Title: Artificial intelligence reveals nuclear pore complexity

Authors:

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

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: jan.kosinski@embl.de, gerhard.hummer@biophys.mpg.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 near-complete structural 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. **(117 words)**

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.

Introduction

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 sub-complexes, most prominently the so-called Y-complex (*7*) arranged in head-to-tail within the outer rings (*8*). The assembly of individual sub-complexes 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 (*12*, *19–24*). Moreover, the NPCs from many other species have a vastly reduced architectural complexity, which limits their usefulness for studying the human biology (*12*, *22*, *23*, *25*, *26*). 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 sub-complexes built using AlphaFold (*27*) and RoseTTAfold (*28*) are consistent with unreleased X-ray crystallography structures, cryo-EM maps, and complementary data. We elucidate the exact 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. Our near-complete structural model enables novel types of analyses. It reveals that linker Nups perform dedicated spatial organization functions within the higher order structure. It enables molecular dynamics simulations of the scaffold embedded into explicit membrane that illustrate how the human NPC scaffold counteracts mechanical forces imposed by the nuclear membranes.

Results

A near-complete 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. S1Fig**. S3**). Recent single-particle analysis of NEs isolated from *Xenopus laevis* oocytes yielded nominally higher resolved maps (*29*), but relied on spread NEs containing NPCs in only top views. 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 (*30*), HeLa (*31*), SupT1 (*32*), and most recently, DLD-1 cells (*20*). 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*, *26*), 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 (*27*) and RoseTTAfold (*28*). We found that most of the Nups can be modeled with high confidence scores (Fig. S5). In addition,

we validated the accuracy of the models by comparison to yet unreleased structures solved by X-ray crystallography for human Nup358, Nup93, Nup88, Nup98, and of Nup205 and Nup188 from *Chaetomium thermophilum* (Fig. S6). The X-ray structures were provided to us by the Hoelz laboratory (*33*, *34*) and were not used as input for the modeling procedure. The AI-based models also excellently fit our EM densities (Fig. S7A-E). 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 (*29*) (Fig. S8). The full-length model of ELYS, for which thus far only the N-terminal β-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. S8).

The capacity of AI-based structure prediction tools to identify and model protein interfaces with high accuracy has recently been demonstrated (*35*, *36*). We therefore attempted to model Nup interfaces using the ColabFold software, a version of AlphaFold adopted for modeling protein complexes (*36*). We found that ColabFold predicts several Nup sub-complexes with inter-domain confidence scores that correlate with the accuracy of the models (Fig. S9-Fig**. S11**). The models of these sub-complexes did not only reproduce the respective, already available X-ray structures X-ray structures, but also agreed with the yet unreleased X-ray structures (*33*, *34*). 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). These structures represent proteins in complex with the respective SLiMs and form relatively small interfaces. However, also for larger sub-complexes we obtained structural models that convincingly fit our cryo-ET maps (Fig. S11). 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. S9). ColabFold predicted the structure of the Nup62 complex with excellent accuracy (Fig. S11) 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 very well, confirming the known structure of Nup85-Seh1 interaction, and revealing how Nup43 interacts with Nup85 (Fig. S11). 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. S11). A previously unknown interface between Nup214 and Nup88 that is biochemically validated in (*33*) has also been predicted correctly (Fig. S6).

With the cryo-EM maps and the repertoire of structural models of individual Nups and their sub-complexes, we built a nearly complete model of the human NPC scaffold (Fig. 1A). 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 the Assembline program (*37*). In addition to fitting the models, we have added several disordered linkers that connect spatially separated domains and SLiMs within the NPC. We then built the model of the dilated state by fitting the constricted NPC map into the dilated NPC map and refining the fits using Assembline. The resulting models (Fig. 1) include 25 out of ~30 Nups, and 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.

The near-complete scaffold model yields novel insights into the organization of the human NPC (Fig. 1). It resolves previous ambiguities (*11*, *19*, *21*, *29*) in localization of similarly shaped Nup205 and Nup188. Within the IR, Nup188 localizes to the outer, while Nup205 localizes to the inner sub-complexes, 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 (*29*), 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 (this study and in (*33*, *34*)). The AI-based model of the Nup214 sub-complex interacts with Nup85 pointing towards the central channel, likely to optimally position the associated helicase crucial for mRNA export.

