side view, cytoplasmic, inner and nuclear rings of the structure of NPC obtained from isolated HeLa NE **(A)**, cryo-FIB milled control HeLa cells **(B)** and cryo-FIB milled control HEK cells **(C)** are shown (from L-R). The *In cellulo* NPCs have a larger central channel diameter in comparison to NPC from isolated HeLa envelope.

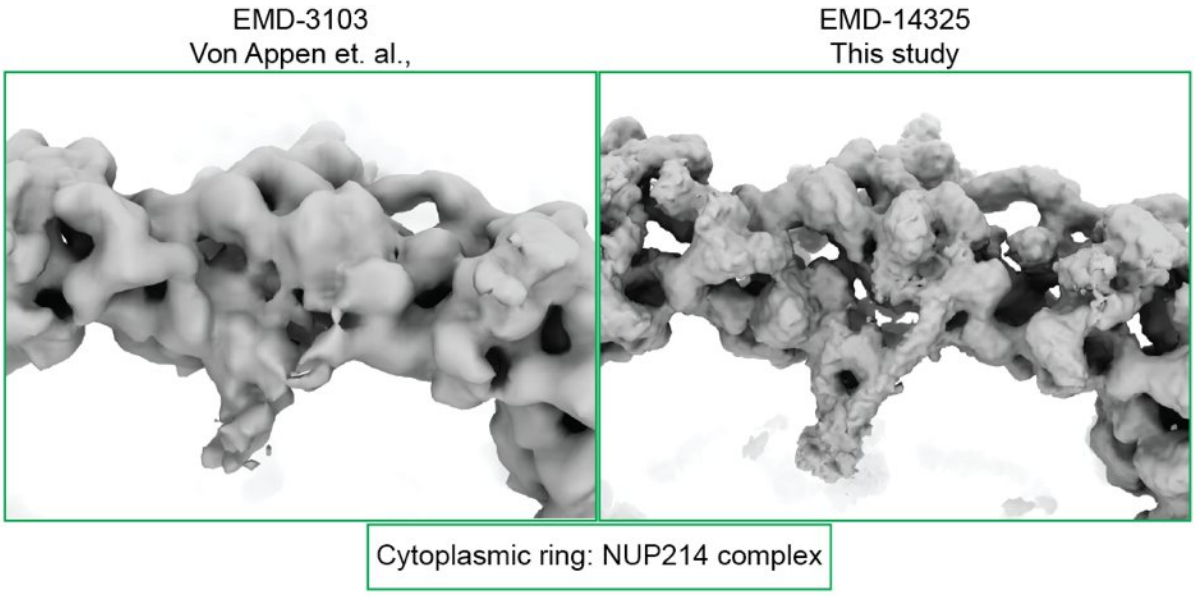


Fig. S3 Comparison of human NPC structures. Resolution improvement in our new structure of NPC from isolated HeLa envelopes (right) in comparison to existing human structure (left). Region of the cytoplasmic ring containing NUP214 subcomplex is highlighted.

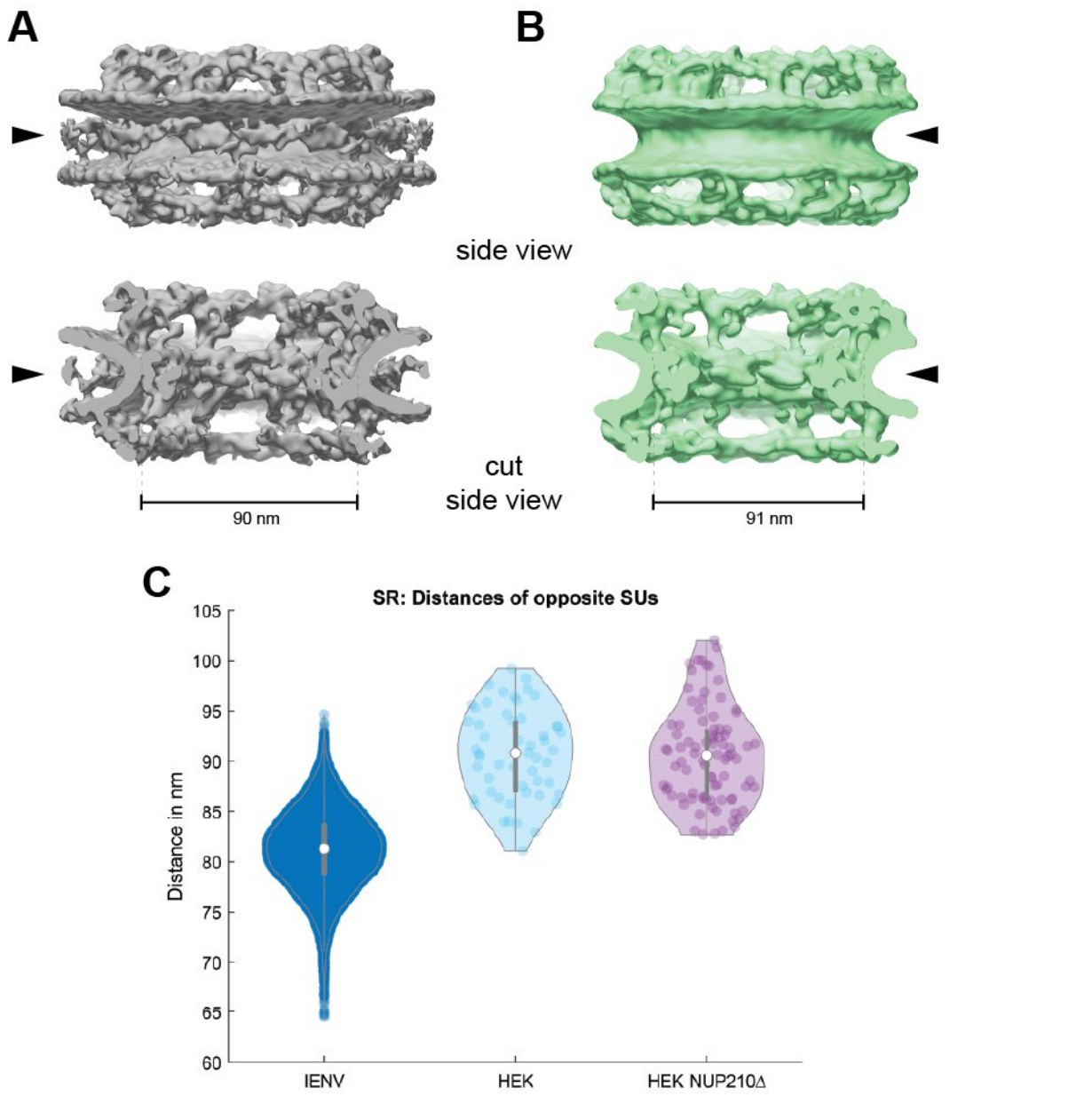


Fig. S4 *In cellulo* human NPC cryo-ET maps from HEK control and NUP210Δ cells. Isosurface representation non-modified (grey, left) and NUP210Δ (green, right) human *in cellulo* NPC cryo-EM maps shown in side view **(A)** and cut-side view **(B)** along the central axis. Black arrowheads highlight the luminal density corresponding to NUP210 which is present in the control (left) and missing in the NUP210Δ map. (**C**) Plot of distances of opposite subunits, measured at the inner/spoke ring highlighting the deviation from perfect circular geometry. Diameter of the NPC from isolated envelopes (IENV) is significantly smaller than NPCs *in cellulo* (HEK). Removal of NUP210 (HEK NUP210Δ) has no effect on the mean diameter of NPCs *in cellulo*.

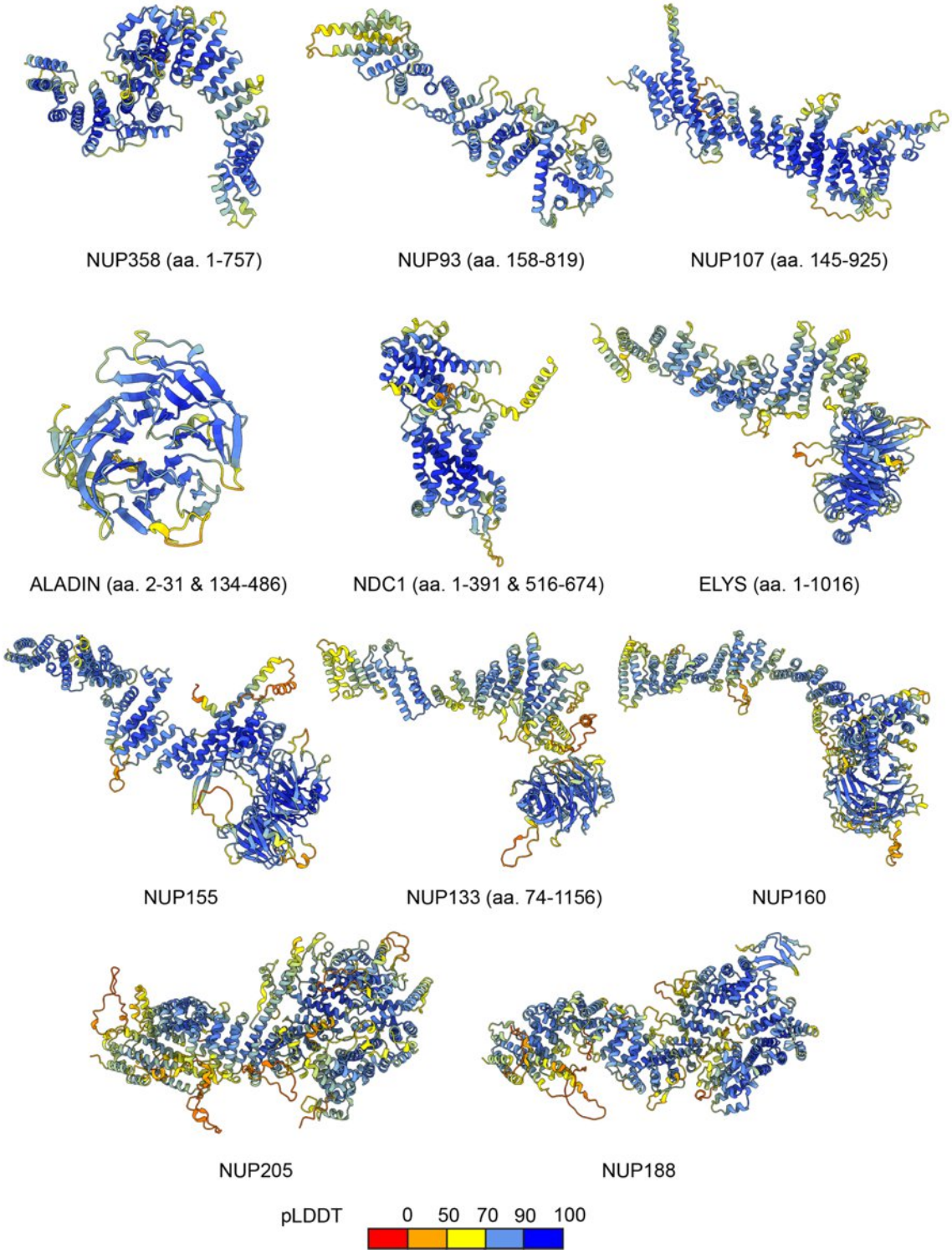


Fig. S5 Estimated quality of monomeric structural models of NUPs built using AlphaFold. The models are colored by local model confidence estimated with predicted local distance difference test (pLDDT), as returned by AlphaFold. The pLDDT > 90 (dark blue) indicates high estimated accuracy of backbone and side chain rotamers whereas pLDDT > 70 (yellow) indicates likely correct backbone prediction (*61*). Long low confidence regions not included in the final NPC models are hidden from view for clarity. aa. – amino acid residues.

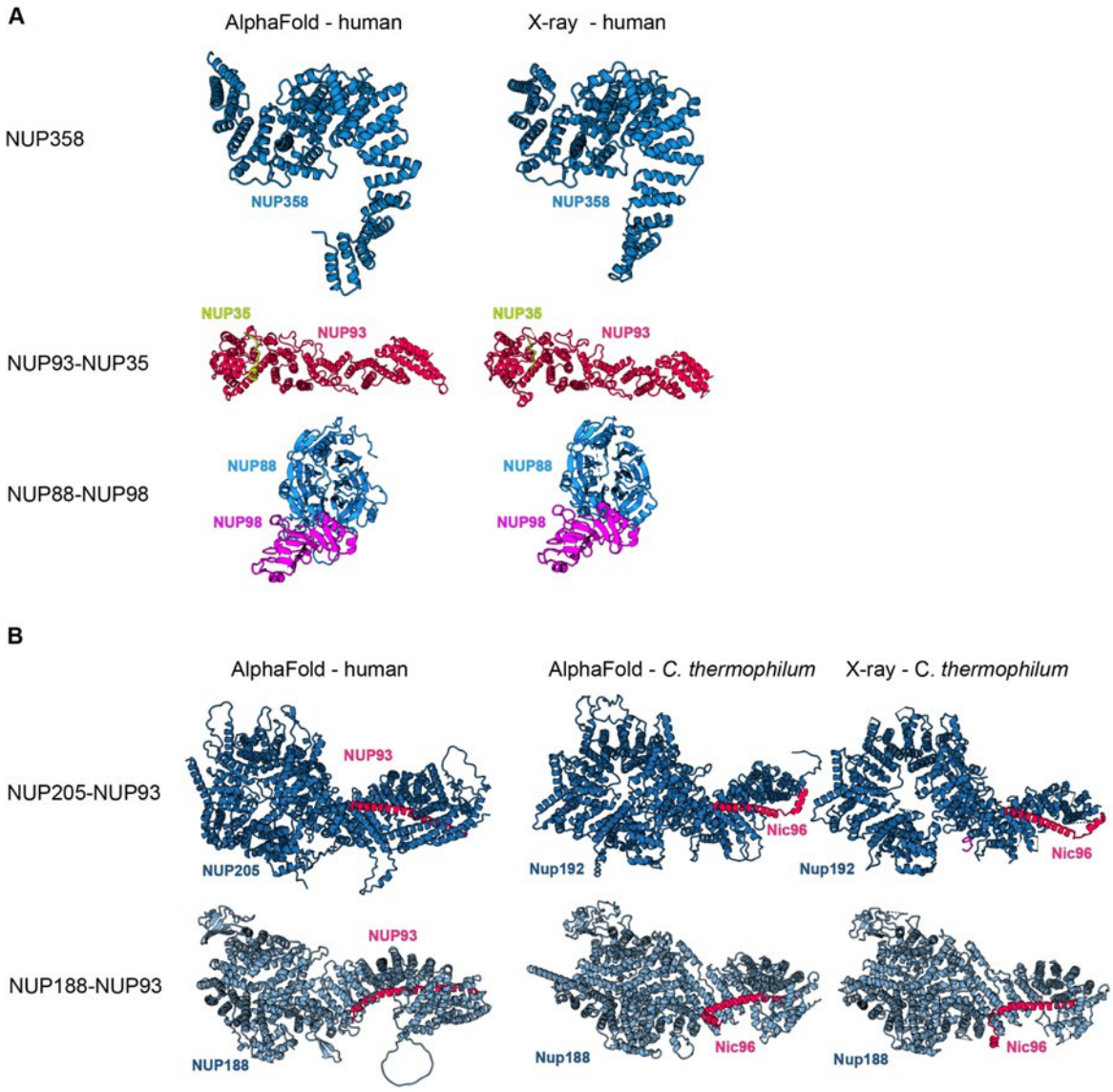


Fig. S6 Validation of selected AlphaFold models with X-ray crystallography structures. The models are compared to X-ray crystallography structures that have been released in the accompanying work (to be published and referenced here upon acceptance) and have not been used as templates for modeling. **(A)** Models are compared to structures of subcomplexes from human. **(B)** Models of NUP205-NUP93 and NUP188-NUP93 complexes are compared to X-ray structures from *C. thermophilum*, as X-ray structures for human are not available. The structures exhibit some differences due to inter-species variation but harbor the same interaction sites. For comparison, AlphaFold models of *C. thermophilum* are also shown, and recapitulate the X-ray structures. Proteins are colored as in **Fig. 1**.

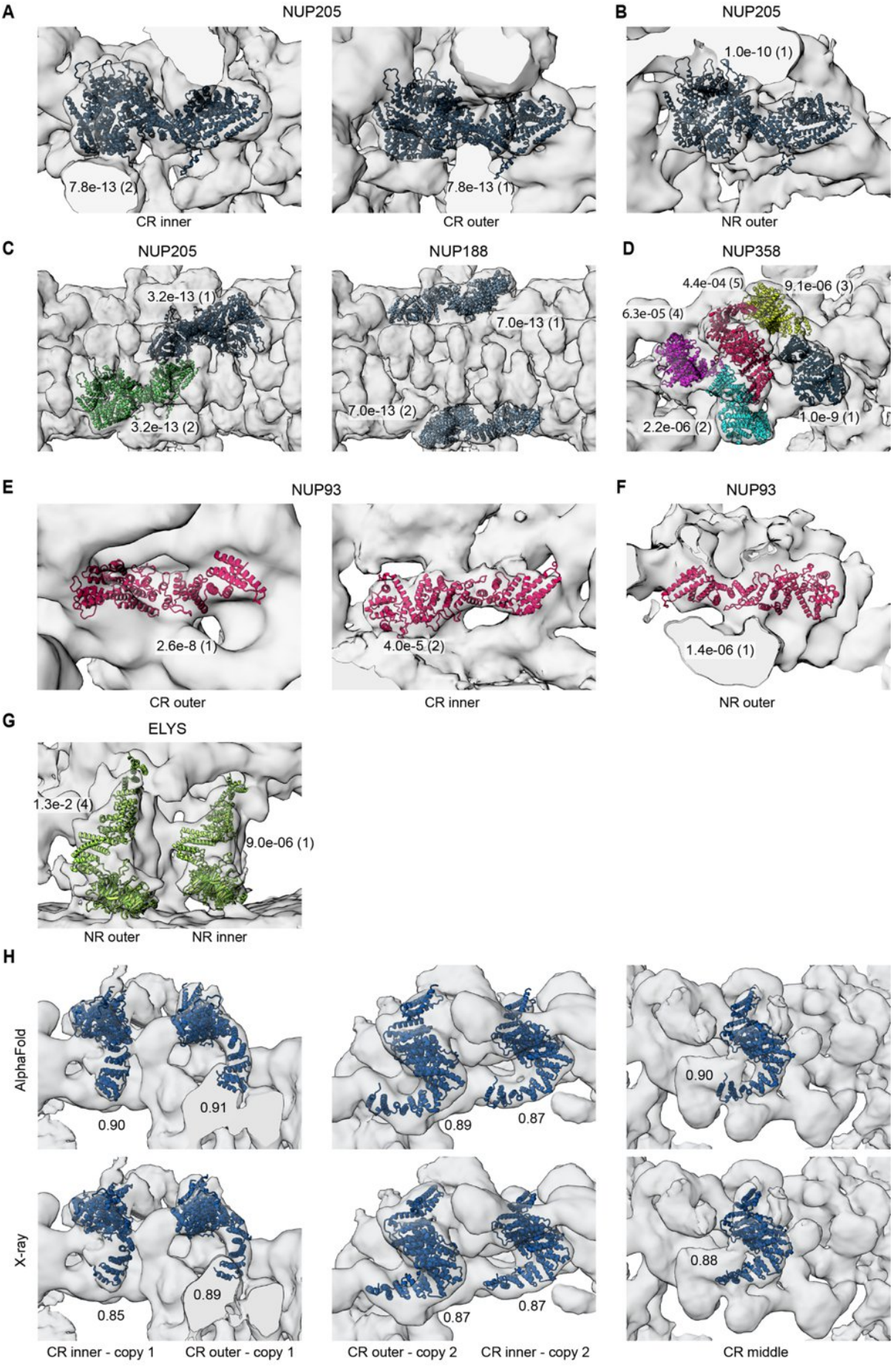

A
NUP205
7.8e-13 (2)
CR inner
7.8e-13 (1)
CR outer
B
NUP205
1.0e-10 (1)
NR outer
C
NUP205
3.2e-13 (1)
3.2e-13 (2)
NUP188
7.0e-13 (1)
7.0e-13 (2)
D
NUP358
4.4e-04 (5)
9.1e-06 (3)
6.3e-05 (4)
2.2e-06 (2)
1.0e-9 (1)
E
NUP93
2.6e-8 (1)
CR outer
4.0e-5 (2)
CR inner
F
NUP93
1.4e-06 (1)
NR outer
G
ELYS
1.3e-2 (4)
9.0e-06 (1)
NR outer
NR inner
H
AlphaFold
0.90
0.91
0.89
0.87
0.90
X-ray
0.85
0.89
0.87
0.87
0.88
CR inner - copy 1
CR outer - copy 1
CR outer - copy 2
CR inner - copy 2
CR middle

Fig. S7 Systematic fitting of monomeric AlphaFold models of NUPs previously not included in the human NPC model or till now not localized unambiguously in the human NPC. Each panel shows the top fits fitted to the EM maps. P-values, calculated as described in the Methods, and ranks of the fits are indicated next to the structures. **(A)** NUP205 fitted to the map of the CR assigned two copies of NUP205 in the CR **(B)** NUP205 fitted to the map of the NR assigned one copy of NUP205 in the NR **(C)** NUP188 and NUP205 fitted to the map of the IR assigned NUP188 to the outer subcomplexes of the IR and NUP205 to the inner subcomplexes of the IR. **(D)** NUP358 was fitted to the EM map of the entire CR spoke assigning five fits, and thus five copies, of NUP358 per NPC spoke (*left)*. **(E and F)** NUP93 fitted to the difference maps of the CR and NR obtained after subtracting the density of Y complexes identifies two locations of NUP93 in CR and one in the NR. **(G)** NUP93 fitted to the difference map of the NR confirms two copies of ELYS in the NR. **(H)** AlphaFold model of NUP358 fits the EM density with high cross-correlation scores than the NUP358 X-ray structures (only the better fitting of the two X-ray structure is shown). Numbers next to the fits indicate cross-correlation with the EM density as calculated with UCSF Chimera.

