Published October 3, 2016

Simple rules for passive diffusion through the nuclear pore complex

Benjamin L. Timney,[1]* Barak Raveh,[2,3]* Roxana Mironska,[1] Jill M. Trivedi,[1] Seung Joong Kim,[2,3] Daniel Russel,[2,3] Susan R. Wente,[4] Andrej Sali,[2,3] and Michael P. Rout[1]

[1]Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, NY 10065
[2]Department of Bioengineering and Therapeutic Sciences and [3]Department of Pharmaceutical Chemistry, California Institute for Quantitative Biosciences, University of California, San Francisco, San Francisco, CA 94158
[4]Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN 37232

Passive macromolecular diffusion through nuclear pore complexes (NPCs) is thought to decrease dramatically beyond a 30–60-kD size threshold. Using thousands of independent time-resolved fluorescence microscopy measurements in vivo, we show that the NPC lacks such a firm size threshold; instead, it forms a soft barrier to passive diffusion that intensifies gradually with increasing molecular mass in both the wild-type and mutant strains with various subsets of phenylalanine-glycine (FG) domains and different levels of baseline passive permeability. Brownian dynamics simulations replicate these findings and indicate that the soft barrier results from the highly dynamic FG repeat domains and the diffusing macromolecules mutually constraining and competing for available volume in the interior of the NPC, setting up entropic repulsion forces. We found that FG domains with exceptionally high net charge and low hydropathy near the cytoplasmic end of the central channel contribute more strongly to obstruction of passive diffusion than to facilitated transport, revealing a compartmentalized functional arrangement within the NPC.

Introduction

The nuclear pore complex (NPC), residing in the nuclear envelope (NE), is the main mediator of transport between the nucleus and the cytoplasm in all eukaryotic cells (Christie et al., 2016; Knockenhauer and Schwartz, 2016; Musser and Grünwald, 2016). It also plays an important part in transcriptional regulation, signal transduction, cell growth, and the cell cycle (Strambio-De-Castillia et al., 2010; Schellhaus et al., 2016). The NPC is one of the largest macromolecular complexes found in eukaryotic cells. It is made of hundreds of copies of broadly conserved proteins, collectively referred to as nucleoporins, or nups. Approximately one third of these are classed as FG nups, because they contain domains with long stretches of repeated phenylalanine-glycine (FG) motifs separated by hydrophilic flexible linkers. FG nups play a key role in forming the selective barrier to transport. Although the NPC is remarkably robust to extensive deletions of FG repeat domains in *Saccharomyces cerevisiae* (Strawn et al., 2004; Terry and Wente, 2007), alterations to FG nups are associated with many defects, including the onset of numerous types of cancers (Simon and Rout, 2014), as well as with viral infection, in which they might allow undesirable access to the nucleus or compromise the cell's antiviral defenses (Simon and Rout, 2014; Mettenleiter, 2016).

Macromolecules traverse the NPC through either facilitated or passive diffusion. Facilitated diffusion is the rapid translocation of cargo-carrying nuclear transport receptors (NTRs) through interactions with FG domains, and it is generally but not necessarily coupled to energy input from the nuclear GTPase Ran (Quimby and Dasso, 2003). Passive diffusion, in contrast, is the equilibration of ions, small molecules, and small- to medium-sized macromolecules between the cytoplasm and the nucleoplasm, independent of NTRs and GTP hydrolysis. FG domains obstruct the passive diffusion of large macromolecules (Paine and Feldherr, 1972; Paine et al., 1975; Keminer and Peters, 1999; Ribbeck and Görlich, 2001; Mohr et al., 2009; Ma et al., 2012), restricting access to the nucleus and contributing to the retention of nuclear and cytoplasmic factors in their appropriate subcellular compartments. Hence, whether a given macromolecule can passively diffuse through the NPC on a physiologically relevant time scale has an important bearing on our understanding of nuclear maintenance and regulation.

A rigid or soft barrier to passive diffusion?

A well-established determinant of passive diffusion rates is molecular mass, which was observed to affect passive permeability

*B.L. Timney and B. Raveh contributed equally to this paper.