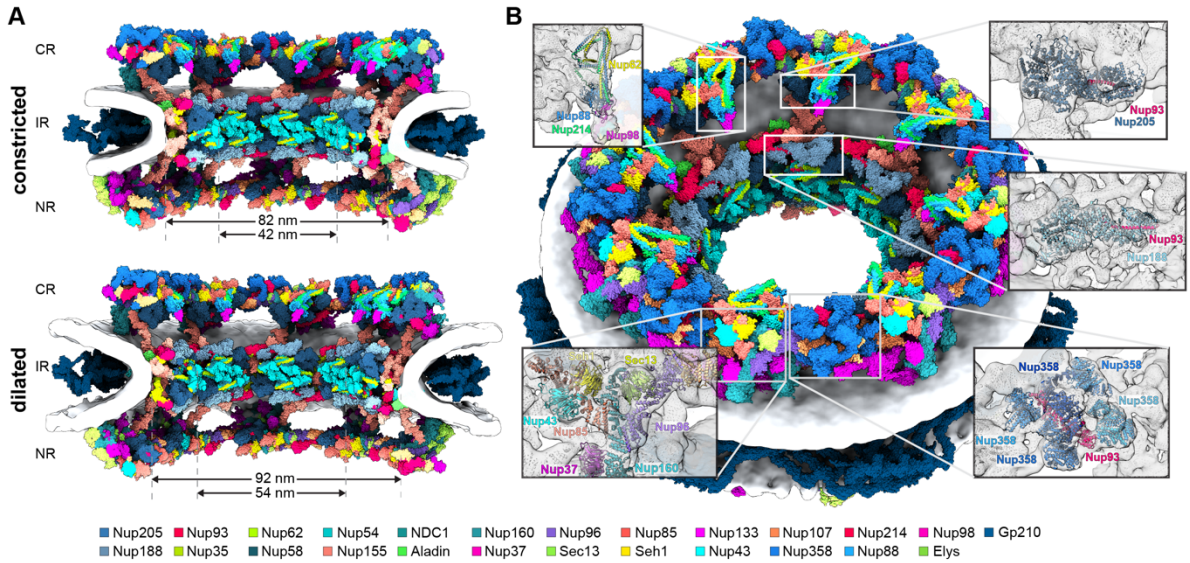

Fig. 1 Scaffold architecture of the human NPC. (A) The near-complete model of the human NPC scaffold is shown for the constricted and dilated states as cut-away views. High-resolution models are color-coded as indicated in the color bar. The nuclear envelope is shown as a gray isosurface. **(B)** Same 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).

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 additional 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 connecting between the anchor points and restricting their possible conformational freedom within our model. We therefore traced the linkers as ensembles of their approximate paths in explicit atomic representation using coarse-grained modeling followed by atomic refinement (Supplementary Materials and Methods).

The resulting connectivity map (Fig. 2) 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, and each of the two copies reaches out with its SLiMs to Nup155 and Nup93 of the adjacent spokes (Fig. 2A). The amphipathic helices of

Nup35 and a SLiM binding to Aladin (see below) limit the distance to the membrane. A Nup35 bridge between adjacent spokes is the only arrangement geometrically possible for both the constricted and the dilated NPC state. These bridges are consistent with the previous interaction and modeling studies in fungi (*10*, *12*, *24*) and the biochemical analysis of the human proteins (*34*). It strongly suggests that the Nup35 dimer, which is critical during early NPC biogenesis (*39*), 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 in either the constricted or the dilated state, and thus connect subunits within a single IR sub-complex inside the same spoke. This is the only geometrically possible solution we found for all four instances of Nup93 in the IR, with the two outer copies binding to Nup188 and the two inner copies binding 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, emphasizing the similar architectural design of the respective complexes. Thereby, the Nup93 SLiM that binds the homologs Nup62 complex could also facilitate linkage between the Nup93 and the 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 EM map of the constricted state, and thus remains to be further investigated.

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.