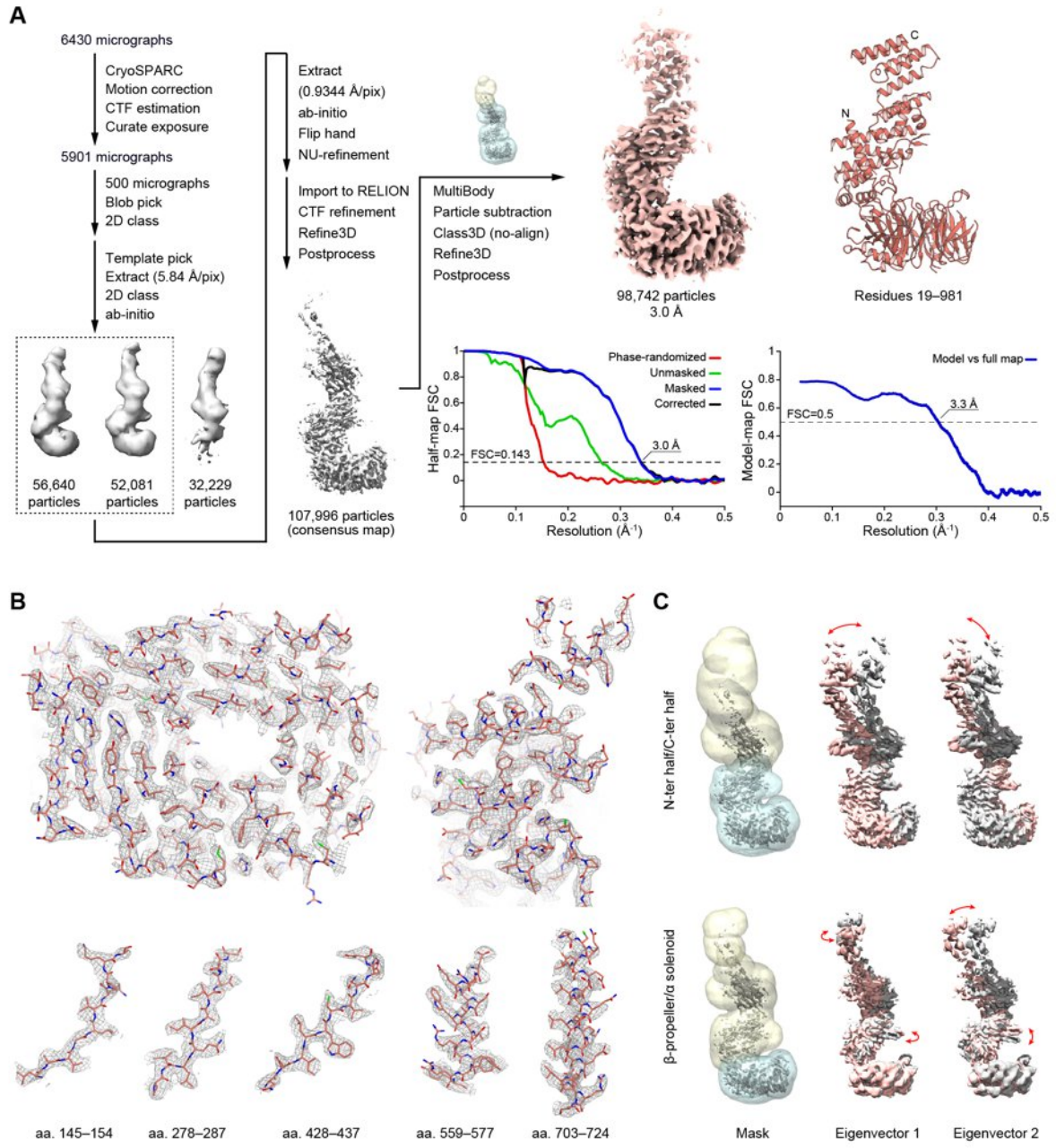


Fig. S8 Single particle cryo-EM analysis of human NUP155. (A) The data processing workflow depicted left resulted in the respective map and model. Fourier shell correlation (FSC) curves are shown for two half-maps and map-to-model. **(B)** Cryo-EM density maps of human NUP155. Sections through the β-propeller domain and the α helical domain are shown with representative density maps of individual β strands and α helices. **(C)** Multibody refinement and principal component analysis of the consensus map depicted in (A). Different sets of masks were used. Motions represented by the first and second eigenvectors are shown.

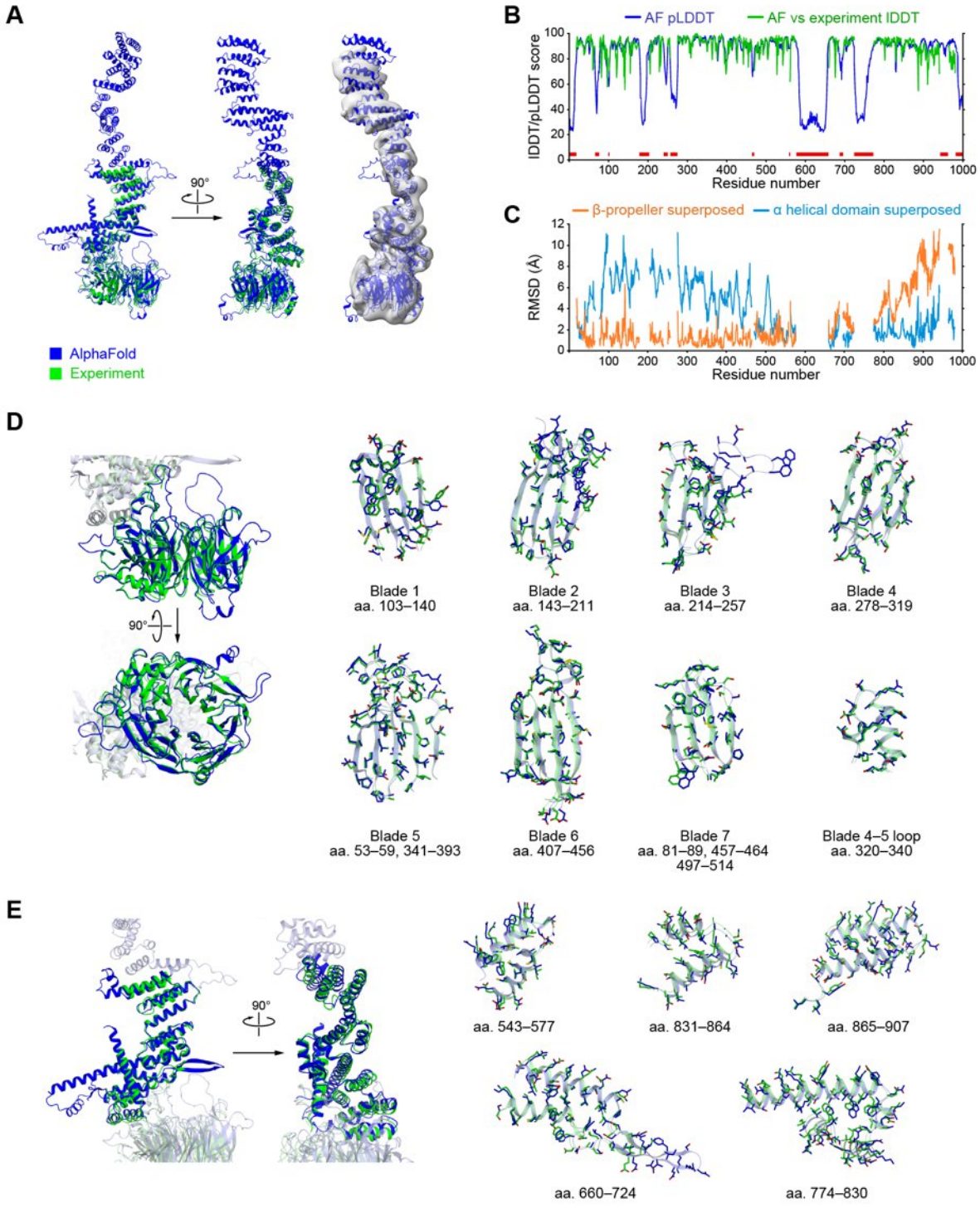


Fig. S9 Structural comparison between the experimentally determined human NUP155 structure and the AlphaFold model. (A) The AlphaFold model (blue) superimposed onto the whole chain of the experimentally determined structure (green) and fitted into the cryo-EM consensus map (lowpass filtered to 10 Å). Long extended loops (residues 1–10, 585–653) are omitted from the AlphaFold model for clarity. **(B)** Plot of the pLDDT score of the AlphaFold model and the lDDT score between the AlphaFold model and the experimentally determined structure (scaled from 0 to 100). Red bars indicate regions that remain unresolved in the experimentally determined model. **(C)** Per-residue RMSD values between both models. The AlphaFold model was superimposed onto the β-propeller domain (residues 41–514, orange) and α helical domain (residues 19–40 and 515–980, blue) of the experimentally determined

model, and per-residue RMSD values were calculated using UCSF Chimera. **(D)** Structural comparison of the β-propeller domain. The AlphaFold model was superimposed onto the β-propeller domain (residues 41–514), and the structures of individual blades are compared. **(E)** Structural comparison of the α helical domain. The AlphaFold model was superimposed onto the α helical domain (residues 19–40 and 515–980), and the structures are compared for representative helices. In (C) and (D), figures are colored as in (A).

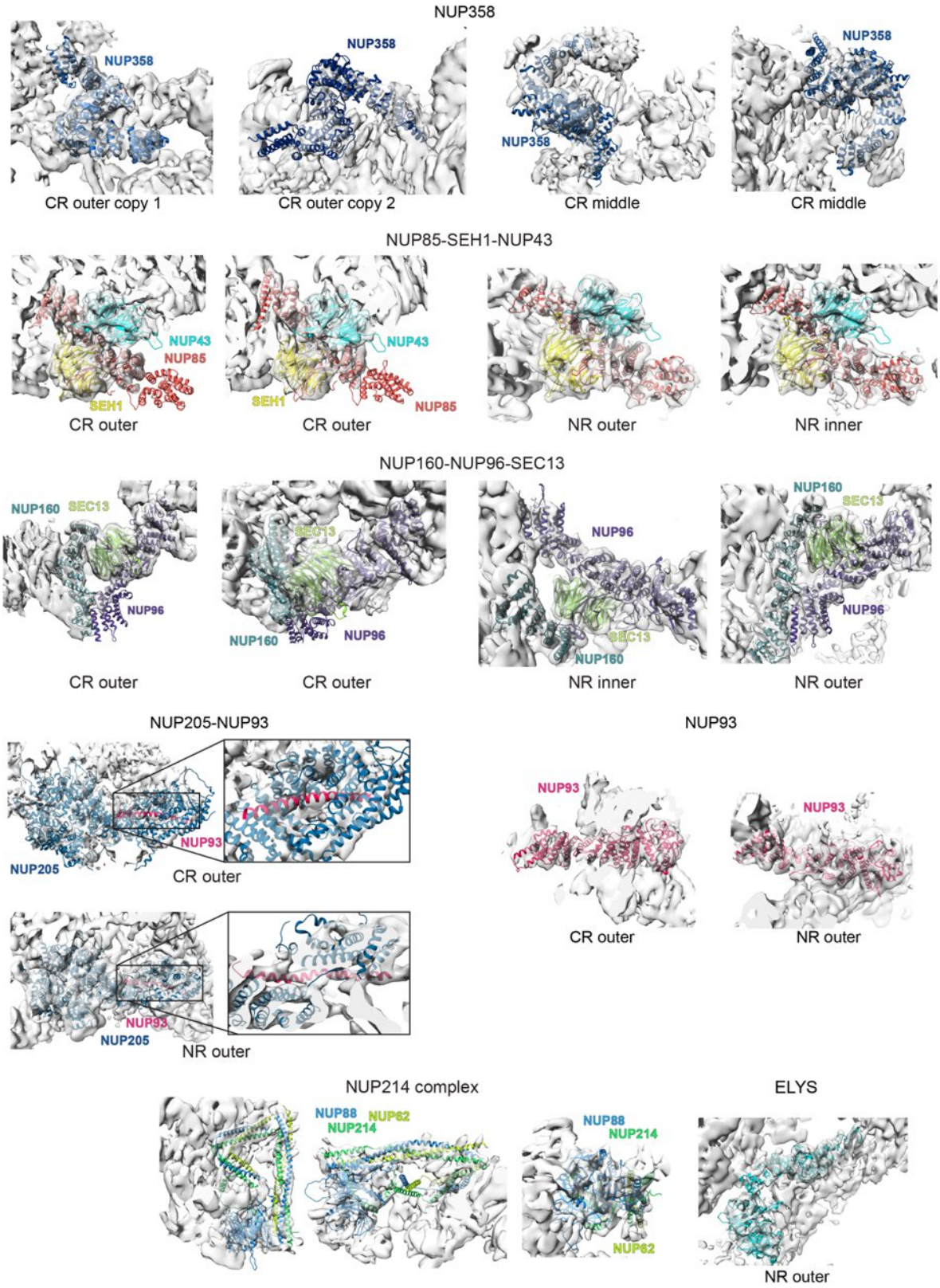
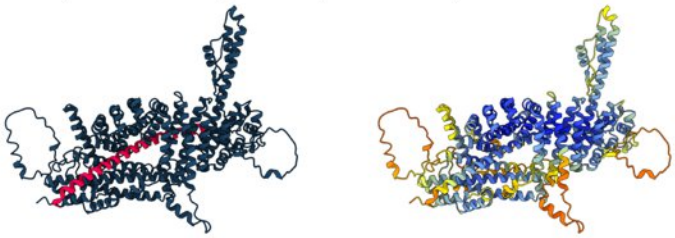
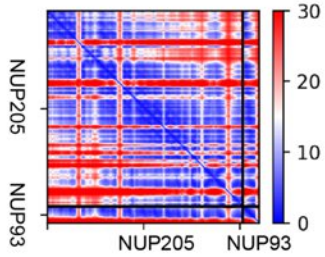


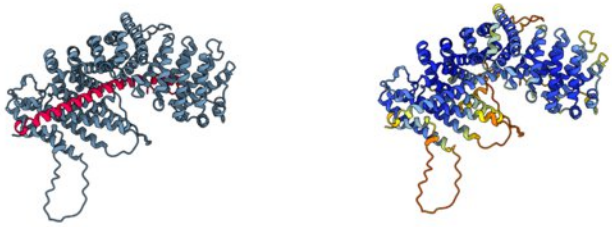
Fig. S10 Consistency of secondary structure observed in cryo-EM maps of the *Xenopus laevis* NPC (*33*, *111*), with AI-based models of human NUPs at the respective fit positions into the human NPC. AI-based structure predictions are fitted into the EM maps of the *Xenopus laevis* NPC (EMD-0910, EMD-0911, EMD-31891). The quality of the AlphaFold models can be appreciated by their match to secondary structure observed in the EM maps. Proteins are colored as in Fig. 1.