Correspondence to Andrej Sali: sali@salilab.org; or Michael P. Rout: rout@rockefeller.edu

Abbreviations used: DLS, dynamic-light scattering; FG, phenylalanine-glycine; LASSO, Least Absolute Shrinkage and Selection Operator; N/C, nucleus/cytoplasm; NE, nuclear envelope; NPC, nuclear pore complex; NTR, nuclear transport receptor; nup, nucleoporin; PCA, principal component analysis; PrA, Protein A; SAXS, small-angle x-ray scattering; WT, wild type.

Supplemental Material can be found at:
/content/suppl/2016/09/27/jcb.201601004.DC1.html

The Rockefeller University Press $30.00
J. Cell Biol. Vol. 215 No. 1 57–76
www.jcb.org/cgi/doi/10.1083/jcb.201601004

Published October 3, 2016

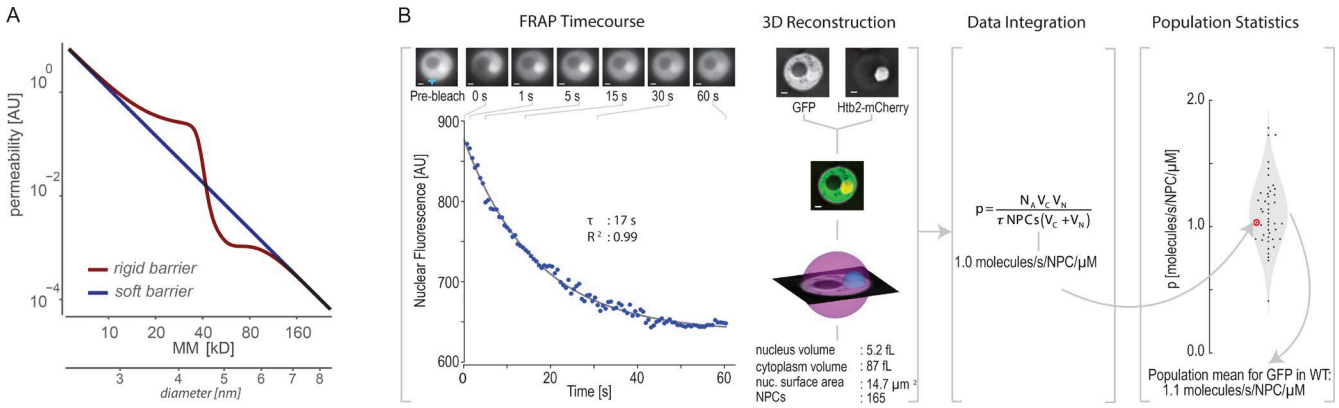


Figure 1. **Characterizing the passive permeability barrier of yeast NPCs.** (A) Expected decay in passive permeability rates with molecular mass in a rigid versus soft barrier to passive transport. (B) Permeability of yeast NPCs determined in vivo using a FRAP approach. From left to right: At time $t = 0$ s, the cytoplasmic pool of reporter macromolecules is photobleached with a spot-bleaching laser (cyan crosshair), and the nuclear fluorescence in a single cell is plotted over time, as the nuclear pool of reporters equilibrates with the cytoplasm through passive diffusion. Bars, 1 µm. Note, at steady state, the cytoplasm can appear ~20% less bright than the nucleus because of vesicles and other structures below imaging resolution that do not contain GFP. We have not adjusted for this small deviance from the ideal N/C of 1, as the effect is small and consistent between strains. The mean transport time τ is the time constant of the exponential decay in nuclear fluorescence; the nuclear volume, cytoplasmic volumes, and number of NPCs are then quantified based on 3D reconstruction of the cell, using a complete image series through the z axis acquired after the time course. These data are integrated to compute the permeability coefficient in a single cell. The population mean of several such measurements from many cells is computed from multiple independent measurements, where each dot in the violin plot indicates a permeability coefficient from a single cell. AU, arbitrary units.