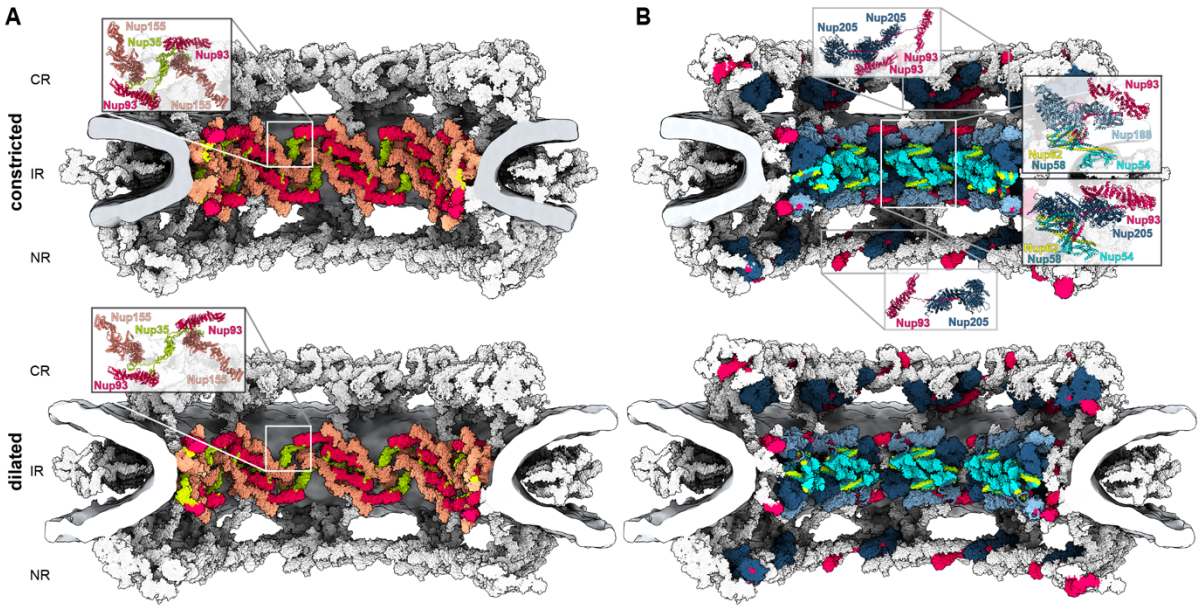

Fig. 2 The connectivity of protein linkers within the human NPC. (A) The Nup35 dimer interconnects adjacent spokes across different sub-complex species, thus facilitating cylindrical assembly of the IR in both constricted (top) and dilated (bottom) states. The Nup205, Nup188, Nup62 complex, and the N-terminus (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 sub-complex 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.

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 remained 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. S12). NDC1 is predicted to comprise six transmembrane helices followed by a cytosolic domain containing mainly α-helices, while Aladin is predicted to have a β-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. S9). Systematic fitting of the hetero-dimeric models to the EM map unambiguously identified two locations within the IR (Fig. S11). The EM density was not used as a restraint for modeling but matches the structure of the model very well and is consistent with the only patches of density spanning the bilayer (Fig. 3, Fig. S11), 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 also 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. S12B). We also predicted an Aladin-binding SLiM in Nup35, located between the amphipathic helix and Nup155-binding SLiM (Fig. 3, Fig. S9). 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 the *Aladin*/*AAAS* that are implicated in the AAA syndrome (*46*–*48*) and is consistent with the absence of Aladin in fungi, which lack the Nup155 connectors (*25*).