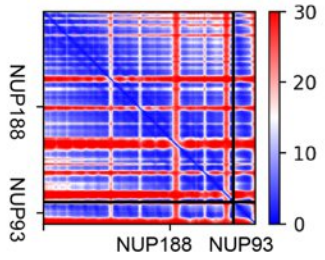
NUP205(aa.1001- 2012)-NUP93(aa. 156-176)



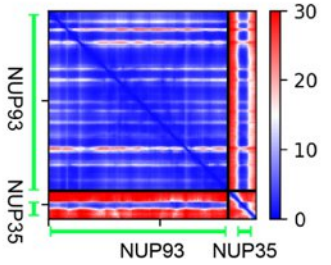


NUP188(aa.1001-1749)-NUP93(aa.156-176)

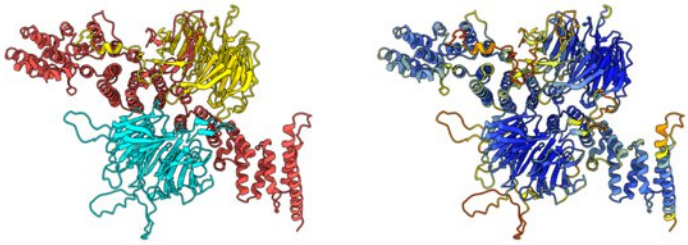


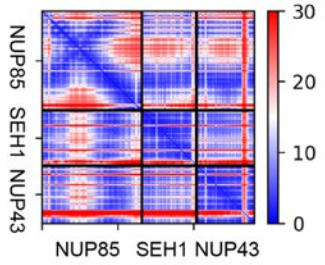


NUP93(aa. 182-502)-NUP35(aa.86-99)

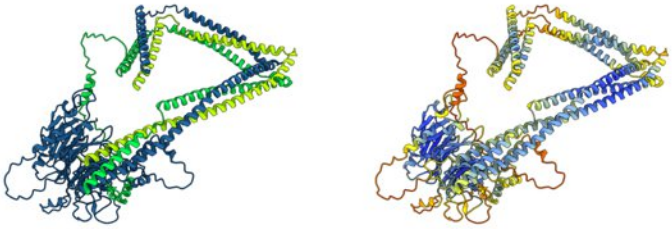


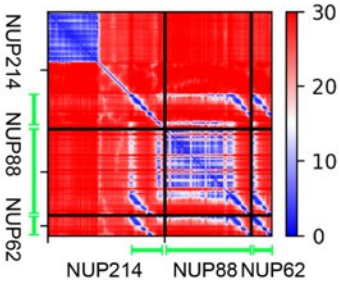
NUP85(aa. 1-656)-SEH1-NUP43



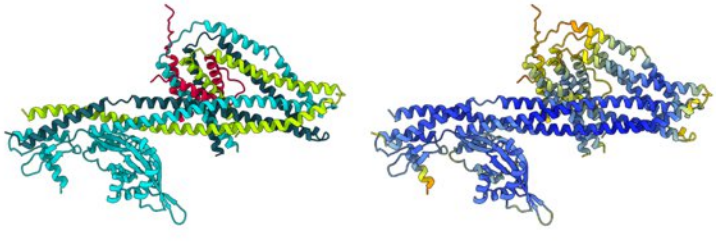


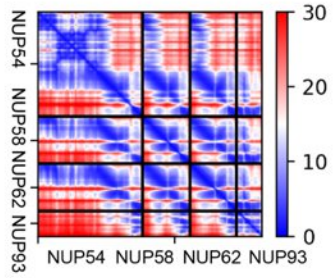
NUP214(aa. 1-428,700-1000)-NUP88-NUP62(aa. 332-502)



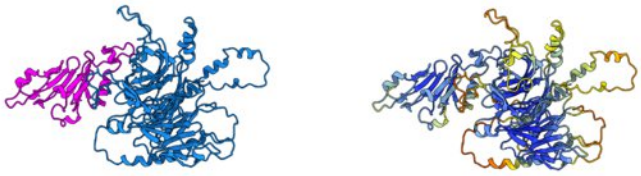


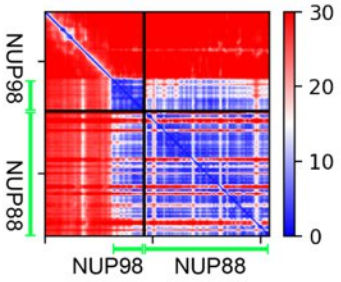
NUP62(aa. 332-502)-NUP58(aa. 246-418)-NUP54(aa. 111-493)-NUP93(aa. 1-94)



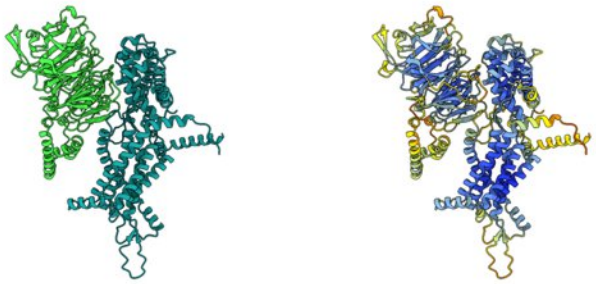


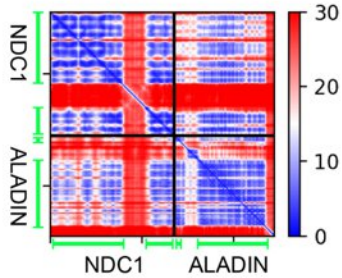
NUP98(aa. 730-880)-NUP88(aa. 1-579)



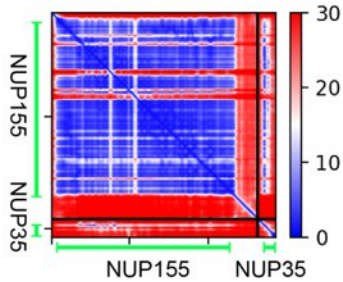


NDC1(aa. 1-391, 516-674)-ALADIN(aa. 2-31,134-486)

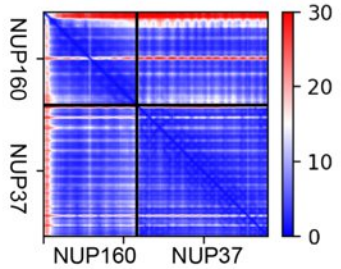




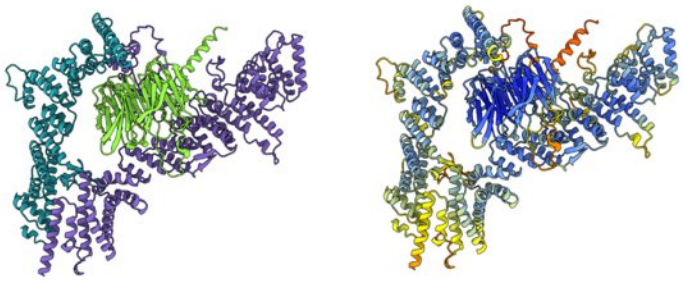
NUP155(aa. 40-577)-NUP35(aa. 285-326)

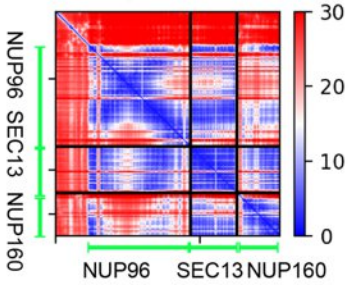


NUP160(aa. 875-1069)-NUP37

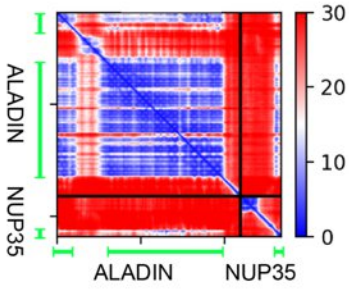


NUP160(aa. 1017-1436)-NUP96(aa. 226-937)-SEC13

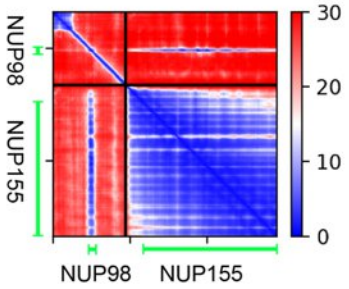




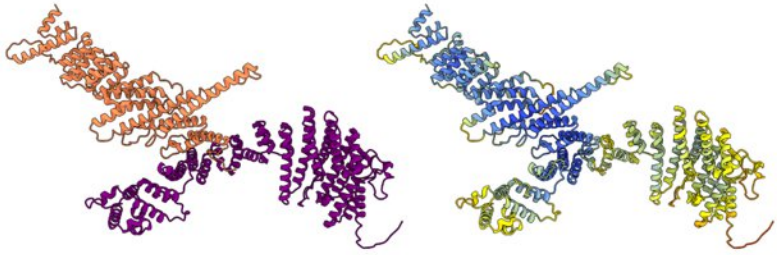
ALADIN(aa. 2-31,134-486)-NUP35(aa. 307-326)

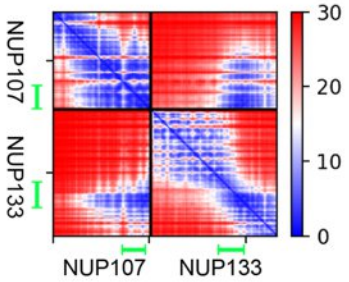


NUP155(aa. 1001-1391)-NUP98(aa.592-613)

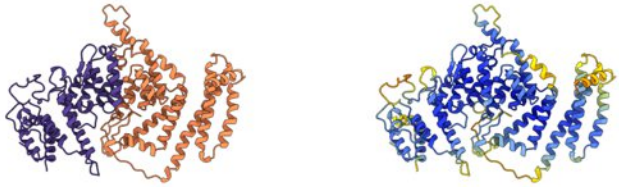


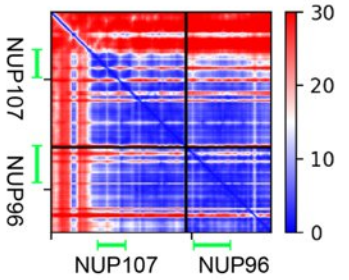
* NUP133(aa. 506-1156)-NUP107(aa. 250-925)





* NUP96(aa. 361-660)-NUP107(129-480)





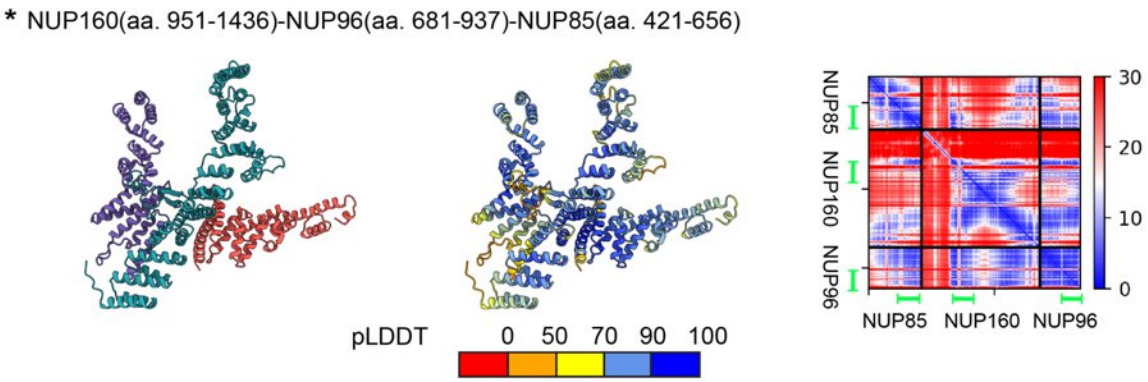


Fig. S11 Estimated quality of structural models of human NUP subcomplexes built using AlphaFold/ColabFold. Each row shows from left to right: model colored by the NUP color code as in Fig. 1 (*left*); model colored by local confidence estimated with predicted local distance difference test (pLDDT), as returned by AlphaFold. The pLDDT > 90 (dark blue) indicates high estimated accuracy of backbone and side chain rotamers whereas pLDDT > 70 (yellow) indicates confident backbone prediction (*61*) (*middle*); The confidence of inter-domain and inter-chain orientations estimated with the expected distance error in Å between all pairs of residues in the complex, as returned by ColabFold. The color at each (x, y) position of the matrix corresponds to the expected distance error in residue x's position, when the prediction and true (unknown) structure are aligned on residue y. Blue indicates low error (*right*). The heat maps show the sequence covered in the initial prediction, while green bars in the heat map indicate the sequence coverage of the ribbon models (left) that was subsequently used. The sequence coverage of the ribbon models is indicated for each protein in brackets. aa. – amino acid residues. The asterisks indicate that the given model was used only as a source of elastic network restraints (see Methods).

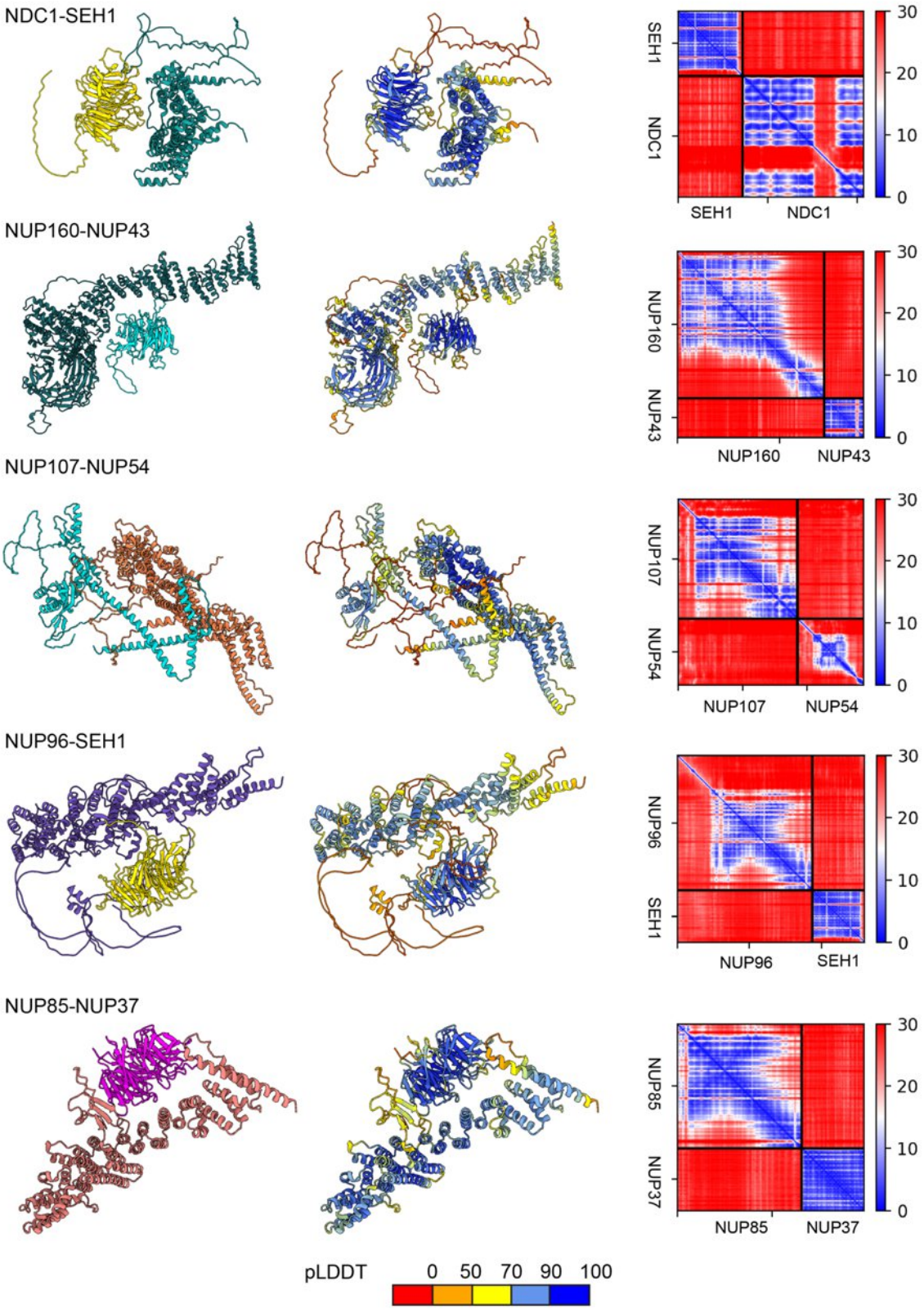


Fig. S12 A negative control set of AlphaFold/ColabFold predictions for NUP subcomplex structures. Various pairs of NUPs not expected to interact (based on their distance in the current NPC models) were modeled. The protein chains form only loose contacts and their relative orientation is not confidently predicted as expected for true negative pairs (red areas in the plots on the right). Each row shows from left to right: (*left*) model colored by the NUP

color code as in Fig. 1; (*middle*) model colored by local confidence estimated with predicted local distance difference test (pLDDT), as returned by AlphaFold. The pLDDT > 90 (dark blue) indicates very high estimated accuracy of backbone and side chain rotamers whereas pLDDT > 70 (yellow) indicates confident backbone prediction (*61*); (*right*). The confidence of inter-domain and inter-chain orientations estimated with the expected distance error in Å between all pairs of residues in the complex, as returned by ColabFold. The color at each (x, y) position of the matrix corresponds to the expected distance error in residue x's position, when the prediction and true (unknown) structure are aligned on residue y. Blue indicates low error.

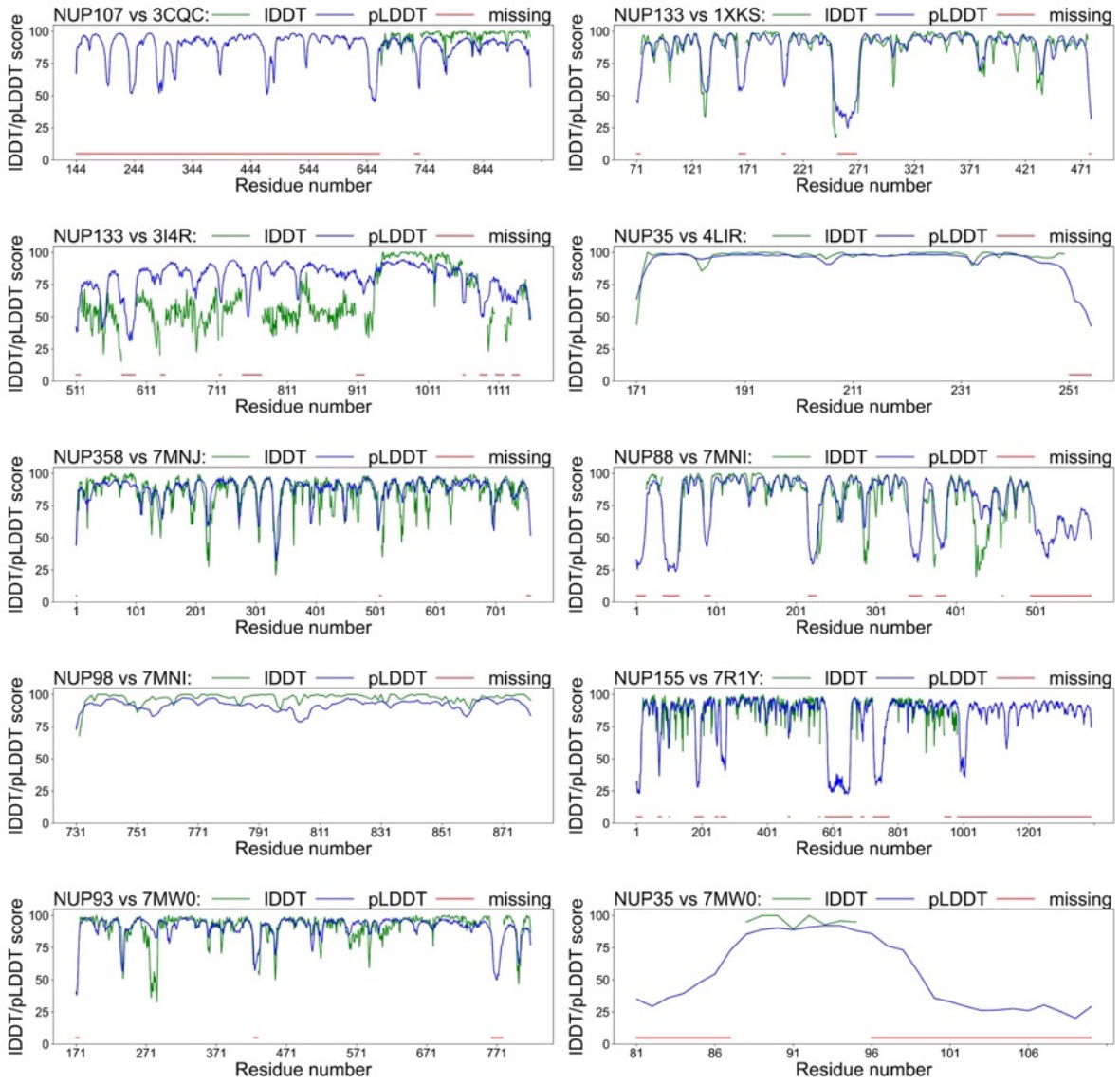


Fig. S13 Comparison of predicted lDDT scores (blue) with the lDDT scores calculated by comparing models to the available structures of human NUPs (green). Regions missing in the reference structures are indicated in red. The predicted and measured lDDT scores are generally above 90 in the structured regions, while low values correspond to loops and regions missing in the reference structures. The low lDDT values in the comparison of the NUP133 with the 3I4R structure likely corresponds to the low quality this structure (with resolution of 3.53 Å and very poor structure quality metrics). The short fragment of NUP35 in the lower right panel corresponds to the SLiM binding to NUP93, while the poorly scoring flanking regions are disordered linkers flanking this SLiM. The pLDDT scores are also reported as a B-factor of the models deposited in the Protein Data Bank (PDB ID: 7R5K and 7R5J).

