ARTICLES

The molecular architecture of the nuclear pore complex

Frank Alber[1]*, Svetlana Dokudovskaya[2]*†, Liesbeth M. Veenhoff[2]*†, Wenzhu Zhang[3], Julia Kipper[2]†, Damien Devos[1]†, Adisetyantari Suprapto[2]†, Orit Karni-Schmidt[2]†, Rosemary Williams[2], Brian T. Chait[3], Andrej Sali[1] & Michael P. Rout[2]

Nuclear pore complexes (NPCs) are proteinaceous assemblies of approximately 50 MDa that selectively transport cargoes across the nuclear envelope. To determine the molecular architecture of the yeast NPC, we collected a diverse set of biophysical and proteomic data, and developed a method for using these data to localize the NPC's 456 constituent proteins (see the accompanying paper). Our structure reveals that half of the NPC is made up of a core scaffold, which is structurally analogous to vesicle-coating complexes. This scaffold forms an interlaced network that coats the entire curved surface of the nuclear envelope membrane within which the NPC is embedded. The selective barrier for transport is formed by large numbers of proteins with disordered regions that line the inner face of the scaffold. The NPC consists of only a few structural modules that resemble each other in terms of the configuration of their homologous constituents, the most striking of these being a 16-fold repetition of 'columns'. These findings provide clues to the evolutionary origins of the NPC.

Nuclear pore complexes (NPCs) are large (~50 MDa) proteinaceous assemblies spanning the nuclear envelope, where they function as mediators of bidirectional exchange between the nucleoplasmic and cytoplasmic compartments[1]. Nucleocytoplasmic transport of macromolecular cargoes depends on their recognition by transport factors, which interact with the NPC to carry cargoes across the nuclear envelope[2,3].

NPCs show a broad degree of compositional and structural conservation among all eukaryotes studied[4,5]. Each NPC contains at least 456 individual protein molecules and is composed of ~30 distinct proteins (nucleoporins)[6,7]. Electron microscopy studies in several organisms have revealed that the general morphology of the NPC is conserved[8–13]. These studies show the NPC to be a doughnut-shaped structure, consisting of eight spokes arranged radially around a central channel that serves as the conduit for macromolecular transport. Each NPC spans the nuclear envelope through a pore formed by the fusion of the inner and outer nuclear envelope membranes. Numerous filamentous structures project from the NPC into the cytoplasm and nucleoplasm[1,5].

Using an approach described in the accompanying paper, we have determined the molecular architecture of the *Saccharomyces cerevisiae* NPC[14]. Here, we describe and discuss the major features revealed by this analysis.

Architectural overview of the NPC

Our NPC structure, when viewed as projections of the mass density of all the nucleoporins (Fig. 1a), closely resembles maps and images obtained using electron microscopy, although in the electron microscopy maps the positions of individual nucleoporins were not defined[8–13]. We have excluded from our map proteins not previously considered 'integral' components of the NPC, such as Mlp1, Mlp2, transport factors and Nup2 (ref. 6). There is no basket in our map, probably because Mlp1 and Mlp2, the probable basket-forming proteins[15], are absent. We do, however, observe a low-density cloud, corresponding to the unstructured regions of FG (phenylalanine-glycine) repeat nucleoporins (see below), filling the central channel and projecting into the nucleoplasm and cytoplasm. The diameter of the central channel in the absence of the unstructured components (Fig. 1b) is ~38 nm, matching the known maximum sizes of transported particles[16]. The diameter of the whole NPC is ~98 nm, and the height of the structured portion is ~37 nm (Fig. 1b), again close to the dimensions observed by electron microscopy[11,13].

A fundamental symmetry unit of the NPC is the spoke[13]. Each spoke can be divided into two almost identical nucleoplasmic and cytoplasmic halves, joined at the NPC's equator and related to each other by a dyad axis (Fig. 2 and Supplementary Movie). Together, eight spokes connect to form several co-axial rings (Fig. 2); we refer to these as the membrane rings (not to be confused with the nuclear envelope membrane itself), the two outer rings at the nucleoplasmic and cytoplasmic periphery, and the two adjacent inner rings. Groups of nucleoporins that we term linker nucleoporins are attached between both sets of outer and inner rings (Fig. 2). Another group of related proteins is collectively termed FG nucleoporins. Each FG nucleoporin contains a small structured domain that serves as the anchor site to the rest of the NPC, and a larger unstructured region that contains characteristic repetitive sequence motifs termed FG repeats. The FG nucleoporins are largely exposed on the inner surface of the spokes and are anchored either to the inner rings or to the

[1]Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, and California Institute for Quantitative Biosciences, Mission Bay QB3, 1700 4th Street, Suite 503B, University of California at San Francisco, San Francisco, California 94158-2330, USA. [2]Laboratory of Cellular and Structural Biology, and [3]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA. †Present addresses: Laboratory of Nucleocytoplasmic Transport, Institut Jacques Monod, 2 place Jussieu, Tour 43, Paris 75251, France (S.D.); Department of Biochemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands (L.M.V.); German Aerospace Center (PT-DLR), Heinrich-Konen-Strasse 1, D-53227 Bonn, Germany (J.K.); Structural Bioinformatics, EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany (D.D.); Office of Technology Transfer, The Rockefeller University, 1230 York Avenue, New York, New York 10065, USA (A.S.); Herbert Irving Comprehensive Cancer Centre, Columbia University, 1130 St Nicholas Avenue, New York, New York 10032, USA (O.K.-S.).
*These authors contributed equally to this work.

linker nucleoporins; curiously, few make connections directly to the outer rings (Fig. 2). Among the most prominent features in electron micrographs of the NPC are the co-axial rings, the positions of which are similar to those of rings in our structure: the nucleoplasmic and cytoplasmic rings seen by electron microscopy coincide with the position of the two outer rings in our map, whereas the spoke ring complex and luminal ring coincide respectively with our map's inner rings and membrane rings (Figs 1 and 2, and Supplementary Fig. 1). Two copies per spoke of Nup82 (a linker nucleoporin) are found adjacent to the cytoplasmic outer ring. They add significant extra mass to the cytoplasmic side of the NPC compared with the nucleoplasmic side, probably accounting for much of the extra density of the cytoplasmic ring seen by electron microscopy (Figs 1 and 2)[8–13]. Details in some of these electron microscopy maps differ from our map, such as the presence of a 'central transporter' and weaker densities at the nucleoplasmic and cytoplasmic surfaces of the yeast NPC map[13]; such differences are probably due to the limits of our map's precision, as well as issues with NPC preservation, imaging and averaging for electron microscopy.