in seminal studies by Paine et al. in the 1970s (Paine and Feldherr, 1972; Paine et al., 1975) using state-of-the-art methods for that time, suggesting a molecular mass size threshold of 30–60 kD for permeation of the nucleus. Recent in vitro studies over the last two decades also suggested a relatively sharp threshold of 30–60 kD (Keminer and Peters, 1999; Ribbeck and Görlich, 2001; Mohr et al., 2009; Ma et al., 2012), although such in vitro systems have recently been shown to be quite sensitive to the precise experimental conditions (e.g., the concentration of transport receptors and the nuclear factor Ran; Ma et al., 2012; Lowe et al., 2015). Larger molecules up to 230 kD in size have also been observed to permeate the NPCs in vivo on the time scale of many minutes to hours (Wang and Brattain, 2007; Popken et al., 2015). Nonetheless, the prevailing functional model of passive transport, which we term the "rigid barrier" model, is that of a barrier with a firm size threshold of 40 kD for passive diffusion (Christie et al., 2016; Knockenhauer and Schwartz, 2016; Musser and Grünwald, 2016; Schmidt and Görlich, 2016). According to this prevailing view, the slow permeation of larger macromolecules across the NPC is regarded as residual leakage beyond such a firm threshold (Kirli et al., 2015; Fig. 1 A, red curve). However, this supposed "leakage" may be also interpreted by an alternative "soft barrier" model of transport. In this model, passive transport rates are expected to decrease much more gradually with increasing molecular mass than previously realized (Fig. 1 A, black curve).

Whether the barrier to passive transport is functionally rigid or soft is also closely related to the biophysical mechanism by which the NPC performs its function. Several biophysically inspired models have attempted over the years to explain how the FG domains form the selective barrier to transport, stirring up lively debate in the literature about their biophysical form and behavior (Ribbeck and Görlich, 2002; Rout et al., 2003; Peters, 2005; Lim et al., 2006; Patel et al., 2007; Zilman et al., 2007, 2010; Jovanovic-Talisman et al., 2009; Tetenbaum-Novatt and Rout, 2010; Yamada et al., 2010; Atkinson et al., 2013; Kapinos et al., 2014; Tran et al., 2014; Vovk et al., 2016; Zahn et al., 2016). However, the evidence for these various mechanistic models is generally indirect, because existing experimental technologies are limited in their ability to probe FG repeats within NPCs of living cells at sufficient spatial and temporal resolutions and because of possible differences in the system behavior under in vivo and in vitro conditions. Finally, although many of these models postulate a relatively homogenous distribution of FG repeats in the NPC, it has also been suggested that FG nups in different regions of the NPC may play specific functional roles (Rout and Wente, 1994; Patel et al., 2007; Yamada et al., 2010; Popken et al., 2015). The extent of this functional partitioning and the precise roles of different FG nups remain unclear.

To discern between the rigid barrier and soft barrier functional models, we have used time-resolved fluorescence microscopy that allowed us to quantify the permeability of NPCs to macromolecules of varying molecular sizes, in living *S. cerevisiae* cells, resulting in a rigorous characterization of the relation between molecular size and passive permeability. We also used theoretical modeling and extensive coarse-grained computer simulations to assess possible biophysical models that could explain our empirical observations. To study the effect of various FG domains on passive transport, we measured changes from wild-type (WT) passive permeability in a set of combinatorially deleted mutant *S. cerevisiae* strains, in which substantial subsets of FG domains were deleted from the NPC. We compared these measurements to simulations of NPCs with depleted FG domains. Finally, we used statistical data analyses to identify the most prominent features of FG domains that form the barrier to passive permeability.

Results

Quantification of NPC permeability in vivo

To characterize the in vivo function of the NPC as a barrier to passive diffusion of macromolecules, we enhanced our FRAP (Reits and Neefjes, 2001) assay of transport rates of GFP fusion proteins (Timney et al., 2006; Lord et al., 2015),

Published October 3, 2016

making it substantially more automated and quantitative by integrating several cell biological measurements in a unified pipeline (Fig. 1 B).

In each FRAP experiment, we expressed GFP-labeled substrates in live cells, photobleached the cytoplasmic pool at time $t = 0$, and measured the nuclear fluorescence over time. As expected of Fick's law applied to porous membranes (Renkin, 1954), the levels of nuclear fluorescence decayed exponentially over time, following first-order kinetics with a characteristic time constant τ as the labeled substrate equilibrated between the nuclear volume V_N and the cytoplasmic volume V_C through N NPCs (Fig. 1 B). At the microscopic level, τ is the mean timespan between two consecutive translocation events for a single substrate macromolecule. In bulk, τ is proportional to $t_{1/2}$, the $t_{1/2}$ time of nuclear fluorescence levels: $\tau = t_{1/2}/ln2$.