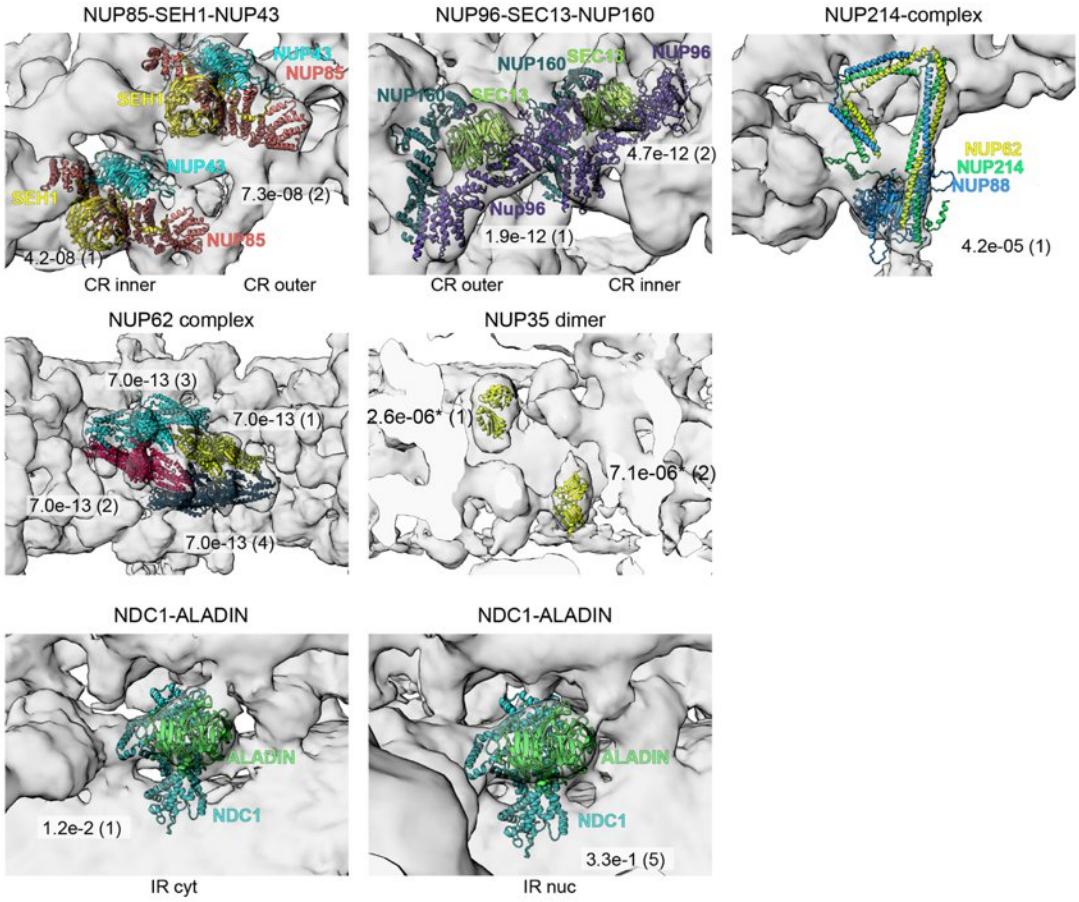
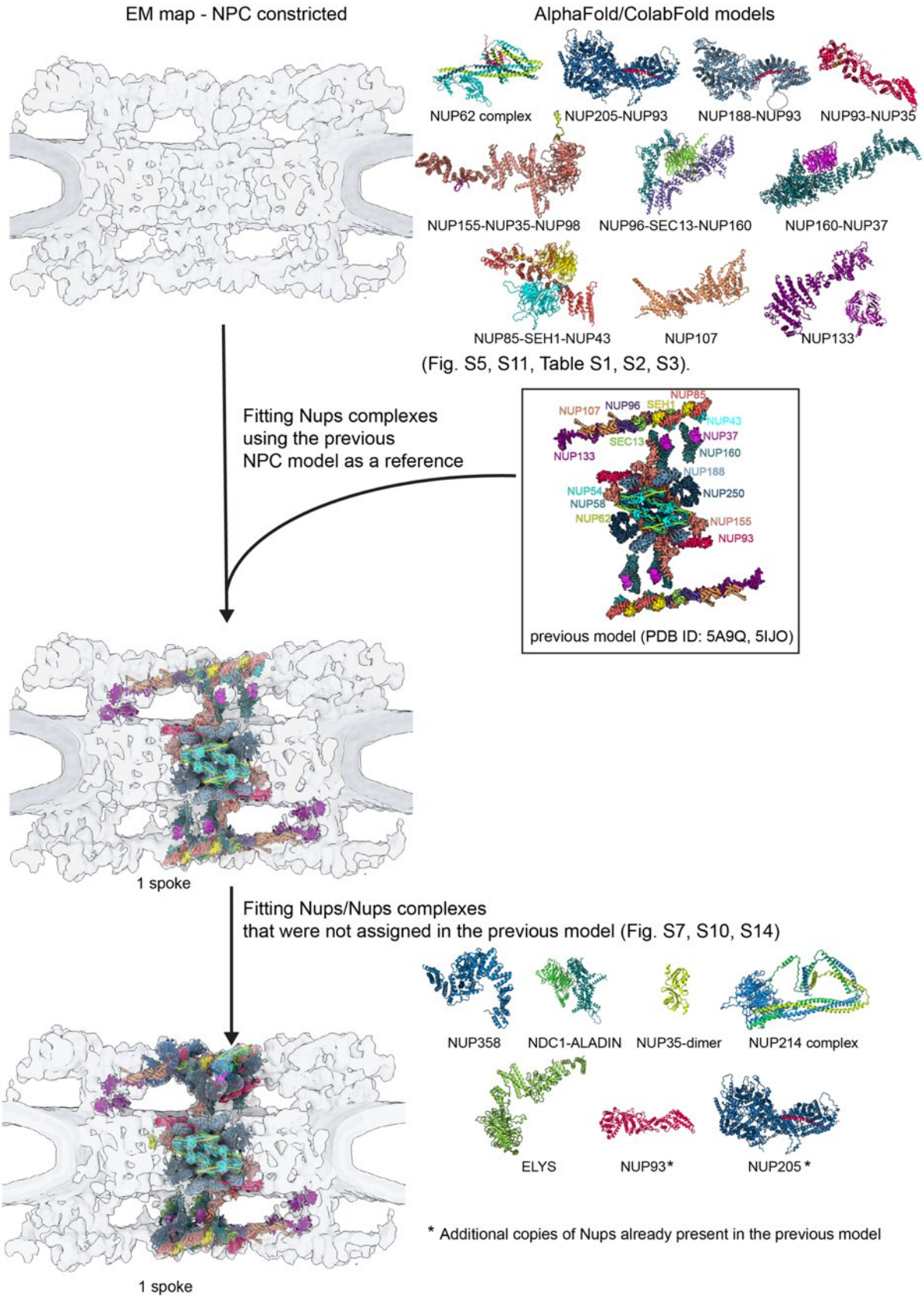
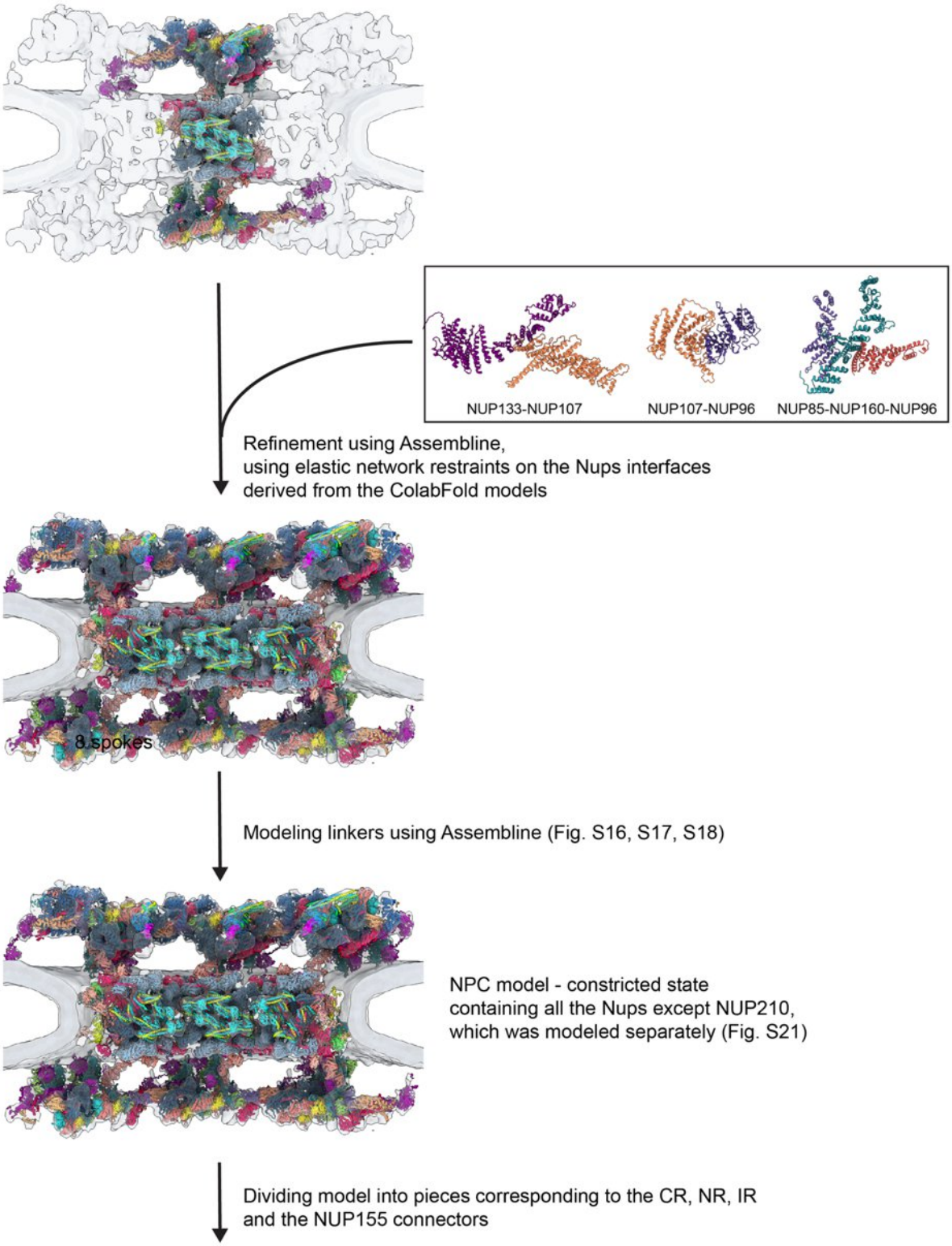
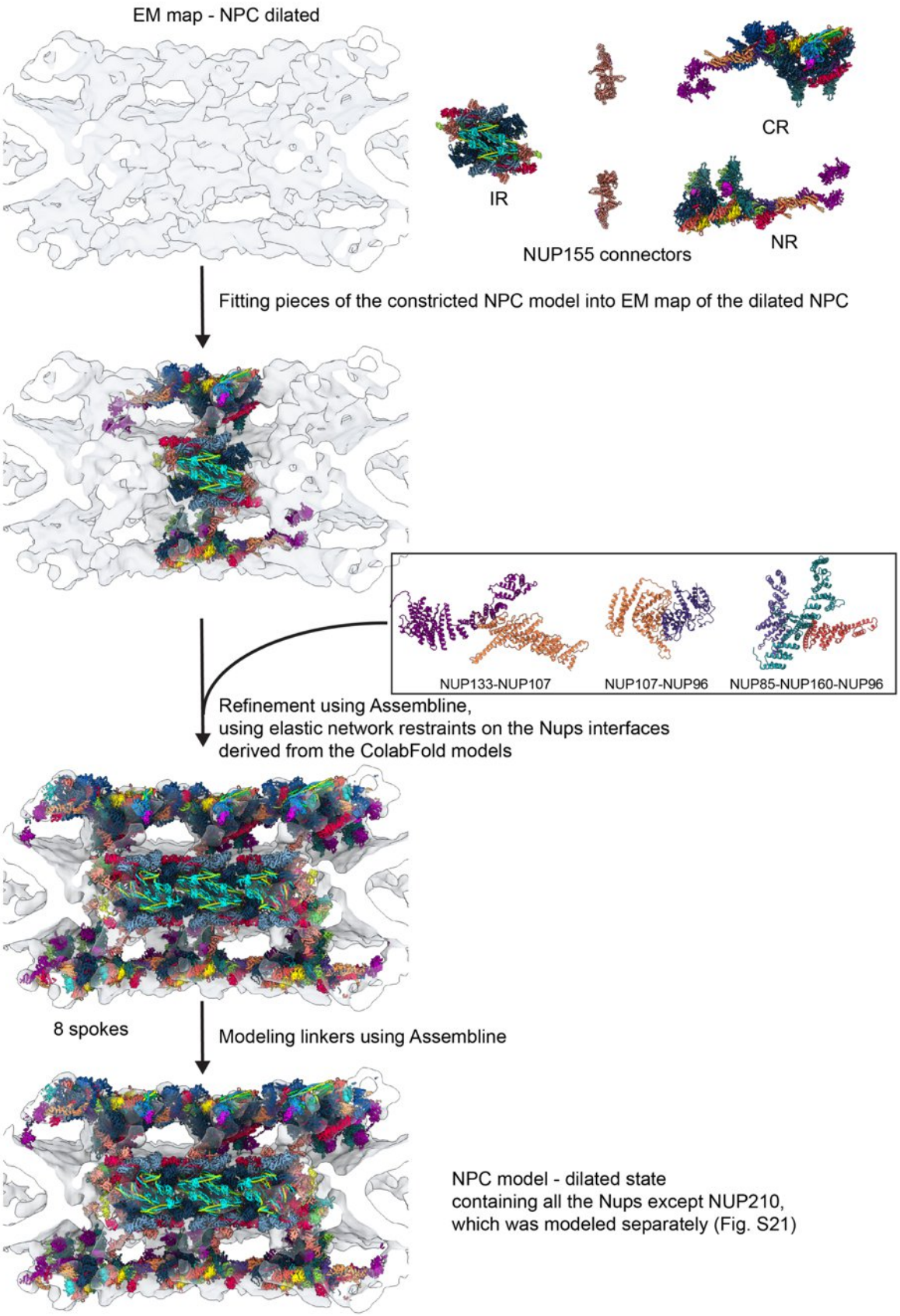


Fig. S14 Systematic fitting of AlphaFold subcomplex models into the hNPC EM maps. Each panel shows the top fits fitted to the EM maps with proteins are colored as in Fig. 1. P-values, calculated as described in the Methods, and ranks of the fits are indicated next to the structures. All subcomplexes have been fitted to the full CR and IR maps, except NDC1-ALADIN and NUP35 dimer, which were fitted to the difference map of the IR. The asterisk next to the NUP35 dimer P-values indicates uncertainty in the P-value assessment in this case due to the non-normal distribution of the fits and a long tail of P-values better than 0.05.


EM map - NPC constricted
AlphaFold/ColabFold models
NUP62 complex
NUP205-NUP93
NUP188-NUP93
NUP93-NUP35
NUP155-NUP35-NUP98
NUP96-SEC13-NUP160
NUP160-NUP37
NUP85-SEH1-NUP43
NUP107
NUP133
(Fig. S5, S11, Table S1, S2, S3).
Fitting Nups complexes
using the previous
NPC model as a reference
NUP107
NUP96
SEH1
NUP85
NUP43
NUP37
SEC13
NUP133
NUP160
NUP188
NUP54
NUP58
NUP62
NUP250
NUP155
NUP93
previous model (PDB ID: 5A9Q, 5IJO)
1 spoke
Fitting Nups/Nups complexes
that were not assigned in the previous model (Fig. S7, S10, S14)
NUP358
NDC1-ALADIN
NUP35-dimer
NUP214 complex
ELYS
NUP93*
NUP205 *
* Additional copies of Nups already present in the previous model
1 spoke


NUP133-NUP107
NUP107-NUP96
NUP85-NUP160-NUP96
Refinement using Assembline,
using elastic network restraints on the Nups interfaces
derived from the ColabFold models
8 spokes
Modeling linkers using Assembline (Fig. S16, S17, S18)
NPC model - constricted state
containing all the Nups except NUP210,
which was modeled separately (Fig. S21)
Dividing model into pieces corresponding to the CR, NR, IR
and the NUP155 connectors


EM map - NPC dilated
CR
IR
NR
NUP155 connectors
Fitting pieces of the constricted NPC model into EM map of the dilated NPC
NUP133-NUP107
NUP107-NUP96
NUP85-NUP160-NUP96
Refinement using Assembline,
using elastic network restraints on the Nups interfaces
derived from the ColabFold models
8 spokes
Modeling linkers using Assembline
NPC model - dilated state
containing all the Nups except NUP210,
which was modeled separately (Fig. S21)