Figure 2 also illustrates how a series of discrete clusters of nucleoporins represents the 'building blocks' of the NPC. As described in the accompanying paper[14], these complexes can frequently be isolated from the NPC, demonstrating that they are indeed coherent NPC substructures. Most of the interactions holding the NPC together are heterotypic, with only a few homotypic interactions. Some nucleoporins act as 'keystones', forming significantly larger numbers of interactions than most other nucleoporins, bridging multiple NPC substructures (Fig. 10 in the accompanying paper)[14]. An example is Nic96, which bridges the inner and outer rings and also serves as an anchor site for FG nucleoporins. The FG nucleoporins themselves are mainly peripherally located and therefore each shows only a few interactions.

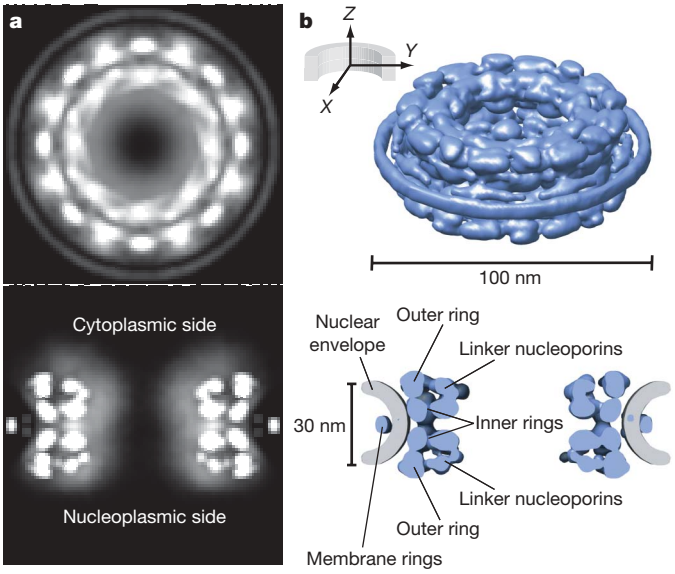

Figure 1 | Architectural overview of the NPC. a, Projections of nucleoporin mass density (white), derived from the combined localization volumes of all structured domains and the normalized localization probability of all unstructured regions[14]. Top, *en face* view showing a density projection along the *Z*-axis from $Z = -50$ nm to $Z = +50$ nm. As in electron microscopy maps of the NPC, radial arms of density correspond to spokes that interconnect to form two strong concentric rings encircling a central region containing low-density unstructured material and bounded by peripheral membrane rings, giving an overall diameter of ~98 nm. Bottom, a slice along the central *z*-axis showing a projection of density from $X = -5$ nm to $X = +5$ nm. More density can be seen on the cytoplasmic side of the NPC. The low-density unstructured material constricts the central channel to ~10 nm diameter. **b**, The structured nucleoporin domains of the NPC, represented by a density contour (blue) such that the volume of the contour corresponds approximately to the combined volume of the 456 nucleoporins comprising the NPC[14]. Top: view from a point ~30° from the equatorial plane of the NPC. Bottom: a slice along the central *Z*-axis between $X = -5$ nm and $X = 5$ nm, in which the nuclear envelope is also shown (in grey). Major features of the NPC are indicated.

A membrane-coating complex

We term the structural core of the NPC the core scaffold, formed from the outer and inner rings (Figs 2 and 3). Most linker nucleoporins and FG nucleoporin attachment sites are on the surface of the core scaffold that faces the central channel; on the outer surface are found the pore membrane (that is, that portion of the nuclear envelope into which the NPC is embedded) and the attachment sites for the membrane rings. The core scaffold probably has a key role in maintaining the stability of the nuclear envelope by ensuring coplanarity of the outer and inner nuclear envelope membranes as they exit the pore. We propose that a major function of the outer rings, by virtue of their location, is to facilitate the smooth transition of the pore membrane into the inner and outer nuclear envelope membranes (Fig. 3).

We have previously assigned fold types to domains in all the nucleoporins[17,18]. We are now in a position to determine the actual locations of these fold types in the NPC. Notably, we find that the core scaffold is made up of nucleoporins containing three distinct arrangements of only two fold types: proteins with a single α-solenoid domain (the α-solenoid fold is a superhelix of α-helices), proteins with a single β-propeller domain, and proteins composed of an amino-terminal β-propeller followed by a carboxy-terminal α-solenoid domain (Fig. 3). This latter pattern is unique to the clathrin-like proteins that coat transport vesicles, such as the clathrin heavy chain itself and Sec31, and all three fold arrangements are found in such coating complexes[17–23]. We have already proposed that the core structures of both NPCs and vesicle-coating complexes have a common evolutionary origin[17,18]. The resemblance of the core scaffold to vesicle-coating complexes does not end here; just like clathrin and the COPII coatomer, the scaffold nucleoporins form an interconnected network, such that the entire core scaffold faces and completely coats the pore membrane, with the clathrin/Sec31-like nucleoporins closer to the membrane than the α-solenoid domain proteins (Fig. 3). The vesicle-coating protein family can form different scaffolds, as demonstrated by the markedly different arrangements found in the COPII and clathrin coats[19,23]. This attribute may have made them well suited for accommodating the combination of negative and positive curvatures characteristic of the pore membrane[24].

Interestingly, five (Nup133, Nup120, Nup85, Nup170 and Nup188) of the scaffold nucleoporins contain an ALPS-like motif that has been experimentally verified to fold into an amphipathic α helix and target curved membrane regions[25]. Our structure is consistent with this observation, since at least four of these nucleoporins are found immediately adjacent to the pore membrane (Fig. 3).

Thus, not only do proteins of the core scaffold and vesicle-coating complexes share a similar fold composition, but they also appear to share a similar assembly architecture and function—namely, to form coats that curve membranes. In the case of the core scaffold, it is the nuclear envelope that is moulded into the sharply curved pore membrane (Fig. 3). The core scaffold is dominated by an evenly distributed meshwork of α-solenoid domains (Fig. 3), a fold type that is expected to be flexible[26], allowing large conformational changes without breaking protein–protein interactions (as is seen in vesicle-coating complexes). This fold flexibility could, in turn, explain the significant degree of flexibility reported for the whole NPC, necessary to accommodate the diverse sizes of nucleocytoplasmic transport cargoes and the malleability of the nuclear envelope[9,27,28].