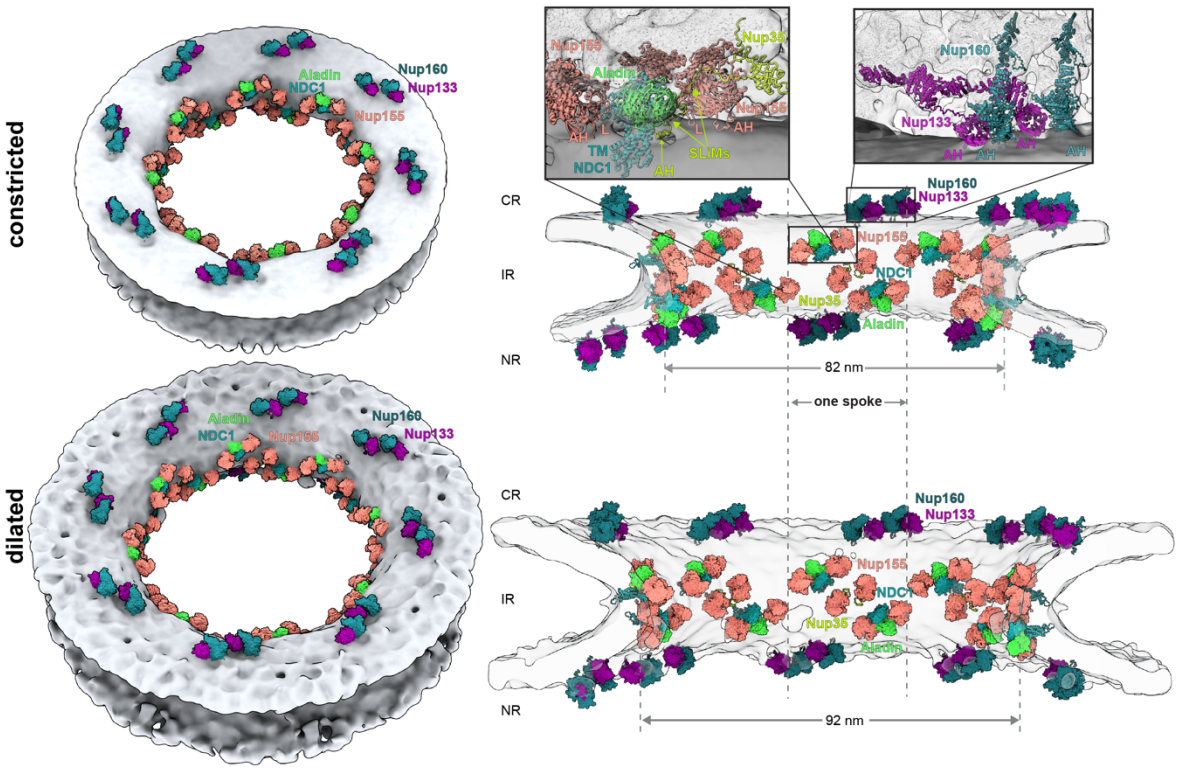

Fig. 3 The membrane-anchoring motifs of the human NPC are distributed over the entire scaffold. The membrane-binding b-propellers of the Y-complex and IR complex are shown color-coded and arranged as pairs of the respective inner and outer copies. Aladin and NDC1 form a transmembrane interaction hub 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.

We next examined the structure of Gp210, 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 Gp210 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 Gp210 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 Gp210 per spoke (Fig. S13). Modeling of individual Gp210 fragments and inter-Gp210 interactions with AlphaFold/ColabFold (Fig. S13) led to an overall model that explained the entire density of the luminal ring of the Xenopus cryo-EM map, including the C-terminal transmembrane helix. This 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 is in good agreement with known interactions of Gp210 homologs (*51*) and our proximity labeling data (Fig. S12). The model of the Gp210 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. S13). To further confirm the Gp210 assignment in human cells, we deleted *gp210* 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 Gp210 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 β-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. S15, Video S1). 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 Notes). Interestingly, the opening of the relaxed double-membrane pore is considerably smaller than even the most constricted NPC conformations. 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. S15-Fig**. S19**, Videos S2-3). 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. S19). Second, by applying lateral tension on the double membrane, we could maintain the pore width or widen it (Fig. 4B, Fig. S15B). At even higher tension, the membrane spontaneously detached from the NPC scaffold (Fig. S20). 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.