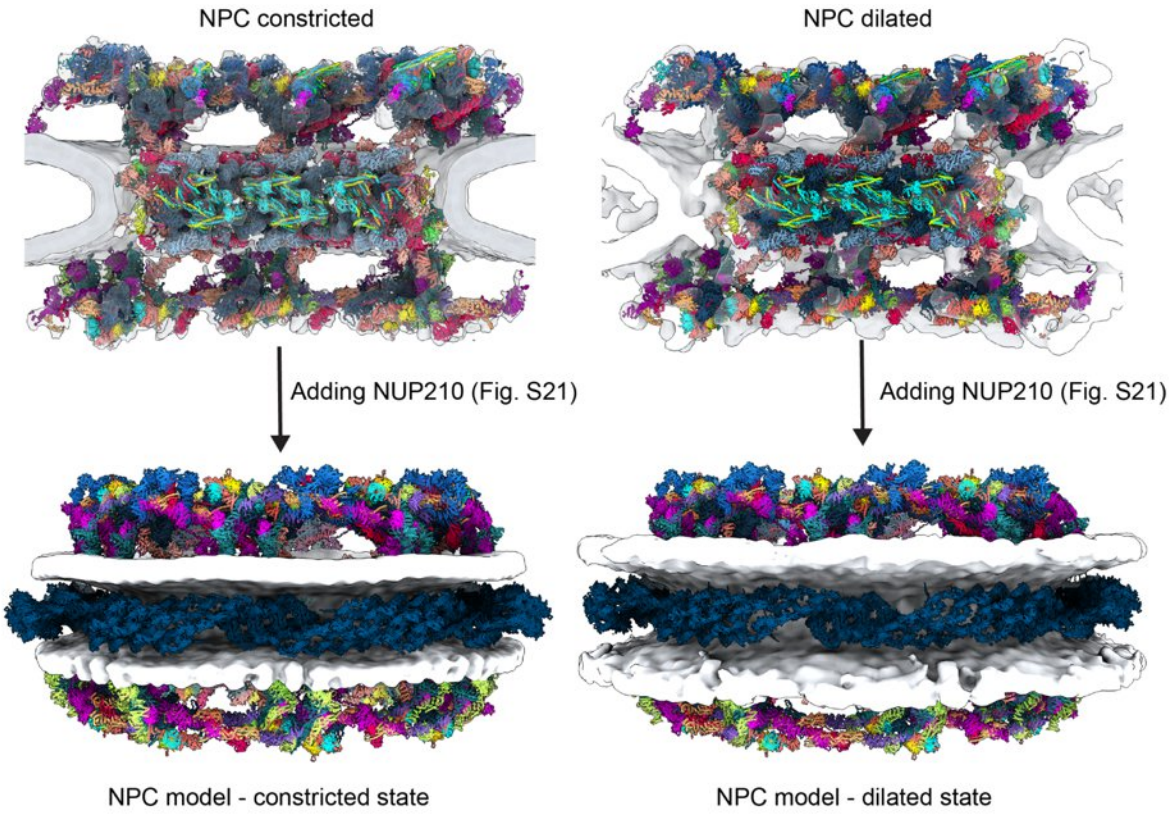
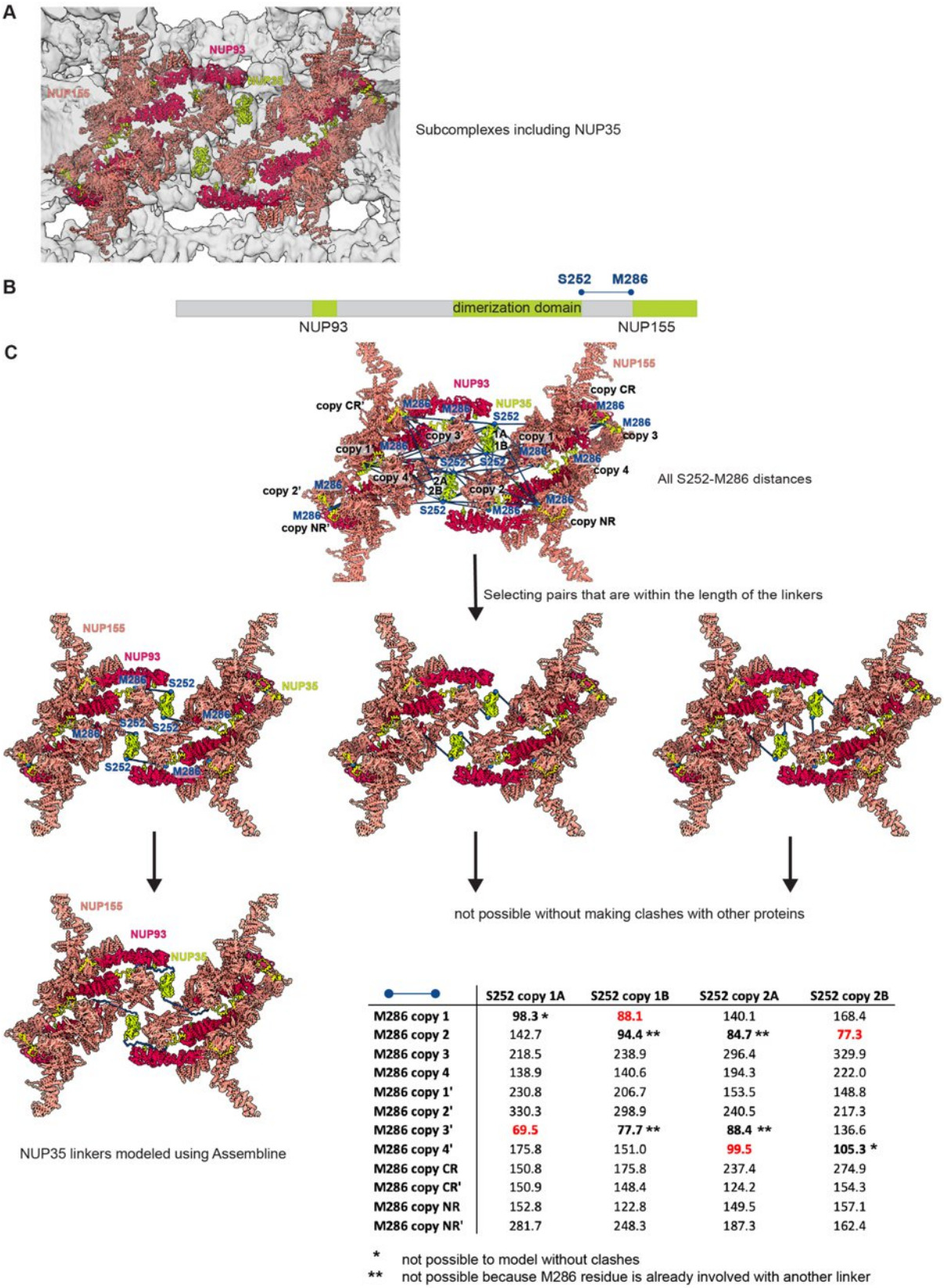
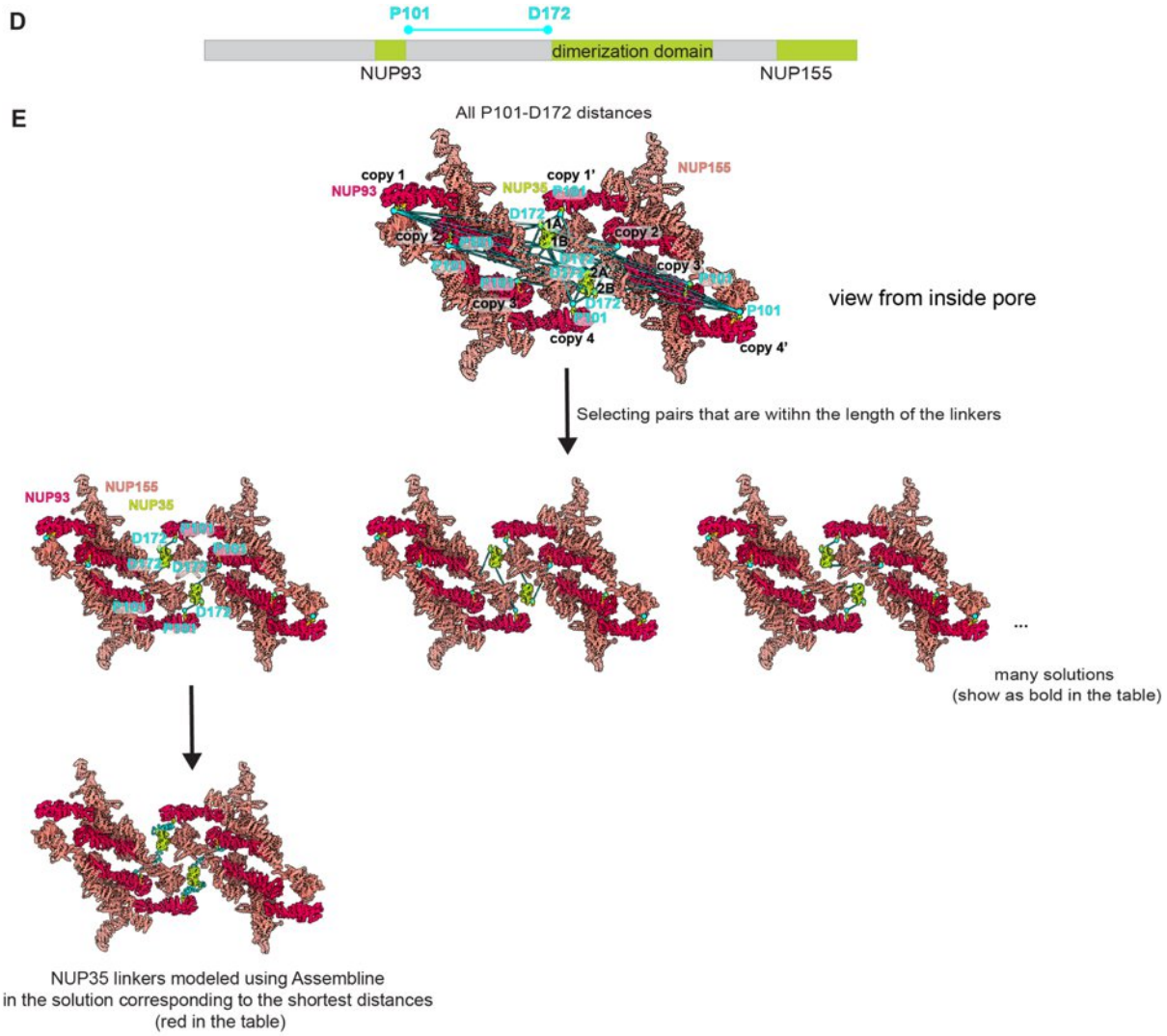


Fig. S15 Overview of the modeling procedure. The structural models are shown in cartoon representation and colored as in Fig. 1. The cryo-ET map is shown as a gray surface. The supplementary figures and tables with additional information about the models and more details about the modeling procedures are indicated at the respective steps.



	S252 copy 1A	S252 copy 1B	S252 copy 2A	S252 copy 2B
M286 copy 1	98.3 *	88.1	140.1	168.4
M286 copy 2	142.7	94.4 **	84.7 **	77.3
M286 copy 3	218.5	238.9	296.4	329.9
M286 copy 4	138.9	140.6	194.3	222.0
M286 copy 1'	230.8	206.7	153.5	148.8
M286 copy 2'	330.3	298.9	240.5	217.3
M286 copy 3'	69.5	77.7 **	88.4 **	136.6
M286 copy 4'	175.8	151.0	99.5	105.3 *
M286 copy CR	150.8	175.8	237.4	274.9
M286 copy CR'	150.9	148.4	124.2	154.3
M286 copy NR	152.8	122.8	149.5	157.1
M286 copy NR'	281.7	248.3	187.3	162.4

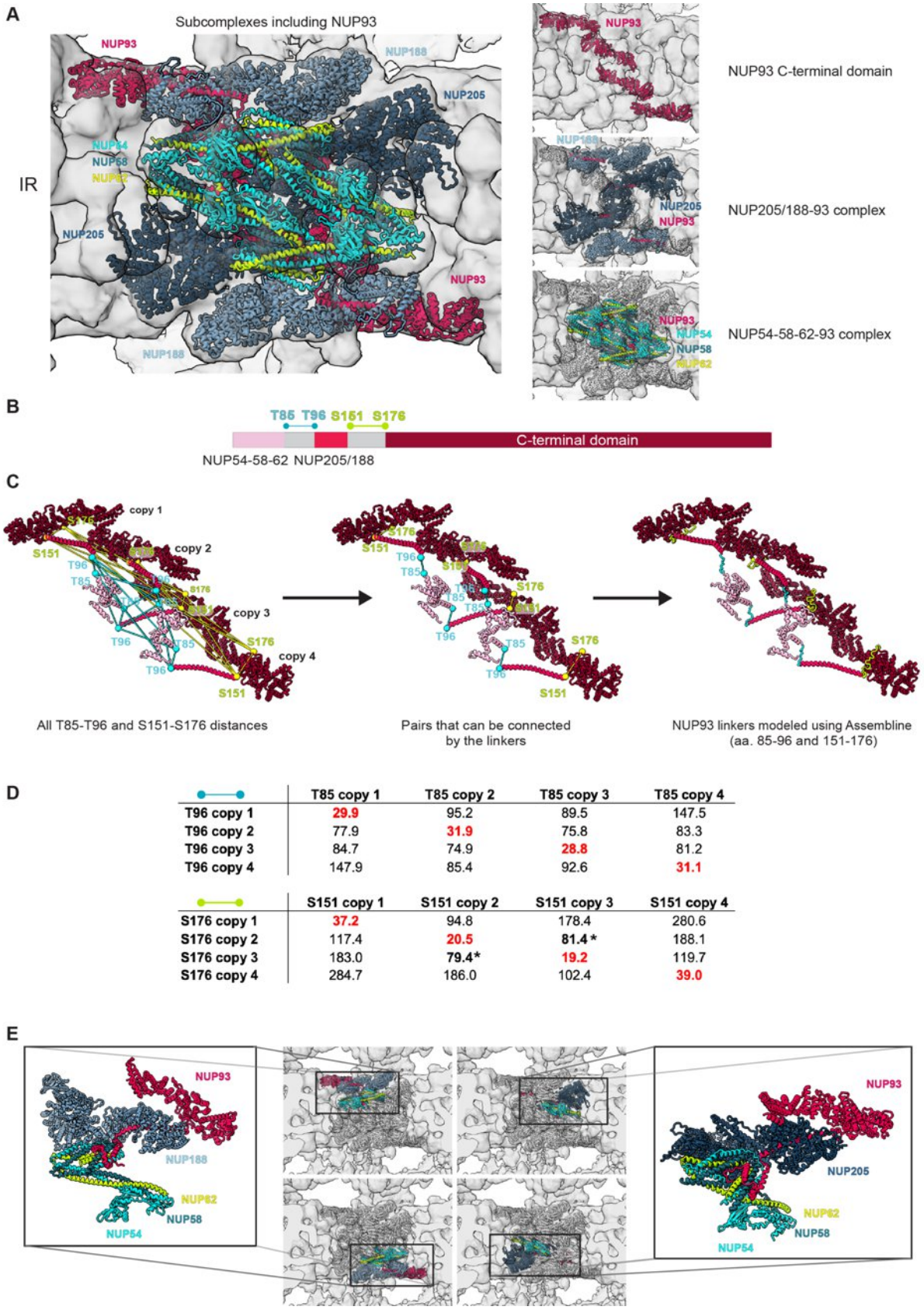
* not possible to model without clashes
** not possible because M286 residue is already involved with another linker



	P172 copy 1A	P172 copy 1B	P172 copy 2A	P172 copy 2B
P101 copy 1	250.6	262.5	333.2	355.0
P101 copy 2	**162.0**	**166.3**	**235.7**	255.2
P101 copy 3	**101.3**	**76.4**	**119.8**	**131.7**
P101 copy 4	**141.2**	**106.1**	**61.6**	**48.2**
P101 copy 1'	**45.2**	**62.5**	**108.9**	**144.0**
P101 copy 2'	**133.6**	**119.8**	**70.0**	**90.9**
P101 copy 3'	262.0	**242.7**	**170.2**	**161.7**
P101 copy 4'	351.0	331.1	260.8	248.7

Fig. S16 Modeling procedure used to trace linkers of NUP35. NUP35 links to NUP93 and NUP155 with its N- and C-terminal linker, respectively. This has been established by previous studies in fungi (*10–12*), the biochemical analysis of the human proteins (*32*), and is detected by ColabFold (**Fig. S11**). **(A)** Subcomplexes containing NUP35 in their positions after the refinement using Assembline are shown. **(B)** Domain organization of NUP35 showing regions of NUP35 interacting with NUP93 (aa. 86-101) and NUP155 (aa. 287-300), and the C-terminal linker connecting residues Ser 252 and Met 286 (dark blue). **(C)** Linker modeling procedure. All distances between Cα atoms of the indicated residues are shown in the tables below (in Å). Pairs of residues that are within a maximum length of the S252-M286 linker (34*3.5 Å = 119 Å) are shown in bold. The connections that were impossible to reconstruct without clashes with other proteins or because one of the residues was already involved in another linker are indicated with asterisks. The linker at the C-terminus of NUP35 reaches to the adjacent copies of NUP155 in only one geometrically possible configuration. NUP35 therefore bridges across

spokes. **(D)** Domain organization of NUP35 showing the N-terminal linker connecting residues Pro 101 and Asp 172 (cyan). The N-terminal linker is 70 aa long and therefore less restrained. **(E)** Linker modeling procedure. Since in this case there are many pairs of Pro 101 and Asp 172 that are within the maximum length of the linker (71*3.5 Å = 248.5 Å), it is not possible to unambiguously determine which copy of NUP93 is connected to which copy of NUP35 dimer. We thus modeled the linkers based on the shortest linkers criterium, that resulted in connecting the NUP93 domains situated in adjacent spokes. We chose the connections corresponding to the shortest distances.



	T85 copy 1	T85 copy 2	T85 copy 3	T85 copy 4
T96 copy 1	29.9	95.2	89.5	147.5
T96 copy 2	77.9	31.9	75.8	83.3
T96 copy 3	84.7	74.9	28.8	81.2
T96 copy 4	147.9	85.4	92.6	31.1

	S151 copy 1	S151 copy 2	S151 copy 3	S151 copy 4
S176 copy 1	37.2	94.8	178.4	280.6
S176 copy 2	117.4	20.5	81.4*	188.1
S176 copy 3	183.0	79.4*	19.2	119.7
S176 copy 4	284.7	186.0	102.4	39.0

Fig. S17 Modeling procedure used to trace the linkers at the N-terminus of NUP93 within the IR. The N-termini of four copies of NUP93 within the IR connect to either NUP205 or NUP188, and the NUP62 complex. **(A)** Subcomplexes containing NUP93 in their positions after the refinement using Assembline are shown. **(B)** Domain organization of NUP93 showing

the regions of NUP93 interacting with NUP62 complex (aa. 1-85, pink) and NUP205/NUP188 (aa. 96-151, red), and the two linkers to be modeled, connecting Thr 85 and Thr 96 (cyan), and Ser 151 and Ser 176 (green). **(C)** Linker modeling procedure. **(D)** All distances between Cα atoms of the indicated residues are shown (in Å). Pairs of residues that are within a maximum length of the linker (38.5 Å (11*3.5 Å) for T85-T96 and 87.5 Å (25*3.5 Å) for S151-S176 are shown in bold. The pairs selected for the final model are marked in red. The connecting pairs marked with asterisk are numerically slightly below the maximal distance and would extensively stretch and clash with neighboring proteins. The geometrically favored configuration is linkage within either of the four IR subcomplexes – the NUP93 linkers cannot reach across spokes. **(E)** Subcomplexes NUP54-58-62-NUP188/205-NUP93 in the IR organized by NUP93.

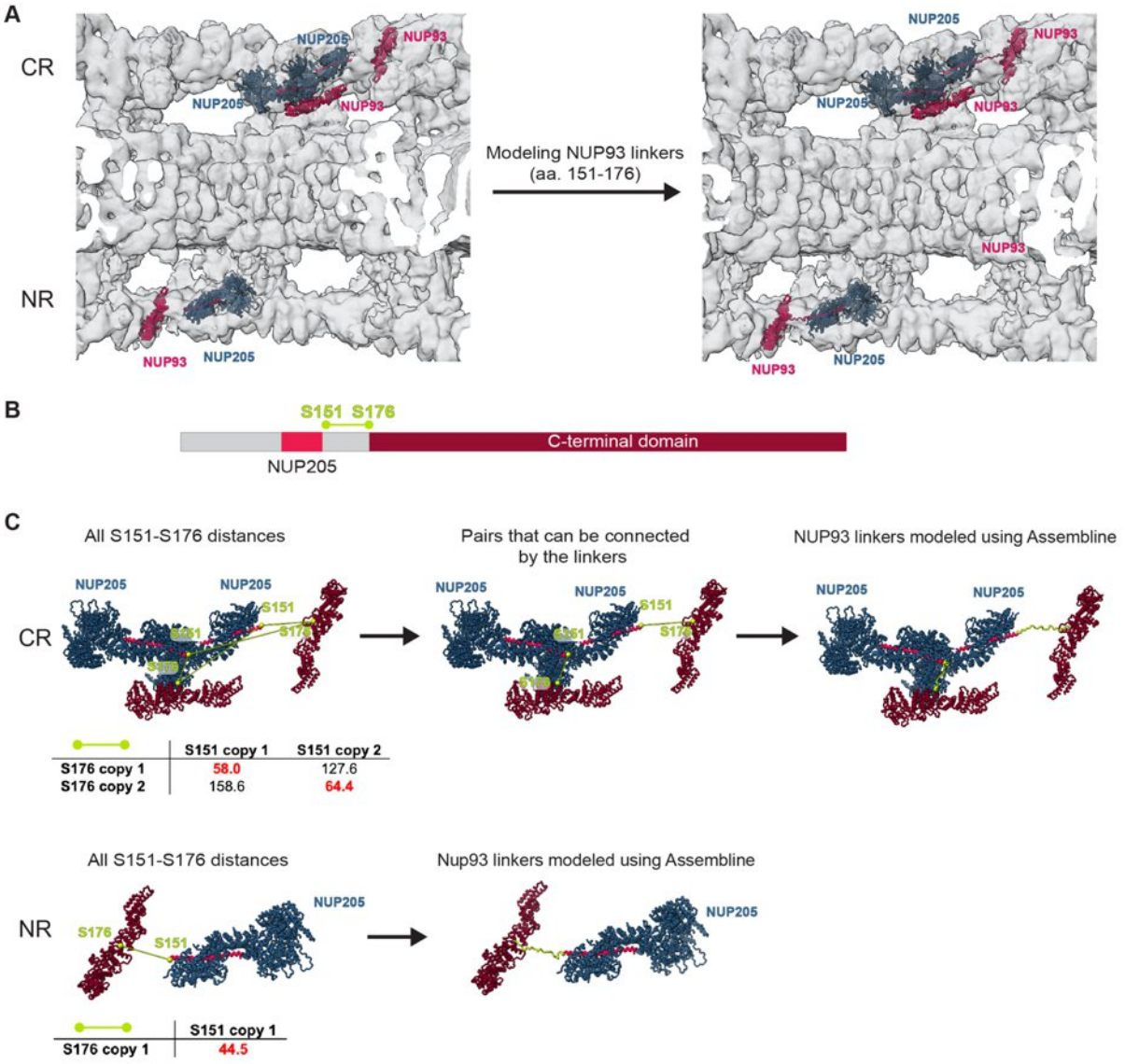


Fig. S18 Modeling procedure used to trace linkers at the N-terminus of NUP93 within the outer rings. (A) Subcomplexes containing NUP93 in their positions after the refinement using Assembline are shown before (left) and after the modeling the NUP93 linkers (aa. 151-176) **(B)** Domain organization of NUP93 showing the regions of NUP93 interacting with NUP205 (aa. 96-151, red), and the linker to be modeled, connecting residues Ser 151 and Ser 176 (green). **(C)** Linker modeling procedure. All distances between Cα atoms of the indicated residues are shown in the tables below (in Å). Pairs of residues that are within a maximum length of the S151-S176 linker (87.5 Å) are shown in red and bold. For each pair, the linker is sufficiently long to reach to the respective copies of NUP205 in only one geometrically possible solution.