To remove cell-to-cell variability in τ caused by fluctuations in V_N, V_C, and N, we used our measurements of these morphological features in each individual FRAP experiment to compute the normalized time constant $\hat{\tau}$, the hypothetical value of τ in a typical cell with mean values of $V_N = 4.8$ fL, $V_C = 60$ fL, and $n = 161$ NPCs, as well as the inversely related permeability coefficient p of a single NPC, specified in units of molecules per second per NPC per micromole, equivalent to 1.86×10^{-6} cm/s, as explained in detail in Materials and methods and in Fig. 1 B (Timney et al., 2006; Lord et al., 2015). We repeated each measurement between 13 and 59 times per condition and computed the mean and 95% confidence intervals for both $\hat{\tau}$ and p.

A gradual and approximately cubic dependence on molecular mass

With this time-resolved FRAP assay, we examined passively diffusing proteins between 27 and 67 kD in mass, spanning the often-cited molecular mass size threshold to passive transport (Fig. 2 A). These substrates each comprised a GFP attached by a short linker to between zero and six repeats of the 7-kD Protein A (PrA) domain and are termed GFP-xPrA in the remainder of this paper (with x denoting the number of PrA repeats). Measurements of permeability coefficients p and normalized time constants $\hat{\tau}$ were acquired from some 236 WT cells (Fig. 2, C–E; and Video 1), each expressing one of the six differently sized GFP-xPrA substrates.

Passive permeability decreased for substrates with increasing molecular masses, as expected (Fig. 2 E). However, we did not observe any sharp cutoff in leakage at any specific threshold. Rather, we observed a strikingly simple functional relationship, in which $\hat{\tau}$ increases with molecular mass [proportional to protein volume (Fischer et al., 2004)] to the power of 3.8 ± 1.0 at the 95% confidence intervals (Fig. 3 A). Hence, although the 27-kD GFP equilibrated in a matter of seconds ($\hat{\tau} = 15 \pm 1.4$ s), the 67-kD GFP-6PrA substrate equilibrated on the timescale of minutes ($\hat{\tau} = 413 \pm 92$ s); a substantially slower rate, yet well within the time scale of biologically relevant processes, including the division time (1.25–2 h) of yeast cells.

Passive transport rates decrease because of reduced translocation efficiency

The increase in the value of $\hat{\tau}$ for each substrate could be a result of an increasing diffusion time between consecutive substrate encounters with the NPC, a decreased translocation probability per each such encounter (the transport efficiency), or a combination of both. However, the 1.6-fold decrease in the diffusion coefficients in solution from the 27-kD GFP to the 67-kD GFP-PrA (Fig. 2 B) is negligible compared with the 27-fold difference in their $\hat{\tau}$ values. Therefore, we attribute most of the increase in $\hat{\tau}$ to changes in the translocation efficiencies of the different substrates.

Approximately cubic dependence extends over a wide range of molecular masses

It was recently illustrated that passive nuclear permeability in yeast is observed on the timescale of hours for substrates with molecular masses of up to 230 kD (Popken et al., 2015), and we were curious whether or not the approximately cubic size dependence extends to these larger substrates. Although our time-resolved FRAP experiments were not practical over time scales >10 min, as cellular growth and new GFP synthesis eventually becomes significant, we could estimate the mean transport time $\hat{\tau}$ of these larger substrates based on the nuclear to cytoplasmic fluorescence ratio after equilibration reported in that study (Materials and methods). Importantly, our estimates were not scaled in any way to fit our own data and were the result of a straightforward and nonparameterized transformation to estimate the exponential decay rate. We plotted the resulting $\hat{\tau}$ values against the molecular masses of the combined dataset of 14 substrates ranging in molecular mass from 27 to 230 kD. Strikingly, the same approximately cubic relation that we observed in our own data applies over this combined dataset, and with a much better confidence interval than if we used either of the datasets alone (Fig. 3 B). Because passive diffusion rates may vary for macromolecules of similar molecular mass, we also examined two fusion proteins in which GFP was attached to either one or two domains of the completely unrelated Protein G from *Streptococcus* (GFP-1PrG and GFP-2PrG), but their passive diffusion rates were consistent with the measurements of similarly sized GFP-xPrA molecules, within 7 kD (Fig. 3, A and B). We conclude that the barrier to passive diffusion lacks any fixed molecular mass threshold and can be predicted remarkably well based on molecular mass within the size range from 27 to 230 kD, at least to first approximation.