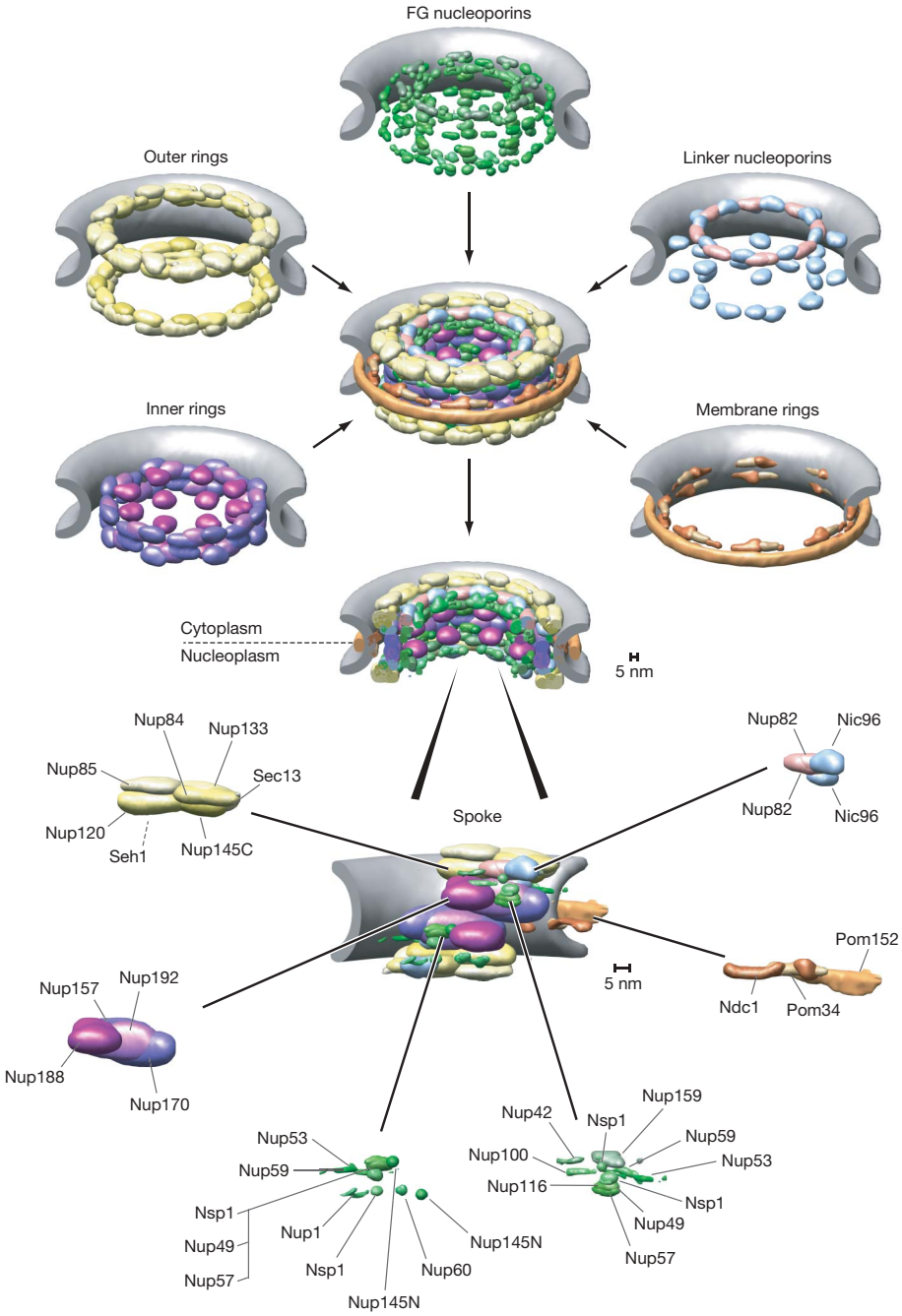

Figure 2 | Localization of major substructures and their component nucleoporins in the NPC. This figure is a single view of data presented in our Supplementary Movie. The nucleoporins are represented by their localization volumes[14] and have been coloured according to their classification into five distinct substructures on the basis of their location and functional properties: the outer rings in yellow, the inner rings in purple, the membrane rings in brown, the linker nucleoporins in blue and pink, and the FG nucleoporins (for which only the structured domains are shown) in green. The pore membrane is shown in grey. A single arbitrary repeat unit, termed the spoke, is shown dissected into its component nucleoporins. Together, the outer and inner rings connect to form the NPC's core scaffold (Fig. 3). Each of the outer rings makes connections with the adjacent linker nucleoporins and inner rings, but connects with few FG nucleoporins and no components of the membrane rings. The two inner rings are closely associated with each other at the NPC's equator and form connections with all three integral membrane proteins in the membrane rings, thereby anchoring the NPC to the nuclear envelope. The bulk of the membrane rings is formed by homo-oligomerization of the C-terminal domain of Pom152. The linker nucleoporins Nic96 and Nup82 are anchored between the inner and outer rings and have a central role in bridging the core scaffold of the NPC with the functionally important FG nucleoporins. On both the cytoplasmic and nucleoplasmic sides of each spoke, one copy of Nic96 is anchored through Nup192 and a second copy through Nup188. Whereas one copy of Nic96 carries the FG nucleoporins Nsp1, Nup57 and Nup49, the second copy forms interactions to another copy of Nsp1 and at the cytoplasmic side also interacts with Nup82. Here, Nup82 associates with the FG nucleoporins Nup159, Nup116, Nsp1 and Nup42. Thus, Nsp1 forms at least two distinct complexes in the NPC: one exclusively cytoplasmic and one disposed symmetrically[52–55]. By contrast, the FG nucleoporins found only on the nucleoplasmic side connect mainly to the inner ring nucleoporins, as do Nup53 and Nup59, both of which also face the pore membrane. The scale bars indicate the average standard deviation of the distance between a pair of neighbouring proteins in the 1,000 best-scoring configurations[14].

Attachment at the nuclear pore membrane

The membrane rings form a discrete region of the NPC, containing the three pore membrane proteins Pom152, Pom34 and Ndc1. It is the core scaffold's inner rings that interact with the membrane rings, thus anchoring the NPC to the pore membrane (Fig. 2). A component of the membrane rings (Pom152) homo-oligomerizes at its C terminus to form the ring that equatorially bounds the NPC in the perinuclear lumen[14]. This luminal portion consists of the C-terminal part of Pom152, containing domains predicted to assume the cadherin fold[18]. Members of the cadherin family are transmembrane receptors that form homophilic binding interfaces[29], probably accounting for the oligomeric luminal ring. Perhaps the NPC carries the remnants of an ancient transmembrane receptor, still attached to its vesicle-coating complex.