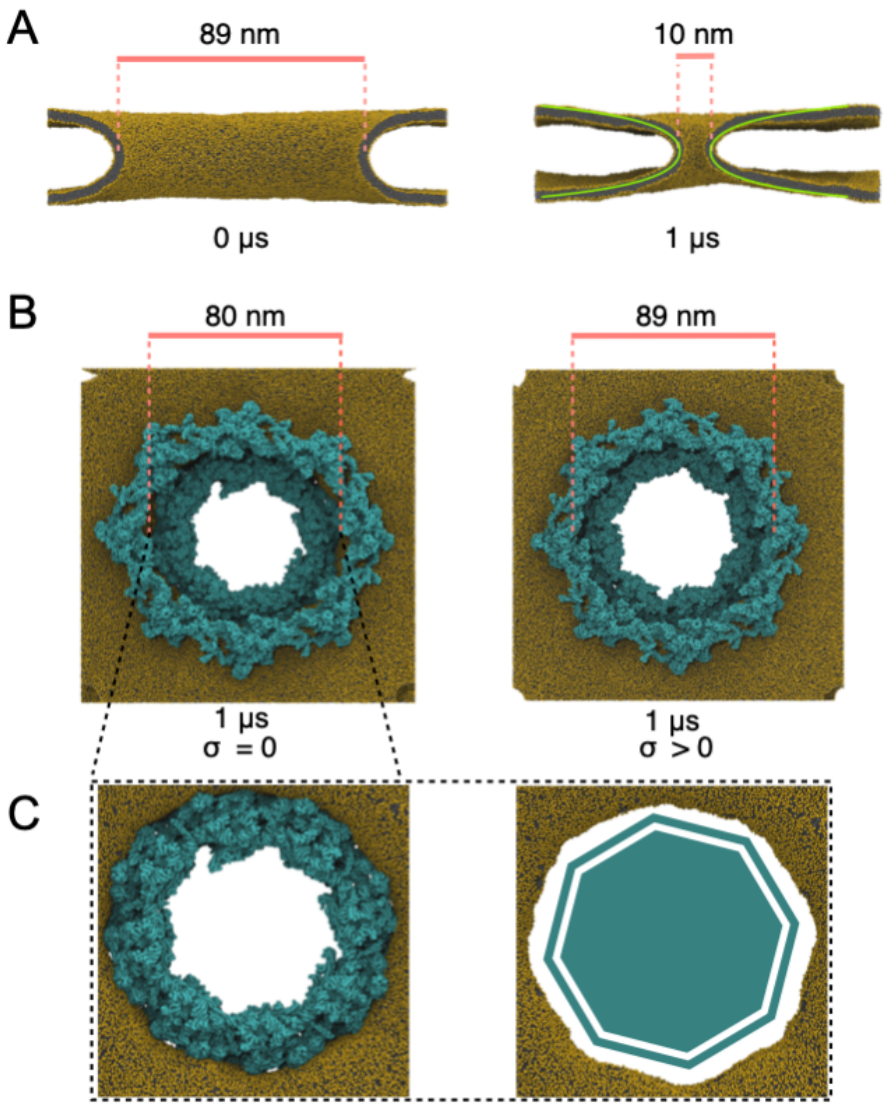

Fig. 4 Dynamics of the NPC from molecular simulations. (A) An isolated half-toroidal double-membrane pore shaped initially as in the tomographic structure of the constricted NPC (left) tightens over the course of 1 µs of MD (right) toward the catenoid-like shape (green) predicted by membrane elastic 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).

Discussion and conclusions

We have built a near-complete 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. S14).

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 Gp210, 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*.

A novel aspect of our work is that we use AI-based structure prediction programs AlphaFold and RoseTTAfold to model all atomic structures used for fitting to the EM maps and modeled the entire NPC scaffold without directly using any X-ray structures or homology models. 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 (*27*, *28*, *61–66*) and complexes (*35*, *62*, *67*, *68*). The programs were also shown to accurately assess their confidence at the level of individual residues and inter-domain contacts (*27*, *28*, *62*). Indeed, we could successfully validate our models by comparing them to unpublished crystal structures, our 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. 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 (*27*, *28*), 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.

References

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**, 490–501 (2010).

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 (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. B. M, H. E, 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. 2015 2210*. **22**, 774–781 (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 nuclear pore complex symmetric core. *Science*. **352**, aaf1015 (2016).

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 (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. *Nat. 2021*, 1–5 (2021).

21. A. von Appen, J. Kosinski, L. Sparks, A. Ori, A. L. DiGuilio, B. Vollmer, M.-T. Mackmull, N. Banterle, L. Parca, P. Kastritis, K. Buczak, S. Mosalaganti, W. Hagen, A. Andres-Pons, E. A. Lemke, P. Bork, W. Antonin, J. S. Glavy, 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, I. Zagoriy, A. Halavatyi, J. Mahamid, J. Kosinski, M. Beck, Nuclear pores constrict upon energy depletion. *bioRxiv* (2020), in press, doi:10.1101/2020.07.30.228585.