Molecular shape has limited effect on passive diffusion

To assess the effects of molecular shape on passive transport rates, we first determined the level of molecular asymmetry for each of the labeled substrates in our 27–67 kD dataset and computed their geometric shapes using three independent methods (Fig. 3 C and Fig. S1 A): (1) analytical centrifugation was used to measure the sedimentation coefficient of each substrate, from which we calculated S_{max}/S values, which increase with increasing molecular asymmetry (Erickson, 2009); (2) dynamic-light scattering (DLS; Schmitz, 1990) was used to measure diffusion coefficients and the inversely related Stokes radii to provide an independent validation of the measured sedimentation coefficients; and (3) small-angle x-ray scattering (SAXS) was used to most directly compute the shape of the enclosing envelope of each substrate. All three methods were in outstanding agreement with each other (Fig. 3 C). Therefore, we could also rely on the sedimentation coefficients for the 110–230 kD dataset (Popken et al., 2015) to quantify the level of asymmetry of these larger substrates based on their S_{max}/S values (Fig. 3 D). The shapes of all molecules in the dataset were computed based on either their SAXS envelopes for the smaller substrates or their S_{max}/S values for the larger ones (Fig. 3 D; and Materials and methods). The GFP-xPrA and -xPrG molecules elongated gradually

Published October 3, 2016

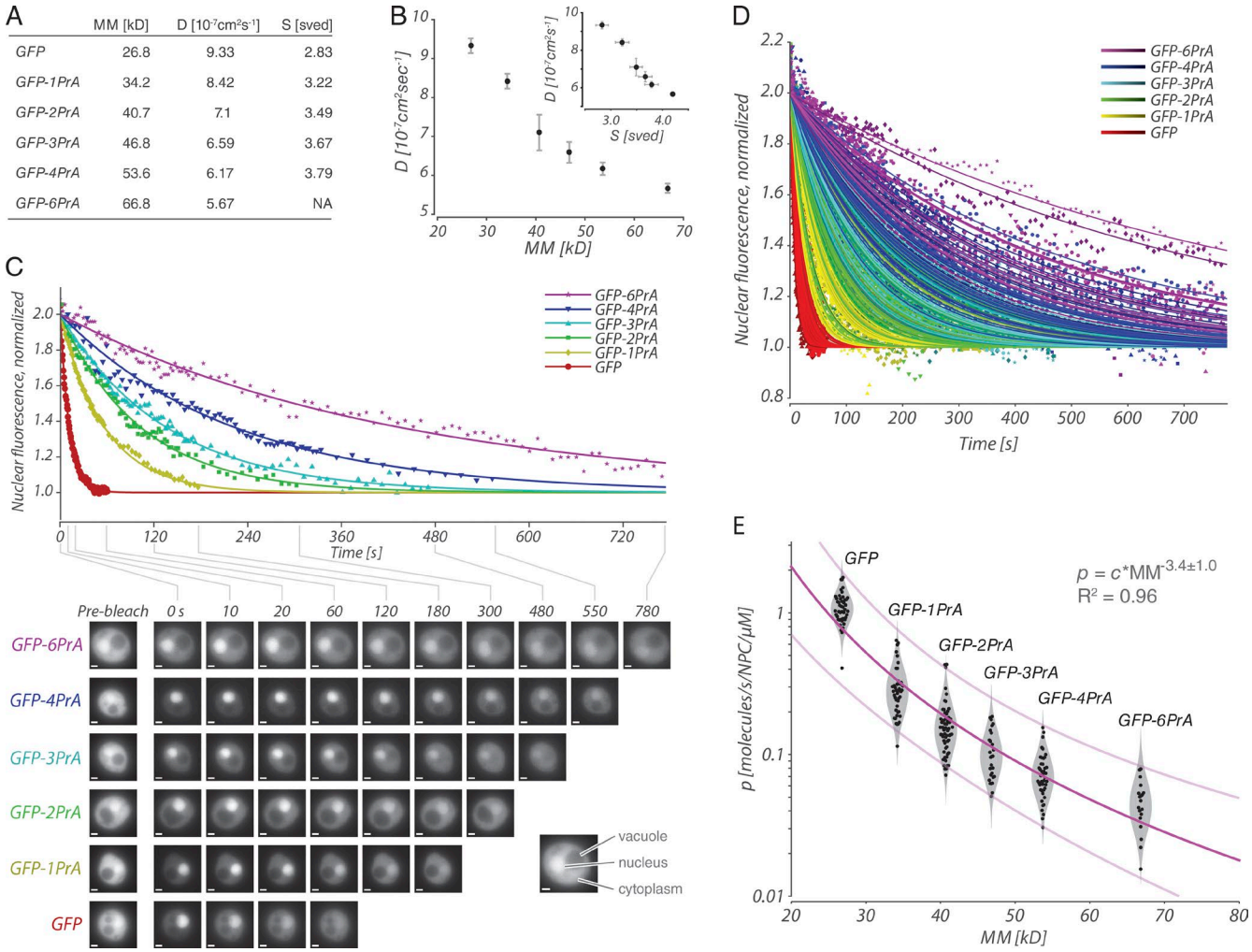


Figure 2. **The relationship between molecular size (and shape) and in vivo diffusion of free proteins.** (A) Passively diffusing reporter macromolecules used in this study, showing their molecular mass, in vitro diffusion constants (*D*) measured using DLS, and sedimentation coefficients (S) measured from sucrose gradient centrifugation. (B) Measured diffusion constant (*D*) in solution plotted against molecular mass and sedimentation coefficients (inset), ±SEM. (C) Images of representative single cell's FRAP experiments (bottom) of six different-sized substrates in WT yeast cells. Bars, 1 µm. The nucleus, cytoplasm, and vacuole of the top cell is indicated. Nuclear fluorescence is plotted after normalizing each exponential fit between 2 (at $t = 0$) and 1 (at $t = \infty$; top). (D) Diffusion curves from all such measurements of WT cells for the different-sized reporters, normalized as in C. (E) Population distributions of permeability coefficients for the different reporters as a function of their size. Each dot indicates the computed permeability coefficient for a single cell.