Transport factor docking sites and nucleocytoplasmic transport

The transport function of the NPC appears to be mediated mainly by the FG nucleoporins. The FG-repeat regions within each FG nucleoporin provide the NPC's docking sites for transport factor–cargo complexes[1,30–33]. The FG nucleoporins and especially their unstructured FG-repeat regions are the least specified part of our structure. Nevertheless, we can still draw conclusions concerning the localization of the FG-repeat regions by using a simplified representation[14]. Because these regions can adopt many different possible configurations in our calculations, on averaging they produce a cloud of low density surrounding their structurally resolved attachment sites, collectively filling and surrounding the central channel and extending into the nucleoplasm and cytoplasm (Figs 1 and 4). This spatial distribution of FG-repeat regions is consistent with 'virtual gating' models explaining the mechanism of nucleocytoplasmic transport[6,31], in which the FG-repeat density represents an effective exclusion filter for macromolecular particles that do not contain FG-repeat binding sites, but is permeable to transport factors that do possess these sites[2,6,31,34–39]. Thus, the cloud of FG-repeat regions forms a zone of selectivity around and across the NPC. The cloud thins radially from the walls of the central channel to the *Z*-axis, limiting the effective diameter of the central channel (Figs 1 and 4). In our structure, this diameter is less than 10 nm, similar to the maximal size of particles that can freely diffuse between the nucleoplasmic and cytoplasmic compartments[2]. Actively transporting cargo–transport factor complexes can displace this diffuse cloud, with the very largest pushing the cloud to the sides of the central channel up to the channel's maximum diameter of ~38 nm.

Nic96 and Nup82 provide anchor points for most of the FG nucleoporins, with connections also being made to the inner ring (Fig. 2). The FG nucleoporins can be divided into three groups according to their localization in the NPC: those that are attached mainly or exclusively to the cytoplasmic or nucleoplasmic side of the NPC, and those attached symmetrically on both sides (Fig. 4)[6]. The distributions of these groups of FG-repeat regions overlap heavily, consistent with the observed long reach of the individual FG-repeat regions[40,41]. The overlap suggests that a transport factor attached to one FG nucleoporin can readily exchange with many other surrounding FG nucleoporins, thus facilitating rapid transit across the NPC.

In contrast to most of the FG nucleoporins, a few transport factor binding sites (in particular Nup53 and Nup59) also face the pore membrane such that they are readily accessible to membrane proteins, as has been previously suggested[42]. These nucleoporins could mediate the transport of transmembrane proteins, in agreement with recent studies showing that active transport is responsible for the translocation of integral membrane proteins from the outer to the inner nuclear membrane[43,44].

Modular duplication in the evolution of the NPC

A striking pattern is revealed when we map the nucleoporins into our NPC structure based on their previously assigned fold types[18]. We find that each spoke can be divided into two parallel columns, in

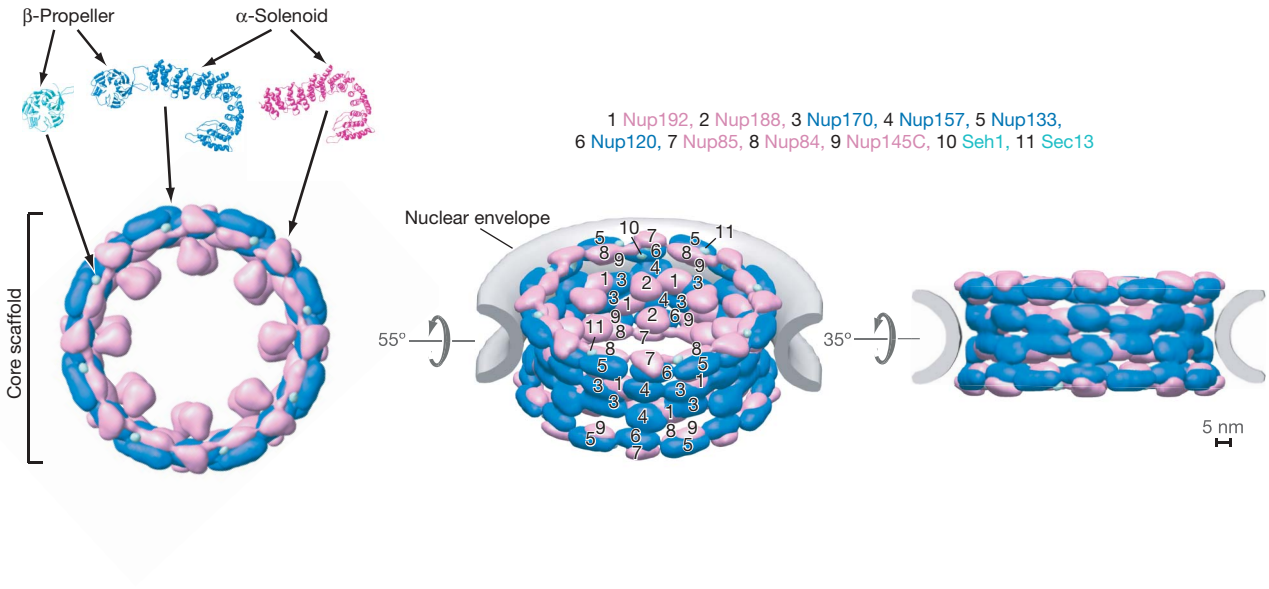

Figure 3 | The core scaffold as a membrane-coating complex. We show here the outer and inner ring nucleoporins comprising the core scaffold. The linker nucleoporins, FG nucleoporins and membrane ring are omitted for clarity. At the top of the left panel are shown the fold types comprising the nucleoporins of the core scaffold: Nup84, Nup85, Nup145C, Nup188 and Nup192 consist mainly of α-solenoid folds (pink); Sec13 and Seh1 are composed of β-propeller folds (cyan); Nup120, Nup133, Nup157 and Nup170 contain both N-terminal β-propeller folds and C-terminal α-solenoid folds (blue), an arrangement shared with clathrin and Sec31. Each of these nucleoporins is present in 16 copies to make the full 176-nucleoporin core scaffold, which is shown in three views related by the indicated rotation around an axis parallel with the NPC's equatorial plane. The localization volumes of all the α-solenoid nucleoporins (pink), all β-propeller nucleoporins (cyan), and all clathrin-like nucleoporins (blue) are indicated. The clathrin-like nucleoporins appear to be located at the outer surface of the core scaffold, adjacent to the surface of the nuclear envelope's pore membrane. Numbers on the middle panel indicate the approximate positions of each nucleoporin. The scale bar indicates the standard deviation of the distance between a pair of neighbouring proteins in the 1,000 best-scoring configurations[14].