24. 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*. **352** (2016), doi:10.1126/science.aaf1015.

25. 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**, 32–40 (2018).

26. 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).

27. 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, 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).

28. M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. 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. C. 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).

29. 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).

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

31. J. Mahamid, S. Pfeffer, M. Schaffer, E. Villa, R. Danev, L. Kuhn 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).

32. 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).

33. Bley et al., Submitted. *bioRxiv* (2021).

34. Petrovic et al., Submitted. *bioArxiv* (2021).

35. 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, Protein complex prediction with AlphaFold-Multimer. *bioRxiv* (2021), doi:10.1101/2021.10.04.463034.

36. M. Mirdita, S. Ovchinnikov, M. Steinegger, ColabFold - Making protein folding accessible to all. *bioRxiv* (2021), doi:10.1101/2021.08.15.456425.

37. V. Rantos, K. Karius, J. Kosinski, Integrative structural modelling of macromolecular complexes using Assembline, *bioRxiv* (2021), in press, doi:10.1101/2021.04.06.438590.

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, in *Endocrine Research* (Endocr Res, 2004), vol. 30, pp. 891–899.

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. 2020 306*. **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. Trevisan Doimo De Azevedo, C. M. Fusco, E. Kaindl, J. Baumbach, J. D. Langer, E. M. Schuman, K. R. Patil, G. Hummer, V. Benes, M. Beck, Co-translational assembly counteracts promiscuous interactions. *bioRxiv* (2021), doi:10.1101/2021.07.13.452229.

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. 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).

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, *Proteins Struct. Funct. Bioinforma.*, in press, doi:10.1002/prot.26171.

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.* (2021), doi:10.1002/prot.26248.

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.* (2021), doi:10.1002/prot.26257.

67. P. Bryant, G. Pozzati, A. Elofsson, Improved prediction of protein-protein interactions using AlphaFold2 and extended multiple-sequence alignments. *bioRxiv* (2021).

68. I. R. Humphreys, J. Pei, M. Baek, A. Krishnakumar, I. Anishchenko, S. Ovchinnikov, J. Zhang, T. J. Ness, S. Banjade, S. Bagde, V. G. Stancheva, X.-H. Li, K. Liu, Z. Zheng, D. J. Barrero, U. Roy, I. S. Fernández, B. Szakal, D. Branzei, E. C. Greene, S. Biggins, S. Keeney, E. A. Miller, J. C. Fromme, T. L. Hendrickson, Q. Cong, D. Baker, Structures of core eukaryotic protein complexes, *bioRxiv* (2021), doi:10.1101/2021.09.30.462231.

69. 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).

70. 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).

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

72. 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).

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

74. 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).

75. B. Turonova, turonova/novaSTA: novaSTA (2020), , doi:10.5281/ZENODO.3973623.

76. 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).

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

78. 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 (2017).

79. 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).

80. 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.

81. 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).

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

83. 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).

84. 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).

85. 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.

86. 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).

87. B. Webb, A. Sali, *Curr. Protoc. Bioinforma.*, in press, doi:10.1002/cpbi.3.

88. 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).

89. 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).

90. 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).

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

92. 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).

93. 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).

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

95. 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).

96. K. W, S. C, Dictionary of protein secondary structure: pattern recognition of hydrogen-

bonded and geometrical features. *Biopolymers*. **22**, 2577–2637 (1983).

97. 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).

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

99. 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).

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

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

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

103. 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).

Acknowledgments

We acknowledge support from the Electron Microscopy Core Facility (EMCF) and IT services of EMBL Heidelberg.

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 Hesse.

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 analysed the HEK control and GP210Δ data. MTM performed the BioID experiments. FS and KB prepared the HEK GP210Δ 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, MB, and JK conceptualized the study, supervised the project, and wrote the manuscript.

CRediT roles:

Conceptualization: SM, AOK, GH, MB, JK

Methodology: SM, AOK, BT, MS, JK

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

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

Funding acquisition: MB

Project administration: SM, GH, MB, JK

Supervision: SM, GH, MB, JK

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

Prior publication, EM maps associated with the manuscript will be deposited in the Electron Microscopy Data Bank (EMDB) and the models will be deposited into the Protein Data Bank (PDB).