from 27 to 67 kD with the addition of multiple PrA units, with S_{max}/S values increasing from 1.15 to 1.35 (Fig. 3, C and D). However, this trend did not persist for the larger substrates up to 230 kD, which had much lower S_{max}/S values of 1.05–1.10, indicating an estimated 1.5:1 ratio between their long and short axes (Erickson, 2009) and a diameter of ∼10 nm for the 230-kD GFP-5MBP substrate. Hence, the substrates in the combined datasets have a wide variety of shapes, and the smallest and largest substrates are also the most globular. If those shape differences affected passive transport to a significant degree, we would expect that the measured sedimentation coefficient S, which is strongly affected by molecular shape, would be a better predictor of passive transport rates than S_{max}, the theoretical sedimentation coefficient of a perfect sphere of the same molecular mass (Erickson, 2009), which completely lacks shape information. However, S_{max} is in fact better correlated with the measured passive diffusion rates than S (Fig. 3 E; and see also Fig. S1 B regarding a similar analysis of the theoretical minimal radius, which also lacks shape information). Hence, the overall molecular mass of a substrate determines its passive diffusion rates, regardless of its particular shape geometry.

Compatibility of a soft barrier with previously proposed models of transport

Several mechanistic models of the transport system assume that large macromolecules are excluded from the NPC due to lack of space. The "aqueous channel" model suggests that passive transport is restricted to a relatively narrow aqueous channel with pore radius r_0, with estimates of r_0 ranging from 4.4 to 6.1 nm (Keminer and Peters, 1999). Under this assumption, passive diffusion rates are expected to depend on molecular shape more strongly than we observe in this study (Fig. 3 E). Moreover, although the molecules in the combined dataset have different shapes, several of the smallest and largest substrates in the full 27–230-kD range have an exceptionally low S_{max}/S ratio and can be treated as roughly spherical particles (Fig. 3, C and D). For such particles, the diffusion times through an aqueous pore would increase sharply beyond a well-defined size threshold (Renkin, 1954), which is