which almost every nucleoporin in one column contains a counterpart of similar size and fold in a similar position in the adjacent column (Fig. 5). For most of these nucleoporin pairs, the relationship goes beyond structural similarity alone, as they are either clear homologues (for example, Seh1 and Sec13) or duplicate copies (for example, Nic96) (Fig. 5a)[17,18,45–48]. Furthermore, the two columns are analogous in terms of the interactions of their constituents (Fig. 5b). This observation indicates that homologous nucleoporin pairs may share similar functions, which in turn could explain why many single nucleoporin deletions are not lethal in yeast. We can thus consider the NPC as being made of a 16-fold repetition of columns, each with a similar architecture. An underlying 16-fold symmetry of the NPC has been previously suggested by electron microscopy studies[8,49], including yeast NPCs[13] (Supplementary Fig. 22). This pattern may be explained if an ancient duplication event gave rise to the two columns comprising each spoke. In yeast, many of the nucleoporin pairs in the columns (for example, Nup157 and Nup170) resulted from a whole-genome duplication[50]; in other organisms, the duplication is revealed as a single nucleoporin present in two copies (for example, vertebrate Nup155). Moreover, another duplication may have produced the outer and inner rings, as each column can be further divided into units that again are related by pairs of nucleoporins with identical fold types in equivalent positions (one member of each pair in the inner ring, the other in the outer ring). As a result, one can consider the NPC to be made of only a few structural modules, each consisting of only two to four proteins (Fig. 5c). These modules resemble each other, both in terms of being composed of homologous proteins and in their similar arrangement. The architecture of the NPC thus appears to be based on the hierarchical repetition of such structural modules that probably evolved through a series of gene duplications and divergences. Duplications and divergencies of these kinds are seen in other coating complexes, such as the clathrin–adaptin complex and COPI complex[51]. A primordial NPC may have been a multimer containing only a few different modules, rather like the clathrin and COPII coats. It may also have carried only a few kinds of FG nucleoporins, suggesting that the selective barrier was simpler. Just a few different transport factors would have been sufficient to recognize the few kinds of docking sites in this simple barrier; thus, the plethora of transport factors found in modern eukaryotes may have also evolved by duplication events, keeping pace with the evolutionary duplication and diverging specialization of the FG nucleoporins in the NPC's modules.

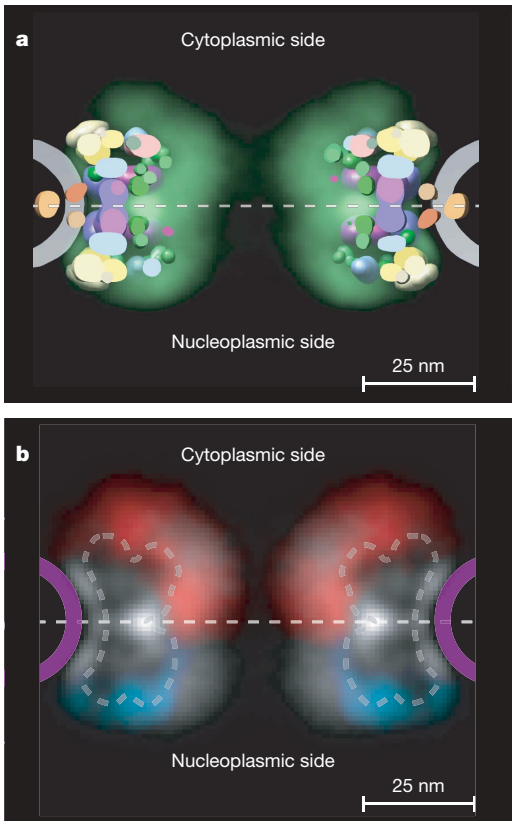

Figure 4 | Distribution of the disordered FG-repeat regions in the NPC. **a**, Slice through the NPC, as in Fig. 1b; here we show the structured domains of all nucleoporins represented by their localization volumes[14] coloured according to their classification into five distinct substructures (Fig. 2). Also shown is the localization probability of the unstructured regions of all FG nucleoporins (green cloud), visualized in Chimera[56]. **b**, Projection of the localization probabilities of the FG-repeat regions from all the FG nucleoporins is shown by a density plot, sampled in a plane perpendicular to the central *Z*-axis from $X = -5$ nm to $X = +5$ nm. Projections from the FG-repeat regions belonging to FG nucleoporins anchored mainly or exclusively on one side of the NPC are indicated: red for those that are cytoplasmically disposed (Nup42, Nup100, Nup116 and Nup159), blue for those nucleoplasmically disposed (Nup60, Nup145N and Nup1) and white for those FG nucleoporins found equally on both sides (Nup49, Nup53, Nup57, Nup59 and Nsp1). The equatorial plane of the NPC is indicated by a dashed white line, the position of the NPC density in Fig. 1 is indicated by a dashed grey line, and the position of the pore membrane is shown in purple. Although each group concentrates around its distinct anchor sites, considerable overlap can be seen between the three groups of FG-repeat regions. Peaks of density, belonging mainly to the repeat regions of Nup53 and Nup59, also face the pore membrane. A scale bar of 25 nm is shown.

Concluding remarks

We have determined the detailed molecular architecture of the NPC in the yeast *Saccharomyces cerevisiae*. Even though the primary sequence conservation between nucleoporins from different model organisms is low, the high conservation of the overall shape and predicted fold types[18] implies that the overall architecture of the NPC described here is highly conserved among eukaryotes.

Although the NPC is a complex structure, our analysis reveals underlying simplicities in its architecture. At its heart, the NPC contains a highly connected scaffold that attaches to and coats the curved pore membrane. The fold composition of the nucleoporins forming the scaffold is remarkably simple, consisting of only two different domain folds, the configurations of which resemble those found in vesicle-coating complexes—to which the NPC may therefore be evolutionarily related. This scaffold anchors disordered clouds of filaments that fill the central channel and regions proximal to the pore membrane, and project into the nucleoplasm and cytoplasm. These clouds act as a selective barrier to mediate nucleocytoplasmic trafficking. We found that proteins in each half-spoke are either present in duplicate copies or homologous pairings, sharing equivalent interaction patterns; thus, the NPC is another example of how a complicated structure can evolve from the duplication, divergence and elaboration of simple ancestral modules.