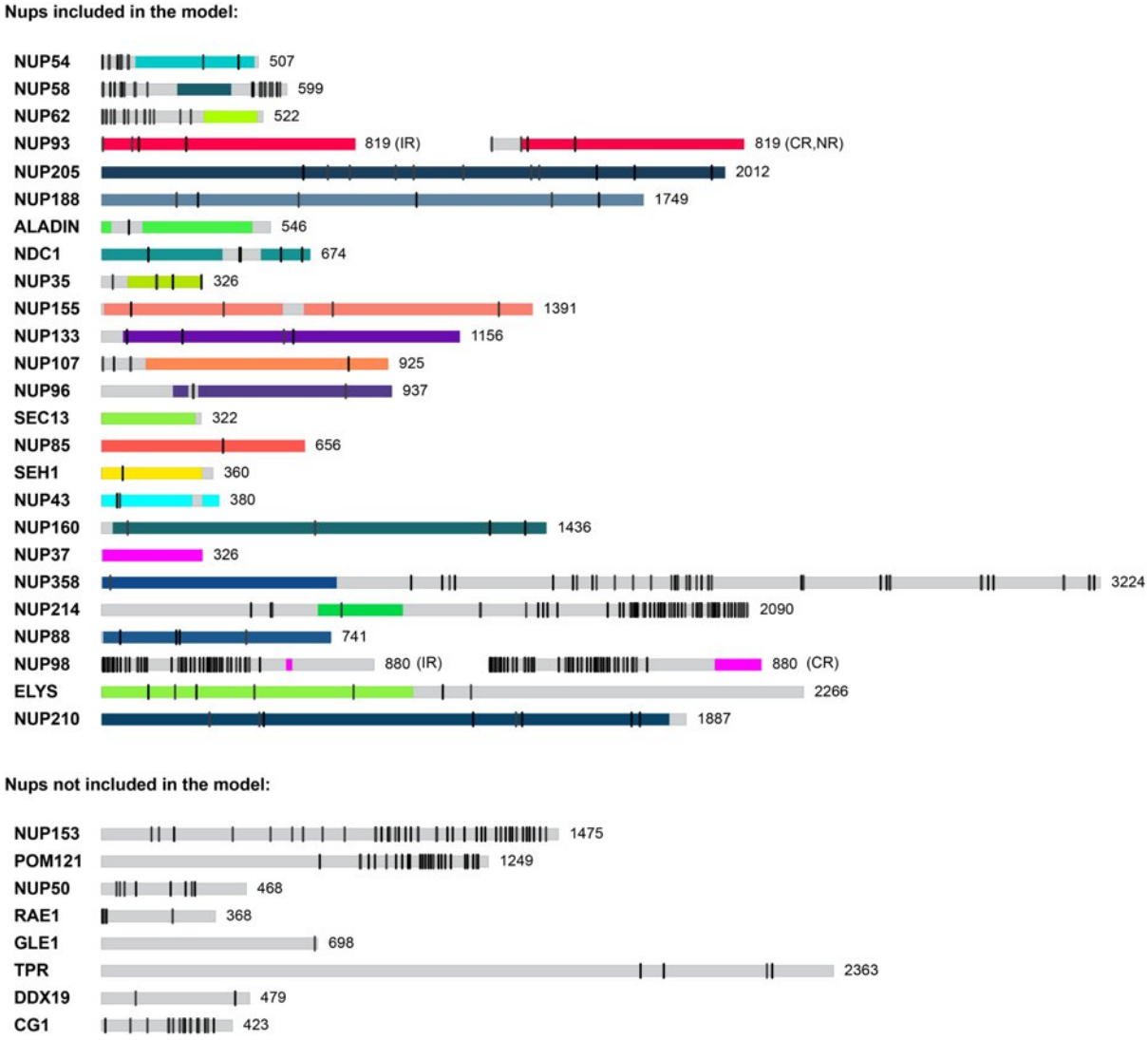


Fig. S19 Coverage of the NUP sequences in the model. Each NUP is shown as a horizontal bar. The colored regions indicate regions included in the model, with the colors corresponding to the scheme in Fig. 1. Regions not included in the model are colored gray. FG and GF motifs are indicated as black vertical lines. The numbers are corresponding to the lengths of nucleoporin sequences. The MD simulations included all indicated regions except the entire NUP210 protein.

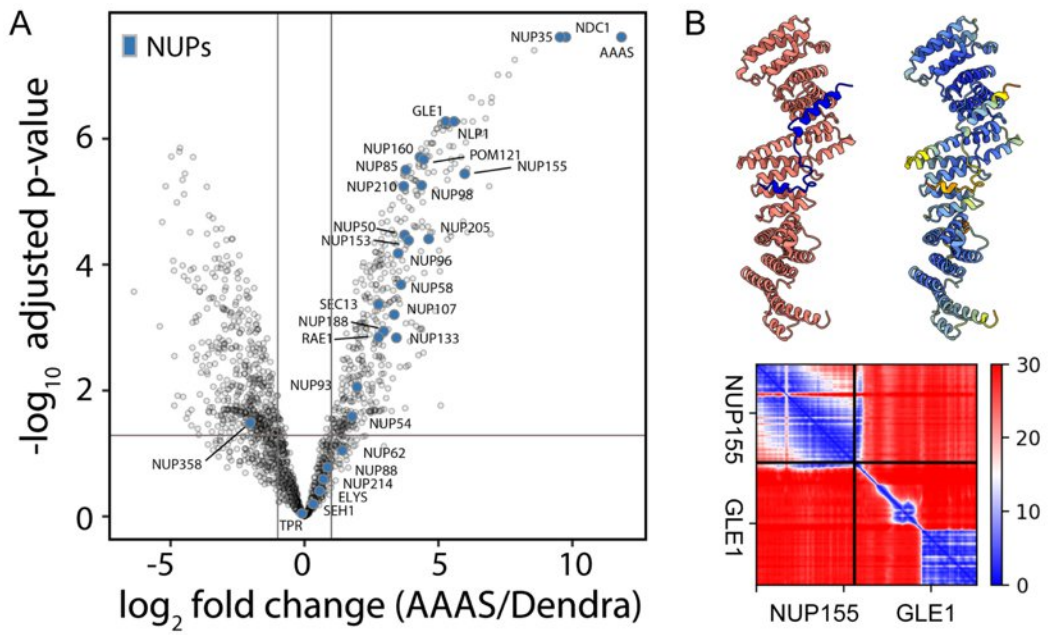


Fig. S20 Proximity labeling experiment with BirA tagged ALADIN. (A) Volcano plot showing the enrichment of individual NUPs in the proximity of ALADIN (AAAS). NDC1 and NUP35 are highly enriched to similar levels as ALADIN itself. NUPs in proximity to the transmembrane interaction hub at the interface of the ORs with the IR are moderately enriched. This includes NUP155, NUP210, NUP155, and NUP160. **(B)** The enriched nucleoporin GLE1 has been shown to interact with NUP155 (*45*), and this interaction is structurally validated by a AlphaFold/ColabFold model. In the AlphaFold model, the N terminus of GLE1 (blue) interacts with the C terminus of NUP155 (salmon, top left). The model has high-local confidence (top right) and confidence of the interaction (bottom). See Fig. S11 for the explanation of the confidence measures.

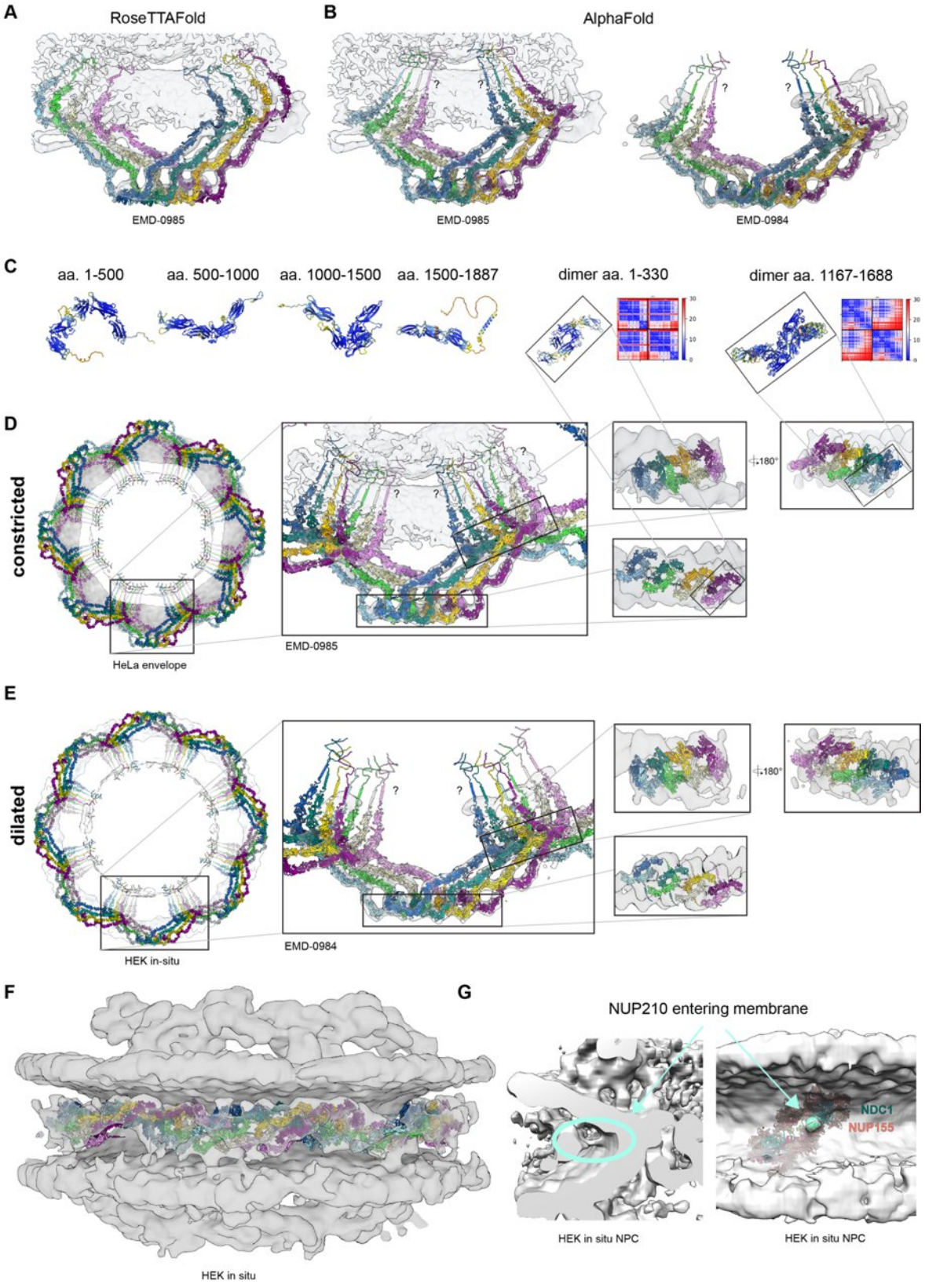


Fig. S21 Modeling NUP210 and overview of the resulting models. (A) A rigid-body fit of NUP210 modelled as full-length with RoseTTAfold. **(B)** AlphaFold did not produce a model fitting as a rigid body but a model could be constructed by superposing smaller AlphaFold fragments onto the RoseTTAfold fit and local refinement of the fits. **(C)** Selected fragments of AlphaFold models for monomers and dimers, which cover dimerization interfaces between

neighboring copies of NUP210. The models are colored by local confidence and the heatmaps represent inter-chain accuracy as described in Fig. S11. **(D)** The full NUP210 model in the constricted NPC. Insets show various zoomed-in views at the structure. The question marks indicate that the positioning of the transmembrane helices is approximate, and they could make helical bundles, which were however not predicted by AlphaFold. **(E)** The full NUP210 model in the dilated NPC. **(F)** The fit of the NUP210 model in the EM density of the dilated NPC as natively in the HEK cells. **(G)** The density connecting NUP210 density with the membrane likely encompasses the transmembrane helices of NUP210 and localizes next to the ALADIN/NDC1/NUP155/NUP35 hub.

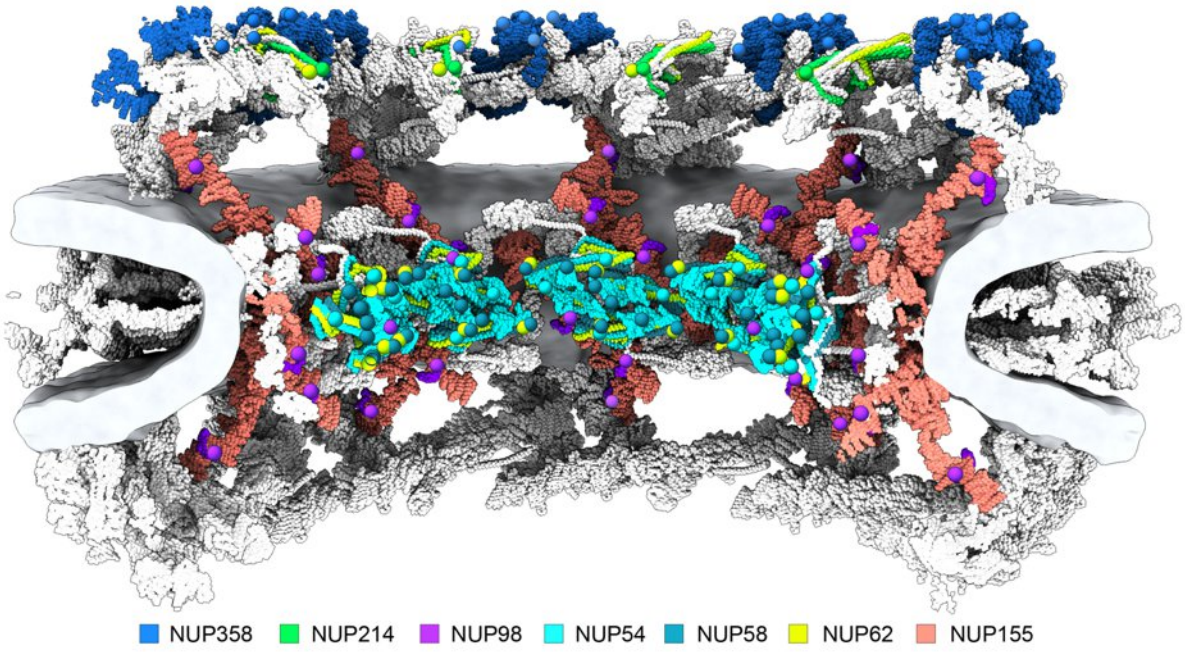


Fig. S22 The map of FG-repeat anchoring points. The model of the constricted NPC is shown in a cut-away view. The termini of NUP structured domains (colored) that anchor flexible FG-repeat tails (not explicitly modeled and thus not shown), are depicted as spheres with the color of the given FG-NUP. The spheres have been placed at the following residues for each NUP: NUP54 - 379, NUP58 - 246, 418, NUP62 – 332, NUP98 – 597, NUP214 – 972, NUP358 – 759.

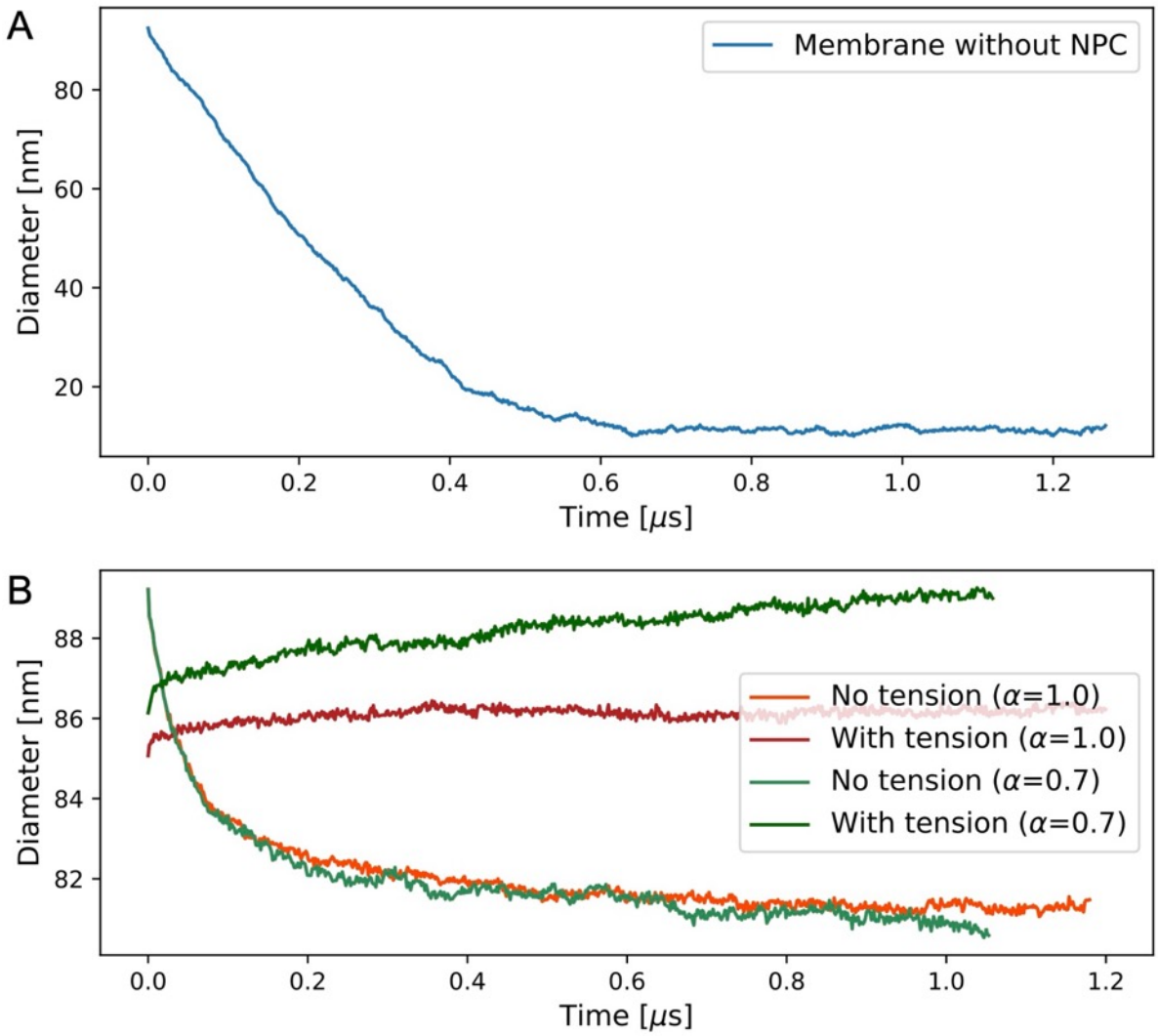


Fig. S23 Timeseries of pore diameters. (A) Diameter of double-membrane pore in MD simulation without NPC as a function of time. See Fig. 4A for snapshots at times t=0 and 1μs. **(B)** Double-membrane pore diameter in MD simulations with hNPC scaffold. See Fig. 4B and Fig. S27 for snapshots of the simulations with $\boldsymbol{\alpha = 0.7}$ and $\boldsymbol{\alpha = 1.0}$, respectively.