The NPC's architecture also suggests a possible assembly process, analogous to the formation of coated vesicles. In this process, the NPC's membrane proteins might serve as the receptors for the attachment of the inner ring to the nuclear envelope and polymerization of the core scaffold, forming a coat on the curved pore membrane. Attachment of the linker and FG nucleoporins to this coat would complete the NPC.

Further elucidation of the evolutionary origin, transport mechanism and assembly pathway of the NPC requires higher resolution information, encompassing the atomic structures of nucleoporins and their intermolecular arrangements. Given that our structure determination method can incorporate such information, we envision continued steady progress in describing the fine architecture of the NPC.

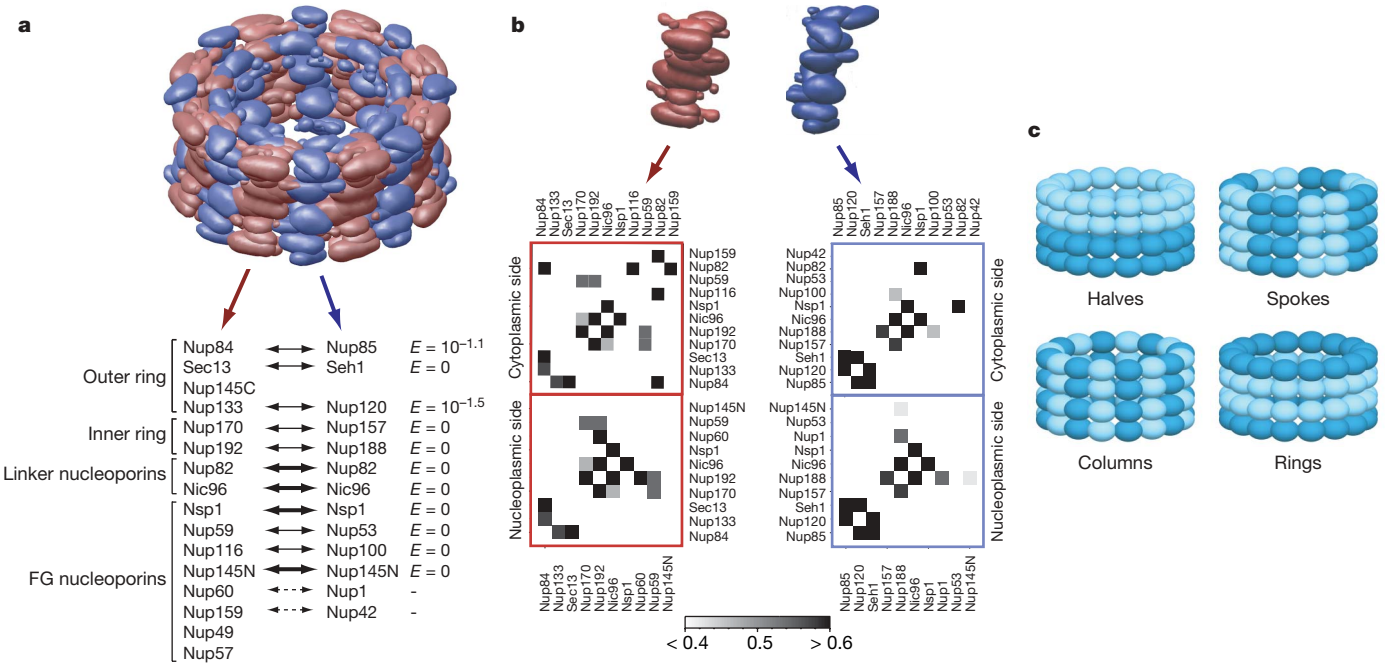

Figure 5 | Modular duplication in the NPC. a, The NPC can be divided into two alternating equivalent groups of nucleoporins arranged in parallel columns, as indicated here in red and blue. The nucleoporins in the red column are listed below left, and those in the blue column are listed below right. Almost every nucleoporin in the red column contains a counterpart in the blue column related by approximate position, as well as size and fold arrangement (pairs linked by dashed arrows), strong sequence similarity (those linked by thin arrows), or as duplicate copies with one copy in each column (thick arrows). The membrane rings and FG repeat regions were removed for clarity. Also shown are the *E*-values generated by HMMsearch[57] for the most significant local matches of the corresponding protein pairs; the *E*-value for a sequence match is the expected number of false positives per database search with a score at least as good as this sequence match. **b**, Network of protein contacts in the cytoplasmic (upper row) and nucleoplasmic (lower row) half-spokes for proteins in the red column (left) and blue column (right), showing that the homologous constituents have equivalent neighbours in each column. The networks are shown as instance contact frequency maps[14] (Supplementary Information). Notably, many contacts among proteins in one column are also present between the equivalent proteins in the second column. **c**, A large portion of the NPC can be divided into pairs of structurally similar modules: nuclear and cytoplasmic halves, eight radially disposed spokes, 16 radially disposed columns, and inner and outer rings.

METHODS SUMMARY

Our approach to structure determination can be seen as an iterative series of four steps: data generation by experiment, translation of the data into spatial restraints, calculation of an ensemble of structures by satisfaction of these restraints, and an analysis of the ensemble to produce the final structure. The experimental and computational methods are described in detail in the accompanying paper[14] and Supplementary Information.

Experimental methods. The localization of the PrA-tagged nucleoporins was determined by immuno-electron microscopy of approximately 10,000 gold particles, obtained from pre-embedding labelled nuclear envelopes isolated from tagged strains[6]. Stokes radii were derived from ultracentrifugation velocity gradient sedimentation experiments on the individual nucleoporins and the Nup84 complex. Affinity purifications of all PrA-tagged nucleoporins were performed from solubilized nuclear envelope, highly enriched NPC fractions[6], or lysates produced by whole cell cryolysis. Approximately 20 variants of extraction buffers were used to generate different complexes on IgG Sepharose resins or IgG-magnetic beads. Co-purified proteins were identified by mass spectrometry[7]. Overlay assays were used to monitor direct binding of a nucleoporin in solution to nucleoporins immobilized on nitrocellulose. The structure and composition of the 'Pom rings' (ref. 14) isolated from NPCs were determined by electron microscopy and mass spectrometry.

Computational methods. The structure calculation is expressed as an optimization problem, a solution of which requires three main components: (1) a representation of the assembly in terms of its constituent parts; (2) a scoring function, consisting of individual spatial restraints that encode all the data; and (3) an optimization of the scoring function, which aims to yield structures that satisfy the restraints.

Each nucleoporin was represented by a flexible chain consisting of a small number of connected spheres. The number and radii of the spheres were chosen to reproduce the nucleoporin mass and Stokes radius. The nuclear envelope in the region of the pore membrane[13] was represented by many partially overlapping spheres, each with a radius of 2.25 nm.

The scoring function captures information about the structure of the NPC and is a sum of restraints of various types. These restraints encode what is known about nucleoporin and nuclear envelope excluded volumes (from the protein sequence and ultracentrifugation), nucleoporin positions (from immuno-electron microscopy), nucleoporin 'connectivity' in composites (from affinity purification), nucleoporin–nucleoporin interactions (from affinity purification and overlay assays), complex diameters (from ultracentrifugation), and the eight-fold and two-fold symmetries of the NPC (from electron microscopy).

The optimization protocol combines the techniques of simulated annealing, molecular dynamics, and conjugate gradient minimization. To sample the space of solutions well, independent optimizations of randomly generated initial configurations were performed until 1,000 structures were generated that satisfied all input restraints (that is, the ensemble); ~200,000 trials were required.

The protein localization probability is the probability that a given volume element of a $100 \times 100 \times 100$ grid with the spacing of 1 nm is occupied by a particular protein. It is calculated for each protein from all superposed configurations in the ensemble. Next, for a given protein, the volume elements are sorted by their localization probability values. The localization volume of the protein then corresponds to the top-ranked elements, the volume of which sums to the protein volume, estimated from its molecular mass. The localization volume of a protein reveals its most probable localization.

Received 20 April; accepted 22 October 2007.

1. Lim, R. Y. & Fahrenkrog, B. The nuclear pore complex up close. *Curr. Opin. Cell Biol.* **18**, 342–347 (2006).
2. Macara, I. G. Transport into and out of the nucleus. *Microbiol. Mol. Biol. Rev.* **65**, 570–594 (2001).
3. Weis, K. Nucleocytoplasmic transport: cargo trafficking across the border. *Curr. Opin. Cell Biol.* **14**, 328–335 (2002).
4. Hetzer, M., Walther, T. C. & Mattaj, I. W. Pushing the envelope: Structure, function, and dynamics of the nuclear periphery. *Annu. Rev. Cell Dev. Biol.* **21**, 347–380 (2005).