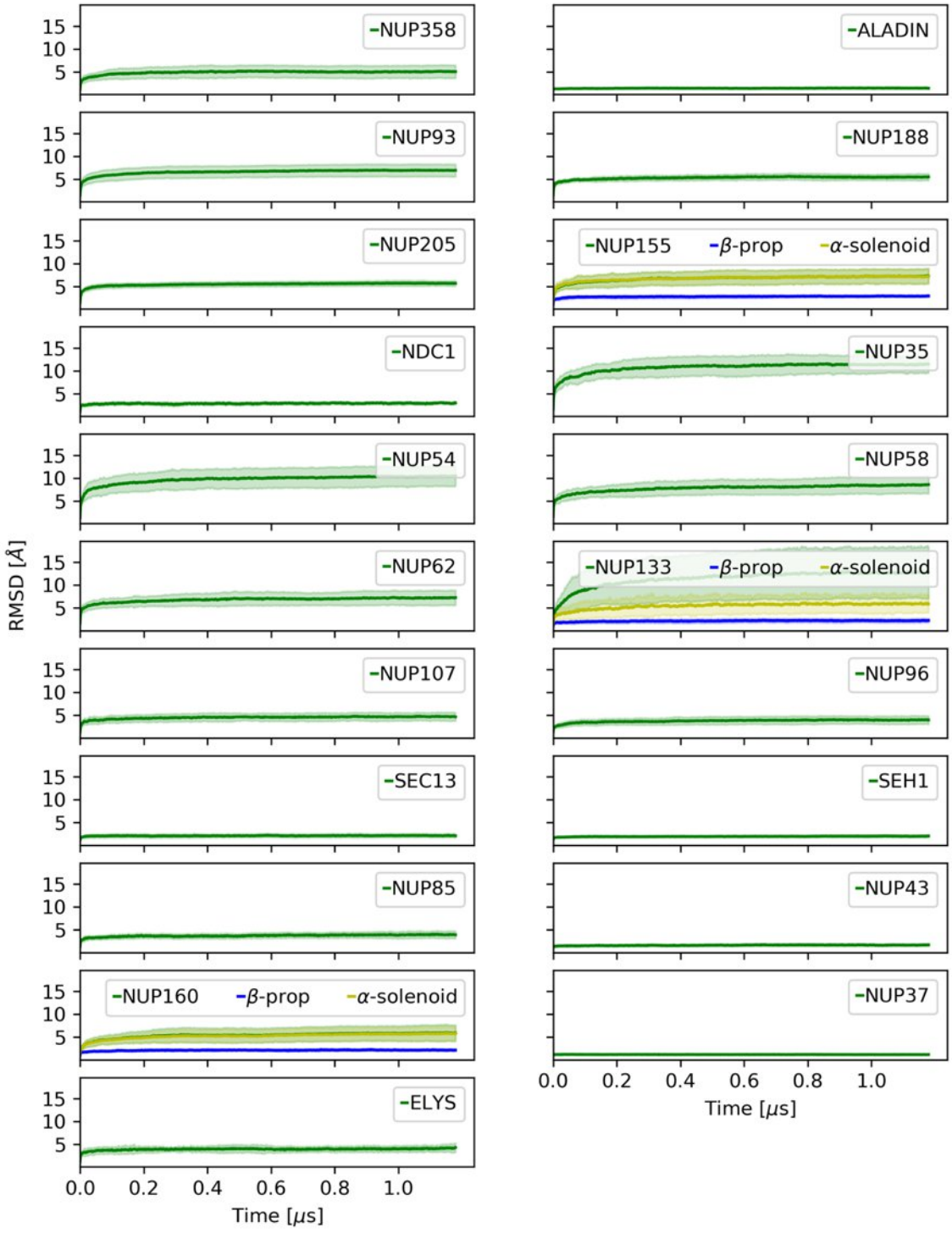


Fig. S24 RMSD of individual proteins from their initial structures in MD simulations of the NPC at a=1.0. The mean value of the RMSD at each given time point across all protein copies is shown as thick green line. Standard deviations are shown in shaded light green. Time points every 1.5 ns were analyzed. Protein names are indicated. The larger apparent flexibility observed in several NUPs such as NUP133, NUP155, NUP35, NUP54, or NUP62 result from more extended regions or flexible linkers. The flexibility between the b-propeller and α-solenoid of NUP155 is consistent with single particle analysis (**Fig. S8**C).

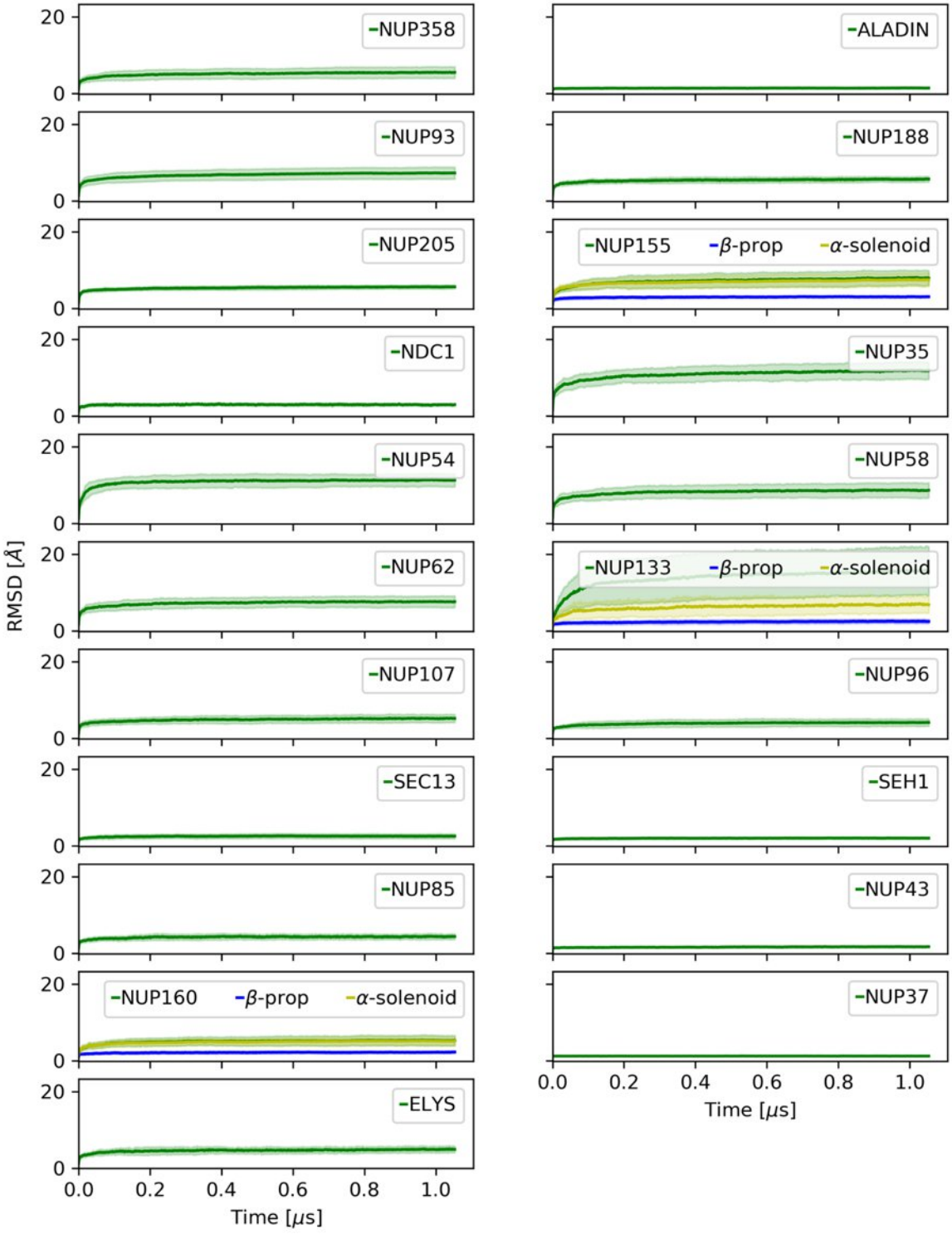


Fig. S25 RMSD of individual proteins from their initial structures in MD simulations of the NPC with a=0.7. The mean value at each given time point across all protein copies is shown as thick green line. Standard deviations are shown in shaded light green. Time points every 1.5 ns were analyzed. Protein names are indicated. The larger apparent flexibility observed in several NUPs such as NUP133, NUP155, NUP35, NUP54, or NUP62 result from more extended regions or flexible linkers.

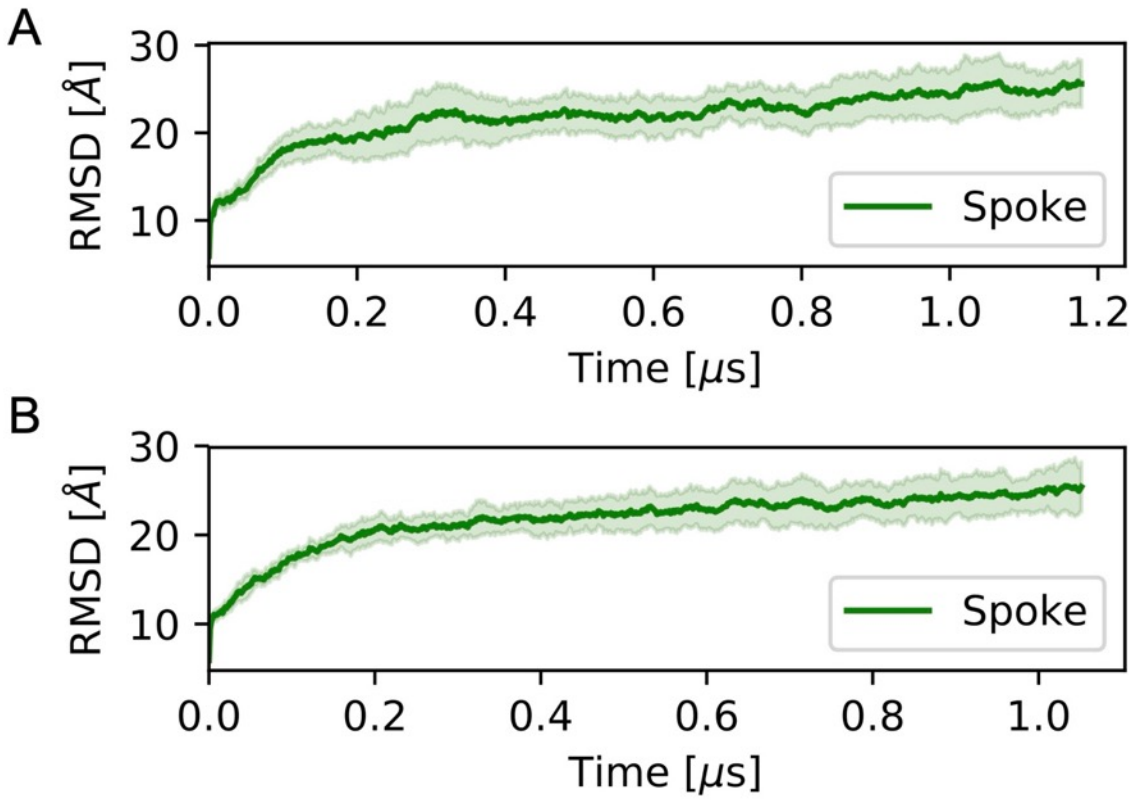


Fig. S26 RMSD of NPC spokes from their initial structures in MD simulations of the NPC with a=1.0 and a=0.7. The mean value of the RMSD at each given time point across all eight copies is shown as thick green line. Standard deviations are shown in shaded light green. Time points every 1.5 ns were analyzed.

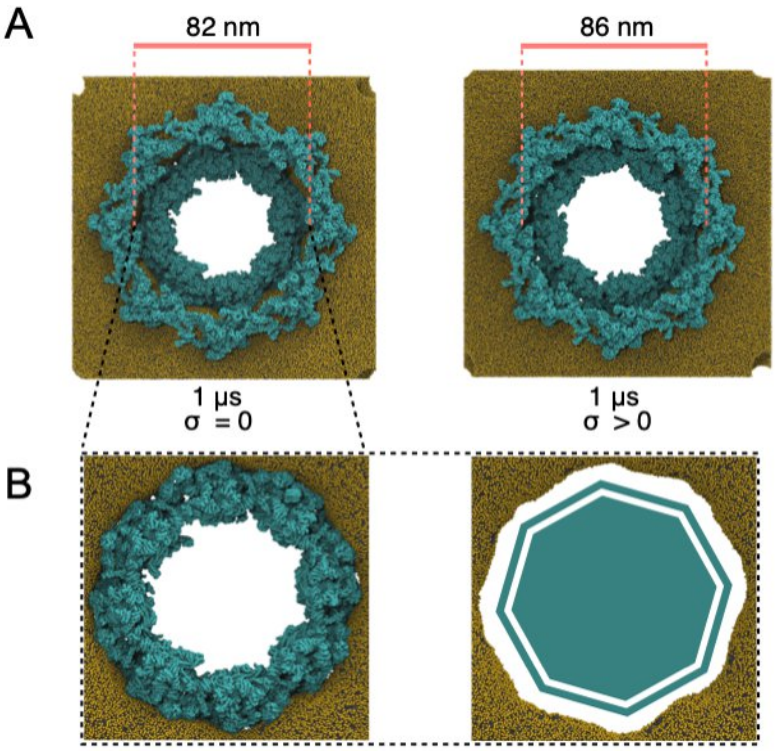


Fig. S27 MD simulations of the complete hNPC with explicit membrane (a=1.0). (A) The NPC (cyan) widens by ~10% in response to lateral membrane tension (right) compared to a zero-tension simulation (left). Shown are snapshots of the relaxed structures after 1 μs of MD with unscaled protein-protein interactions (a=1). **(B)** The membrane fits tightly around the NPC inner ring (cyan, left) and forms an octagonally shaped pore (right, NPC removed).

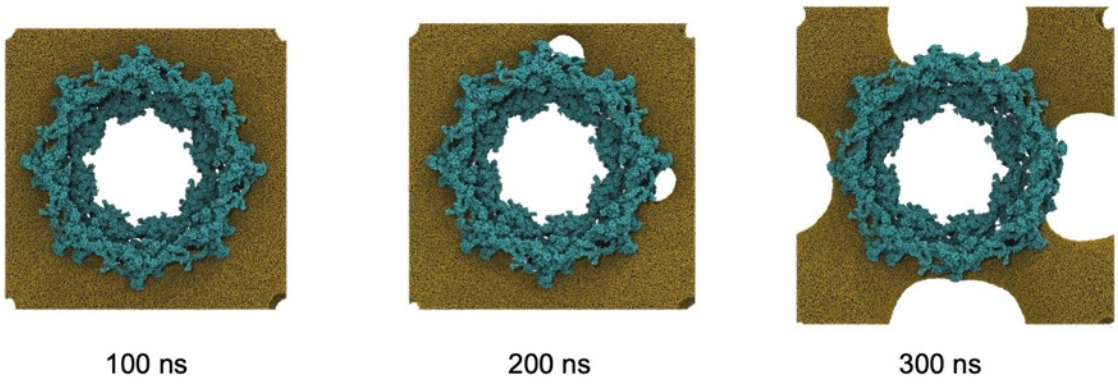


Fig. S28 Snapshots from a representative membrane rupture event under high membrane tension (ΔP = 3 bar). The NPC scaffold backbone beads are shown in cyan. Lipid headgroups are shown as gold spheres. Water and ions are omitted for clarity.

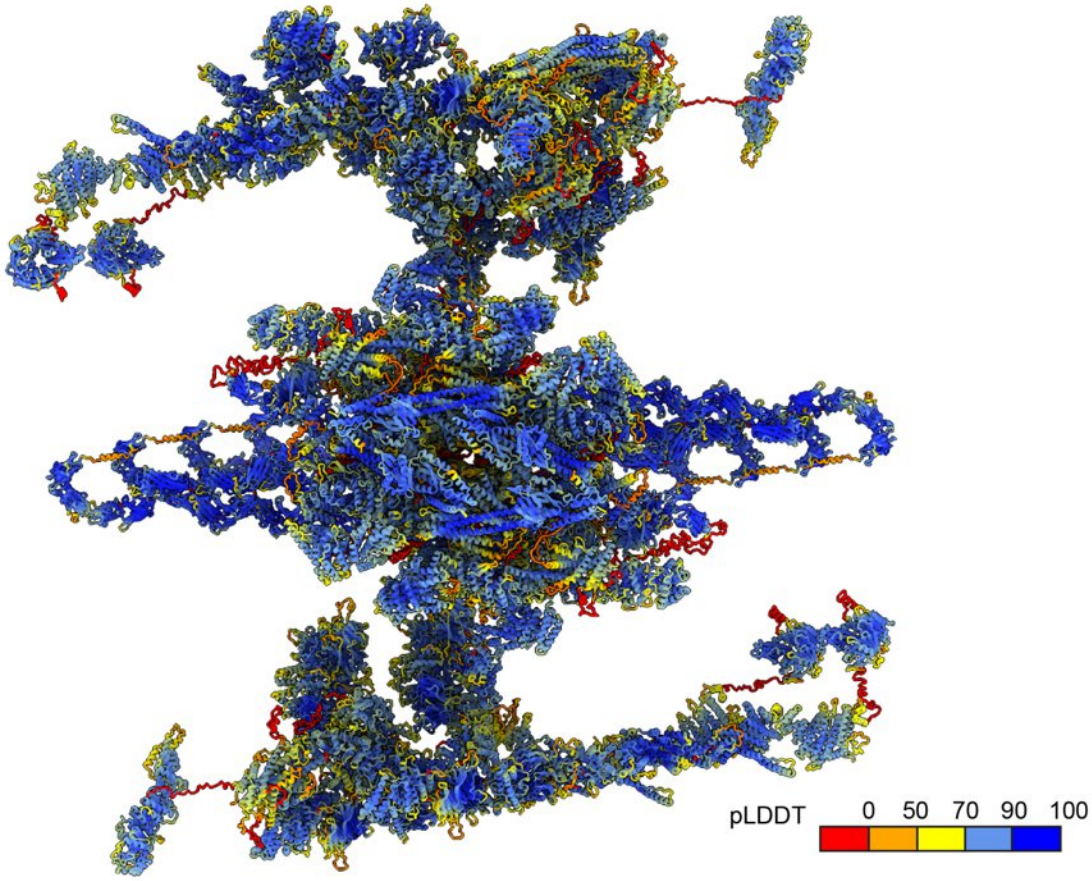


Fig. S29 Projection of the estimated local accuracy onto an asymmetric unit of the NPC. The model is colored by local confidence estimated with predicted local distance difference test (pLDDT), as returned by AlphaFold. The pLDDT > 90 (dark blue) indicates high estimated accuracy of backbone and side chain rotamers whereas pLDDT > 70 (yellow) indicates confident backbone prediction (*61*). Regions that were added subsequently to AI-based structure prediction in Assembline are listed in **Table S4** and their accuracy was set to zero.

Tables S1 to S6

Table S1 Global quality assessment of the AlphaFold models. The columns indicate as follows: "AlphaFold/ColabFold model" – NUPs that were included in the given AlphaFold or ColabFold model and residue ranges included in the model. "pTM-score" and "pLDDT" – TM-scores and pLDDT predicted by AlphaFold. "PDB" – the PDB code of the corresponding structure of the protein/complex from human, if available. "TM-score" – the TM-score calculated by comparing the model to the respective PDB structure calculated for monomers using TM-align (*118*) or for complexes using MM-align (*119*). "LDDT" – the Local Distance Difference Test score calculated using lDDT program (*120*) for the respective model and PDB pairs. For multichain models, LDDT scores are reported for each chain separately as the LDDT program does not support multichain structures. "RMSD" – Root Mean Square Deviation calculated for the pairs, calculated using TM-align and MM-align.

AlphaFold/ColabFold model	pTM-score	pLDDT	PDB	TM-score	LDDT	RMSD
NUP54 (aa. 111-493) - NUP58 (aa. 246-418) - NUP62 (aa. 332-502) - NUP93 (aa. 1-93)	0.72	85.42	-	-	-	-
NUP205 (aa. 1001-2012)- NUP93 (aa. 89-173)	0.82	77.69	-	-	-	-
NUP188 (aa. 1001-1749) - NUP93 (aa. 89-173)	0.82	80.83	-	-	-	-
NUP205	0.73	77.64	-	-	-	-
NUP188	0.68	79.58	-	-	-	-
NUP93 (aa. 171-819) - NUP35 (aa. 81-110)	0.85	89.32	7MW0	0.86	92.81,96.0	2.51
NUP155 (aa. 1-660) - NUP35 (aa. 271-326)	0.79	79.01	-	-	-	-
NUP155 (aa. 1011-1391) - NUP98 (aa. 580-625)	0.83	89.05	-	-	-	-
NUP155 (aa. 10-584,655-1391)	0.71	86.43	7R1Y	0.94	91.1	2.40
NUP35 (aa. 171-255) dimer	0.89	93.70	4LIR	0.96	97.6,93.9	0.88
NDC1 (aa. 1-390,516-674)- ALADIN (aa. 1-31,134-486)	0.84	84.57	-	-	-	-
ALADIN (aa. 1-31,134-486) - NUP35 (aa. 281-326)	0.82	82.85	-	-	-	-
NUP43	0.86	86.24	4I79	0.98	93.04	1.08
NUP85 - SEH1 - NUP43	0.84	82.98	-	-	-	-
NUP160 (aa. 1017-1436) - NUP96 - SEC13	0.74	73.11	-	-	-	-
NUP160 (aa. 838-1069)- NUP37	0.88	90.46	-	-	-	-
NUP160 (aa. 38-1436)	0.72	82.58	-	-	-	-
NUP160 (aa. 951-1436) - NUP85 (aa. 421-656) - NUP96 (aa. 681-937)	0.80	84.34	-	-	-	-
NUP107 (aa. 144-925)	0.83	88.63	3I4R	0.98	96.35	0.82
NUP133 (aa. 71-480)	0.89	86.66	1XKS	0.97	91.0	1.41
NUP133 (aa. 511-1156)	0.68	80.19	3I4R	0.84	63.03	3.89
NUP133 (aa. 511-1156)	0.68	80.19	3CQC	0.97	85.99	0.90
NUP107 (aa. 1-480) - NUP96 (aa. 361-660)	0.79	77.16	-	-	-	-
NUP107 (aa. 421-925) - NUP133 (aa. 501-1156)	0.63	79.52	3I4R	0.81	96.26,62.38	4.53
NUP107 (aa. 421-925) - NUP133 (aa. 501-1156)	0.63	79.52	3CQC	0.97	96.84,83.93	1.41
NUP358 (aa. 1-759)	0.81	87.87	7MNJ	0.85	87.36	4.08
ELYS (aa. 1-1005)	0.75	85.34	-	-		-
NUP88 (aa. 1-579) - NUP98 (731-880)	0.83	80.39	7MNI	0.74	82.65,92.91	2.07
NUP214 (aa. 1-1000) - NUP88 - NUP62 (332-502)	0.56	68.29	-	-	-	-
NUP210 (aa. 1-500) dimer	0.78	85.93	-	-	-	-
NUP210 (aa.978-1287) - NUP210 (aa. 1270-1554)	0.85	92.99	-	-	-	-
NUP210 (aa. 240-500)	0.84	92.77	-	-	-	-
NUP210 (aa. 360-600)	0.78	92.89	-	-	-	-
NUP210 (aa. 500-1000)	0.78	92.21	-	-	-	-
NUP210 (aa. 750-1250)	0.69	90.16	-	-	-	-
NUP210 (aa. 1000-1500)	0.83	92.76	-	-	-	-
NUP210 (aa. 1250-1750)	0.74	89.71	-	-	-	-
NUP210 (aa. 1500-1887)	0.62	79.00	-	-	-	-

Table S2 Comparison of the interfaces modelled by AlphaFold/ColabFold to experimental structures from human and other species. The columns indicate as follows: "NUPs" – NUPs that were included in the given AlphaFold or ColabFold model. "PDB" – the PDB code of the corresponding structure of the complex. "RMSD" – Root Mean Square Deviation calculated by comparing the modelled interface to the respective PDB structure, using the interface residues only. "iTM-score" – the TM-score calculated by comparing the interfaces. The values were calculated using iAlign (*121*). PDBs taken from references (*11*, *13*, *31*, *32*, *122–126*).

NUPs	PDB	iRMSD	iTMscore
NUP107/NUP96	3IKO	1.98	0.73
NUP96/SEC13	3IKO	2.28	0.71
NUP85/SEH1	3F3F	2.73	0.68
NUP160/NUP37	4GQ2	2.88	0.38
NUP62/NUP54	5CWS	5.01	0.44
NUP58/NUP93	5CWS	2.62	0.46
NUP54/NUP93	5CWS	3.23	0.18
NUP62/NUP93	5CWS	2.05	0.55
NUP62/NUP58	5CWS	2.43	0.7
NUP58/NUP54	5CWS	3.19	0.66
NUP88/NUP214	5CWW	3.02	0.44
NUP88/NUP98	5CWW	2.9	0.36
NUP88/NUP214	3PBP	2.74	0.42
NUP88/NUP98	3PBP	2.96	0.46
NUP98/NUP155	5HB0	1.89	0.39
NUP155/NUP35	5HAX	2.71	0.31
NUP93/NUP35	5HB3	3.05	0.16
NUP85/NUP160	4YCZ	1.12	0.55
NUP96/SEC13	4YCZ	2.11	0.69
NUP96/NUP160	4YCZ	2.55	0.2
NUP188/NUP93	7MVX	3.09	0.49
NUP205/NUP93	7MVT	2.54	0.55
NUP88/NUP98	7MNI	1.66	0.67

Table S3 Physical parameters of protein-protein interfaces modeled by AlphaFold, compared to interfaces in crystal structures of NUP complexes and in models of false-negative control complexes (i.e., for proteins not expected to interact). The columns indicate as follows: “Model / PDB” - which NUPs were included in the given AlphaFold model or PDB ID of a crystal structure, “Interface” – the evaluated interface from the model or structure, “Num_intf_residues” – number of interface residues. “Polar”, ”Hydrophobic”, and “Charged” – fractions of the corresponding type of residues at the interface, “Conserved interface” – conservation score according to Rate4Site (*127*) ,“Contact pairs” – number of contact pairs, “hb” – number of potential hydrogen bonds, “sb” – number of potential salt bridges, “int_solv_en” - solvation free energy gain upon formation of the interface, “int_area” – interface area, “pvalue” - the P-value of the observed solvation free energy gain (P<0.5 implies higher likelihood that the interaction is specific (*128*). Units are specified where applicable, “pi_score” – PI score of the interface, with values > 0 indicating that the interface resembles real interfaces. All values were calculated with PI_score pipeline (*129*) using CCP4 (*130*), Rate4Site (*127*), and PISA (*128*) programs included in the pipeline. The models and crystal structures were minimized using GROMACS prior to analysis. “-“ values for some of the false-negative control complexes indicate that the pipeline detected no interface.

Model / PDB	Interface	Num_intf_residues	Polar	Hydrophobic	Charged	Conserved_interface	Contact_pairs	sc	hb	sb	int_solv_en [kcal/M]	int_area [Å^2]	pvalue	pi_score
ALADIN_53	NUP35-ALADIN	21	0.19	0.29	0.33	0.52	26	0.64	6	7	-6.63	912.09	0.58	0.81
NUP210_NTD_dimer	NUP210-NUP210	51	0.35	0.39	0.18	0.49	54	0.72	29	8	-9.14	1612.49	0.53	2.04
NUP210_middle_dimer	NUP210-NUP210	49	0.29	0.43	0.18	0.41	54	0.67	24	2	-24.73	2175.63	0.04	1.82
NDC1_ALADIN	ALADIN-NDC1	24	0.46	0.38	0.12	0.5	23	0.6	10	2	-14.3	1425.28	0.61	1.28
NUP107_NUP133	NUP133-NUP107	29	0.1	0.38	0.34	0.52	25	0.65	10	6	-17.32	1496.07	0.07	2.05
NUP107_NUP96	NUP96-NUP107	54	0.24	0.43	0.17	0.39	55	0.66	21	4	-25.42	2100.78	0.37	1.79
NUP155_NUP35	NUP35-NUP155	23	0.22	0.61	0.04	0.39	32	0.62	10	4	-11.77	1037.9	0.71	1.63
NUP160_NUP37	NUP37-NUP160	12	0.17	0.33	0.17	0.17	9	0.6	8	9	-4.47	1063.94	0.54	0.35
NUP160_NUP96_NUP85	NUP85-NUP160	23	0.35	0.3	0.26	0.39	17	0.58	5	6	-13.11	932.22	0.09	1.86
NUP160_NUP96_SEC13	SEC13-NUP96	89	0.33	0.39	0.16	0.34	109	0.61	34	5	-46.84	4091.32	0.03	1.21
NUP160_NUP96_SEC13	NUP96-NUP160	50	0.3	0.38	0.24	0.4	50	0.64	16	11	-25.9	2216.86	0.1	2.03
NUP188_NUP93	NUP93-NUP188	24	0.38	0.38	0.25	0.54	18	0.57	14	13	-32.53	2745.79	0.23	1.3
NUP205_NUP93	NUP93-NUP205	32	0.34	0.47	0.16	0.5	25	0.68	17	9	-33.71	2623	0.23	1.48
NUP214_NUP88_NUP62	NUP62-NUP88	60	0.22	0.5	0.22	0.47	36	0.61	19	20	-62.22	3944.81	0.06	1.26
NUP214_NUP88_NUP62	NUP62-NUP214	27	0.33	0.52	0.15	0.41	14	0.64	7	6	-53.15	2943.97	0.06	1.42
NUP214_NUP88_NUP62	NUP88-NUP214	77	0.23	0.49	0.2	0.4	64	0.61	23	36	-90.53	5646.72	0.04	0.89
NUP62_NUP58_NUP54_NUP93	NUP62-NUP54	52	0.21	0.54	0.21	0.35	32	0.67	22	23	-66.96	4223.88	0.03	1.16
NUP62_NUP58_NUP54_NUP93	NUP93-NUP54	21	0.24	0.38	0.29	0.43	20	0.56	6	1	-17.32	1152.31	0.16	1.68
NUP62_NUP58_NUP54_NUP93	NUP93-NUP58	40	0.33	0.45	0.17	0.33	35	0.68	12	3	-21.61	1947.84	0.53	2.03
NUP62_NUP58_NUP54_NUP93	NUP93-NUP62	31	0.19	0.48	0.19	0.52	27	0.7	9	5	-19.15	1145.38	0.23	2.62
NUP62_NUP58_NUP54_NUP93	NUP62-NUP58	33	0.27	0.64	0.09	0.61	17	0.68	6	4	-57.33	2676.17	0.02	1.27
NUP62_NUP58_NUP54_NUP93	NUP58-NUP54	37	0.32	0.51	0.11	0.35	22	0.66	11	10	-63.54	3456.3	0.03	1.3
NUP85_SEH1_NUP43	NUP43-NUP85	30	0.5	0.1	0.2	0.4	26	0.57	19	10	1.25	1778.68	0.8	1.66
NUP85_SEH1_NUP43	NUP43-SEH1	21	0.43	0.05	0.19	0.29	23	0.66	5	0	-1.84	403.56	0.51	1.21
NUP85_SEH1_NUP43	SEH1-NUP85	122	0.29	0.38	0.18	0.25	149	0.62	31	10	-58	4320.18	0	1.04
NUP93_NUP35	NUP35-NUP93	16	0.31	0.19	0.31	0.5	18	0.69	11	2	-10.65	969.41	0.49	2.34
NUP98_NUP155	NUP98-NUP155	13	0.08	0.54	0.39	0.61	17	0.84	4	2	-11.02	726.11	0.14	1.95
NUP98_NUP88	NUP88-NUP98	28	0.21	0.32	0.32	0.5	25	0.59	15	2	-7.77	1577.56	0.74	0.71
		37	0.28	0.39	0.21	0.43	35	0.64	14	7	-27.87	2144.17	0.29	1.44
3cqc	NUP133-NUP107	20	0.25	0.4	0.3	0.55	19	0.69	11	6	-16.83	1267.94	0.12	2.53
3fmp	DDX19-NUP214	23	0.3	0.17	0.3	0.43	17	0.63	19	16	4.11	1080.68	0.84	1.58
3iko	Sec13-Nup145C	52	0.39	0.29	0.14	0.19	44	0.67	23	5	-24.5	2601.48	0.09	1.75
3iko	Nup145C-Nup84	49	0.27	0.45	0.22	0.53	45	0.69	28	5	-27.75	2056.62	0.02	1.85
4gq2	Nup37-Nup120	46	0.41	0.3	0.13	0.41	44	0.63	14	9	-19.38	2153.66	0.2	1.68
4ycz	Nup160-Nup85	17	0.29	0.35	0.35	0.47	13	0.6	6	0	-7.57	907.74	0.45	1.07
4ycz	Nup96-Sec13	58	0.33	0.34	0.15	0.34	55	0.46	18	10	-15.28	2477.64	0.5	0.25
4ycz	Nup160-Nup96	25	0.24	0.36	0.32	0.2	16	0.36	8	3	-24.94	1820.85	0.09	0.57
5cws	Nup57-Nsp1	38	0.16	0.55	0.24	0.32	22	0.64	18	19	-66.53	4132.77	0.13	1.17
5cws	Nup49-Nsp1	26	0.31	0.54	0.08	0.58	15	0.59	7	4	-59.45	2735.02	0.04	1.26
5cws	Nup57-Nup49	45	0.2	0.62	0.16	0.51	26	0.64	15	13	-63.4	3574.63	0.13	1.33
5cww	Nup82-Nup145N	31	0.32	0.26	0.13	0.45	33	0.74	15	19	-3.76	1150.59	0.52	2.3
5cww	Nup159-Nup82	33	0.18	0.48	0.24	0.39	30	0.59	11	14	-10.67	1371.14	0.29	1.72
6b4j	DDX19B-GLE1	33	0.18	0.48	0.18	0.46	28	0.71	9	1	-14	996.2	0.25	2.58
		35	0.27	0.4	0.21	0.42	29	0.62	14	8	-25	2023.35	0.26	1.55
NDC1_SEH1	NDC1-SEH1	6	0.17	0.33	0.33	0.5	6	0.66	2	0	-0.78	343.29	0.77	-0.14
NUP160_NUP43	NUP160-NUP43	10	0.3	0.2	0	0.4	10	0.44	1	0	-4.73	564.04	0.35	-1.07
NUP85_NUP37	NUP85-NUP37	5	0.2	0.2	0.2	0	3	0.25	2	0	1.14	385.37	0.67	-1.23
NUP96_SEH1	NUP96-SEH1	-	-	-	-	-	-	-	-	-	-	-	-	-
NUP107_NUP54	NUP107-NUP54	-	-	-	-	-	-	-	-	-	-	-	-	-
		7	0.22	0.24	0.18	0.3	6	0.45	1	0	-1.46	430.9	0.6	-0.81

Table S4 Coverage of the NUPs in the model. All residues were modeled with AlphaFold except residues indicated in the rightmost column, which are either added or re-modeled using Assembline. The MD simulations included all indicated regions except the entire NUP210 protein.

Nucleoporin	Regions covered by the model	Regions added denovo/remodeled using Assembline
NUP54	111-493	
NUP58	246-418	
NUP62	332-502	
NUP93	2-819 (IR) 94-819 (CR,NR)	86-96,151-178 (IR) 151-178 (CR and NR)
NUP205	2-2012	2-9,465-488,1146-1179,1360-1385,1791-1805,1926-1947
NUP188	2-1749	1707-1729
ALADIN	2-31,134-486	
NDC1	2-390,516-674	
NUP35	86-326	101-172,252-286,301-308,310-314
NUP155	10-584,655-1391	863-868,988-1012,1104-1110
NUP133	71-1156	251-270,479-514
NUP107	144-925	583-585,730-736
NUP96	232-280,314-937	281-313,402-403,585-586,761-762
SEC13	2-302	
NUP85	2-656	60-61,388-389,550-552
SEH1	2-324	
NUP43	2-292,327-380	
NUP160	38-1436	991-993,1053-1055,1274-1283
NUP37	5-326	
NUP358	4-759	161-164
NUP214	700-972	
NUP88	7-741	
NUP98	597-615 (IR) 731-880 (CR)	
ELYS	2-1005	555-557
NUP210	2-1832	343,452,641,747,893,977,1166,1553-1554,1688-1689,1787-1789

Table S5 Overview of MD simulations.

Protein	**System size [nm^3]**	**No. of particles**	**Alpha scaling** (protein-protein)	**Pxx=Pyy, Pzz [bar]** (after equil)	**Simulation time**
Without hNPC	140×140×80	13813071	1.0	1, 1	1 µs
With hNPC	140×140×80	13395936	1.0	1, 1	1 µs
With hNPC	140×140×80	13395936	0.7	1, 1	1 µs
With hNPC	140×140×80	13395936	1.0	0.33,2.33	1 µs
With hNPC	140×140×80	13395936	0.7	0.33,2.33	1 µs
With hNPC	140×140×80	13395936	1.0	0, 3	1 µs

Table S6 Cryo-EM data collection, refinement and validation statistics

	NPC – Hela Envelopes EMD-14322	NPC – HeLa *in cellulo* EMD-14327	NPC – HEK *in cellulo* EMD-14321	NPC – HEK *in cellulo (NUP210D)* EMD-14326	NUP155 EMD-14243 PDB ID 7R1Y
Data collection and processing					
Microscope	TFS Krios G2	TFS Krios G2	TFS Krios G4	TFS Krios G4	TFS Krios G4
Voltage (keV)	300	300	300	300	300
Camera	Gatan K2	Gatan K3	Gatan K3	Gatan K3	TFS Falcon4
Magnification	42000	42000	42000	42000	165,000
Pixel size at detector (Å/pixel)	3.37	3.37	3.45	3.45	0.730
Total electron exposure (e^-/ Å^2)	120	120	120-150	120-150	50
Defocus range (mm)	-2 to -4	-2 to -4	-2 to -3	-2 to -3	-1.0 to -2.0
Automation software	SerialEM	SerialEM	SerialEM	SerialEM	EPU
Energy filter slit width (eV)	20	20	20	20	10
Data collected	617 (tilt-series)	13 (tilt-series)	8 (tilt-series)	13 (tilt-series)	6430 (micrographs)
Data used	554 (tilt-series)	13 (tilt-series)	8 (tilt-series)	13 (tilt-series)	5901 (micrographs)
Total particles extracted (no.)	7711 (sub-tomograms)	53 (sub-tomograms)	30 (sub-tomograms)	43 (sub-tomograms)	1,421,476
Symmetry imposed	C8	C8	C8	C8	C1
Final particles (no.)	21604 (CR) 30000 (SR) 11112 (NR)	280 (CR, NR, SR)	150 (CR, NR, SR)	222 (CR, NR, SR)	98,742
Symmetry imposed to final particles	C1	C1	C1	C1	C1
Map Resolution (Å)	-	-	-	-	3.0
- FSC threshold	0.143	0.143	0.5	0.5	0.143
Resolution range (Å)	12-24	44-53	47-54	44-48	2.94–6.13
Map sharpening *B* factor (Å^2)					-33.6
Model refinement					
Initial model used (PDB ID)					5HAX
Refinement package					Phenix
- real or reciprocal					Real space
Model resolution (Å)					3.3
- FSC threshold					0.5
Model composition					
- non-hydrogen atoms					5,704
- Protein residues					723
- Ligands (no.)					0
- water (no.)					0
B factors (Å^2)					
- Protein					52.1
R.m.s deviations					
- Bond lengths (Å)					0.002
- Bond angles (°)					0.528
Validation					
Molprobity score					1.92
Clash score					6.47
Rotamer outliers (%)					3.44
Ramachandran plot					
- Favored (%)					97.14
- Allowed (%)					2.86
- Outliers (%)					0.00