5. Tran, E. J. & Wente, S. R. Dynamic nuclear pore complexes: life on the edge. *Cell* **125,** 1041–1053 (2006).
6. Rout, M. P. *et al.* The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* **148,** 635–651 (2000).
7. Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T. & Matunis, M. J. Proteomic analysis of the mammalian nuclear pore complex. *J. Cell Biol.* **158,** 915–927 (2002).
8. Akey, C. W. & Radermacher, M. Architecture of the *Xenopus* nuclear pore complex revealed by three-dimensional cryo-electron microscopy. *J. Cell Biol.* **122,** 1–19 (1993).
9. Beck, M. *et al.* Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* **306,** 1387–1390 (2004).
10. Hinshaw, J. E., Carragher, B. O. & Milligan, R. A. Architecture and design of the nuclear pore complex. *Cell* **69,** 1133–1141 (1992).
11. Kiseleva, E. *et al.* Yeast nuclear pore complexes have a cytoplasmic ring and internal filaments. *J. Struct. Biol.* **145,** 272–288 (2004).
12. Stoffler, D. *et al.* Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport. *J. Mol. Biol.* **328,** 119–130 (2003).
13. Yang, Q., Rout, M. P. & Akey, C. W. Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Mol. Cell* **1,** 223–234 (1998).
14. Alber, F. *et al.* Determining the architectures of macromolecular assemblies. *Nature* doi:10.1038/nature06404 (this issue).
15. Krull, S., Thyberg, J., Bjorkroth, B., Rackwitz, H. R. & Cordes, V. C. Nucleoporins as components of the nuclear pore complex core structure and Tpr as the architectural element of the nuclear basket. *Mol. Biol. Cell* **15,** 4261–4277 (2004).
16. Pante, N. & Kann, M. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol. Biol. Cell* **13,** 425–434 (2002).
17. Devos, D. *et al.* Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol.* **2,** e380 (2004).
18. Devos, D. *et al.* Simple fold composition and modular architecture of the nuclear pore complex. *Proc. Natl Acad. Sci. USA* **103,** 2172–2177 (2006).
19. Fotin, A. *et al.* Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **432,** 573–579 (2004).
20. Stagg, S. M. *et al.* Structure of the Sec13/31 COPII coat cage. *Nature* **439,** 234–238 (2006).
21. ter Haar, E., Musacchio, A., Harrison, S. C. & Kirchhausen, T. Atomic structure of clathrin: a β propeller terminal domain joins an α zigzag linker. *Cell* **95,** 563–573 (1998).
22. Dokudovskaya, S. *et al.* Protease accessibility laddering: a proteomic tool for probing protein structure. *Structure* **14,** 653–660 (2006).
23. Fath, S., Mancias, J. D., Bi, X. & Goldberg, J. Structure and organization of coat proteins in the COPII cage. *Cell* **129,** 1325–1336 (2007).
24. Antonin, W. & Mattaj, I. W. Nuclear pore complexes: round the bend? *Nature Cell Biol.* **7,** 10–12 (2005).
25. Drin, G. *et al.* A general amphipathic α-helical motif for sensing membrane curvature. *Nature Struct. Mol. Biol.* **14,** 138–146 (2007).
26. Conti, E., Muller, C. W. & Stewart, M. Karyopherin flexibility in nucleocytoplasmic transport. *Curr. Opin. Struct. Biol.* **16,** 237–244 (2006).
27. Akey, C. W. Structural plasticity of the nuclear pore complex. *J. Mol. Biol.* **248,** 273–293 (1995).
28. Hinshaw, J. E. & Milligan, R. A. Nuclear pore complexes exceeding eightfold rotational symmetry. *J. Struct. Biol.* **141,** 259–268 (2003).
29. Bryant, D. M. & Stow, J. L. The ins and outs of E-cadherin trafficking. *Trends Cell Biol.* **14,** 427–434 (2004).
30. Strawn, L. A., Shen, T., Shulga, N., Goldfarb, D. S. & Wente, S. R. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nature Cell Biol.* **6,** 197–206 (2004).
31. Rout, M. P., Aitchison, J. D., Magnasco, M. O. & Chait, B. T. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* **13,** 622–628 (2003).
32. Denning, D. P., Patel, S. S., Uversky, V., Fink, A. L. & Rexach, M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc. Natl Acad. Sci. USA* **100,** 2450–2455 (2003).
33. Liu, S. M. & Stewart, M. Structural basis for the high-affinity binding of nucleoporin Nup1p to the *Saccharomyces cerevisiae* importin-β homologue, Kap95p. *J. Mol. Biol.* **349,** 515–525 (2005).
34. Peters, R. Translocation through the nuclear pore complex: selectivity and speed by reduction-of-dimensionality. *Traffic* **6,** 421–427 (2005).
35. Ribbeck, K. & Gorlich, D. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* **20,** 1320–1330 (2001).
36. Isgro, T. A. & Schulten, K. Binding dynamics of isolated nucleoporin repeat regions to importin-β. *Structure* **13,** 1869–1879 (2005).
37. Isgro, T. A. & Schulten, K. Association of nuclear pore FG-repeat domains to NTF2 import and export complexes. *J. Mol. Biol.* **366,** 330–345 (2007).
38. Stewart, M. Molecular mechanism of the nuclear protein import cycle. *Nature Rev. Mol. Cell Biol.* **8,** 195–208 (2007).
39. Zilman, A., Di Talia, S., Chait, B. T., Rout, M. P. & Magnasco, M. O. Efficiency, selectivity, and robustness of nucleocytoplasmic transport. *PLoS Comput. Biol.* **3,** e125 (2007).
40. Paulillo, S. M. *et al.* Nucleoporin domain topology is linked to the transport status of the nuclear pore complex. *J. Mol. Biol.* **351,** 784–798 (2005).
41. Lim, R. Y. *et al.* Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. *Proc. Natl Acad. Sci. USA* **103,** 9512–9517 (2006).
42. Hawryluk-Gara, L. A., Shibuya, E. K. & Wozniak, R. W. Vertebrate Nup53 interacts with the nuclear lamina and is required for the assembly of a Nup93-containing complex. *Mol. Biol. Cell* **16,** 2382–2394 (2005).
43. King, M. C., Lusk, C. P. & Blobel, G. Karyopherin-mediated import of integral inner nuclear membrane proteins. *Nature* **442,** 1003–1007 (2006).
44. Saksena, S., Summers, M. D., Burks, J. K., Johnson, A. E. & Braunagel, S. C. Importin-α-16 is a translocon-associated protein involved in sorting membrane proteins to the nuclear envelope. *Nature Struct. Mol. Biol.* **13,** 500–508 (2006).
45. Aitchison, J. D., Rout, M. P., Marelli, M., Blobel, G. & Wozniak, R. W. Two novel related yeast nucleoporins Nup170p and Nup157p: complementation with the vertebrate homologue Nup155p and functional interactions with the yeast nuclear pore-membrane protein Pom152p. *J. Cell Biol.* **131,** 1133–1148 (1995).
46. Marelli, M., Aitchison, J. D. & Wozniak, R. W. Specific binding of the karyopherin Kap121p to a subunit of the nuclear pore complex containing Nup53p, Nup59p, and Nup170p. *J. Cell Biol.* **143,** 1813–1830 (1998).
47. Siniossoglou, S. *et al.* A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. *Cell* **84,** 265–275 (1996).
48. Wente, S. R., Rout, M. P. & Blobel, G. A new family of yeast nuclear pore complex proteins. *J. Cell Biol.* **119,** 705–723 (1992).
49. Unwin, P. N. & Milligan, R. A. A large particle associated with the perimeter of the nuclear pore complex. *J. Cell Biol.* **93,** 63–75 (1982).
50. Scannell, D. R., Butler, G. & Wolfe, K. H. Yeast genome evolution-the origin of the species. *Yeast*. (in the press).
51. Schledzewski, K., Brinkmann, H. & Mendel, R. R. Phylogenetic analysis of components of the eukaryotic vesicle transport system reveals a common origin of adaptor protein complexes 1, 2, and 3 and the F subcomplex of the coatomer COPI. *J. Mol. Evol.* **48,** 770–778 (1999).
52. Grandi, P. *et al.* A novel nuclear pore protein Nup82p which specifically binds to a fraction of Nsp1p. *J. Cell Biol.* **130,** 1263–1273 (1995).
53. Belgareh, N. *et al.* Functional characterization of a Nup159p-containing nuclear pore subcomplex. *Mol. Biol. Cell* **9,** 3475–3492 (1998).
54. Bailer, S. M. *et al.* Nup116p associates with the Nup82p-Nsp1p-Nup159p nucleoporin complex. *J. Biol. Chem.* **275,** 23540–23548 (2000).
55. Bailer, S. M., Balduf, C. & Hurt, E. The Nsp1p carboxy-terminal domain is organized into functionally distinct coiled-coil regions required for assembly of nucleoporin subcomplexes and nucleocytoplasmic transport. *Mol. Cell. Biol.* **21,** 7944–7955 (2001).
56. Pettersen, E. F. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).
57. Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33,** W244–W248 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. Shio for performing the electron microscopic studies; J. Fanghänel, M. Niepel and C. Strambio-de-Castillia for help in developing the affinity purification techniques; M. Magnasco for discussions and advice; A. Kruchinsky for assistance with mass spectrometry; M. Topf, D. Korkin, F. Davis, M. S. Madhusudan, M.-Y. Shen, F. Foerster, N. Eswar, M. Kim, D. Russell, B. Peterson and B. Webb for many discussions about structure characterization by satisfaction of spatial restraints; C. Johnson, S. G. Parker, and C. Silva, T. Ferrin and T. Goddard for preparation of some figures; and S. Pulapura and X. J. Zhou for their help with the design of the conditional diameter restraint. We are grateful to J. Aitchison for discussion and suggestions. We also thank all other members of the Chait, Rout and Sali laboratories for their assistance. We acknowledge support from an Irma T. Hirschl Career Scientist Award (M.P.R.), a Sinsheimer Scholar Award (M.P.R.), a grant from the Rita Allen Foundation (M.P.R.), a grant from the American Cancer Society (M.P.R.), the Sandler Family Supporting Foundation (A.S.), the Human Frontier Science Program (A.S., L.M.V.), NSF (A.S.), and grants from the National Institutes of Health (B.T.C., M.P.R., A.S.), as well as computer hardware gifts from R. Conway, M. Homer, Intel, Hewlett-Packard, IBM and Netapp (A.S.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.P.R (rout@rockefeller.edu), A.S. (sali@salilab.org), or B.T.C. (chait@rockefeller.edu).