Published October 3, 2016

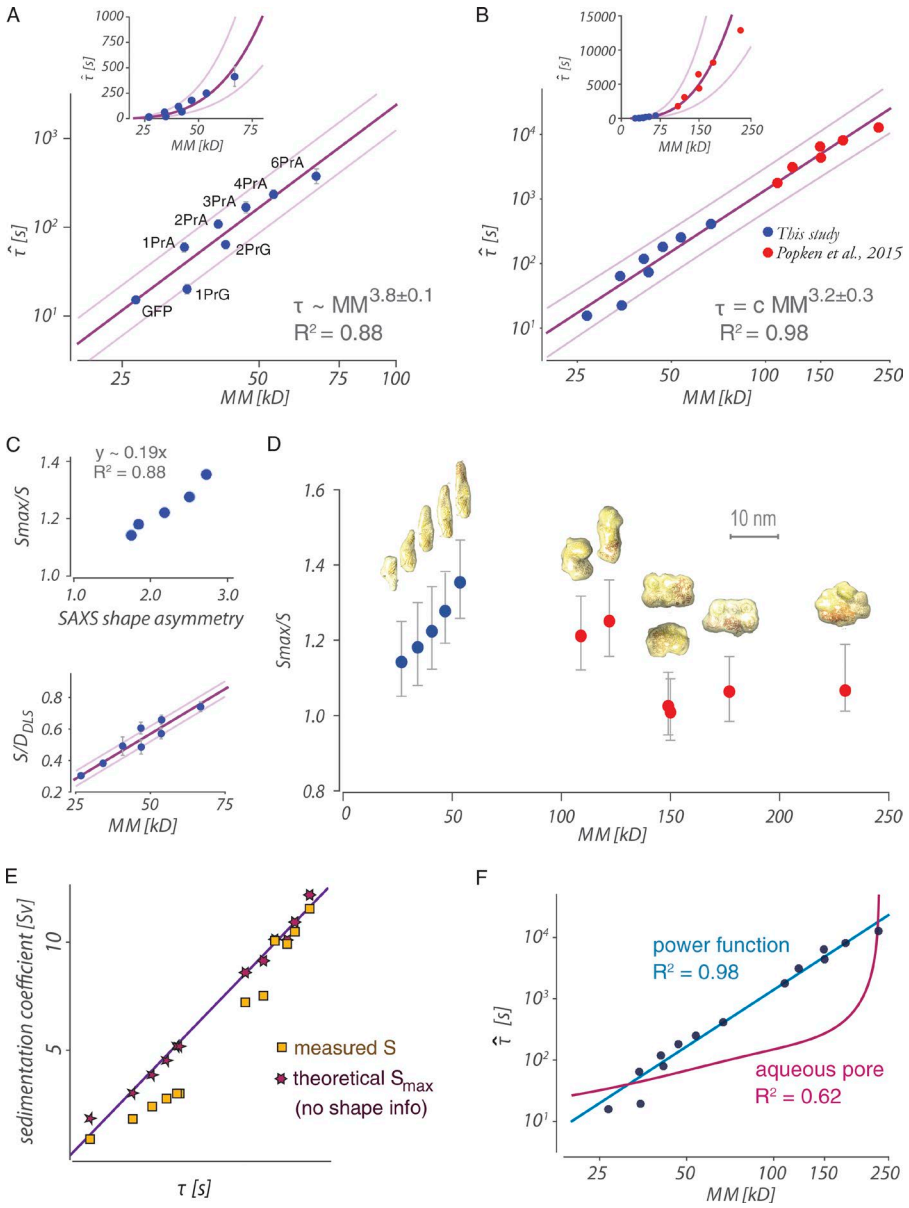


Figure 3. **Effect of molecular mass on the exponential time constant $\hat{\tau}$.** (A) $\hat{\tau}$ Values plotted as a function of the substrate molecular masses in log-log or linear scale (inset), ±95% confidence intervals. The fit to a power function is shown in a solid bold line; the thin gray lines indicate the power function at the 95% confidence intervals of the fit. (B) $\hat{\tau}$ Values as a function of the reporter molecular masses in log-log or linear scale (inset) for the joint set that includes our measurements as in A (blue) and the normalized $\hat{\tau}$ values that we computed for larger cargos, based on the reported ratio between nuclear and cytoplasmic fluorescence after incubation (Popken et al., 2015; red). The fit to a power function is indicated as in A. (C, top) The ratio between S_{max}, the sedimentation coefficient of a perfect sphere, and S, the measured sedimentation coefficient, plotted against the SAXS measured shape asymmetry. (bottom) The ratio between the measured sedimentation coefficient and the measured diffusion coefficient ± 95% confidence intervals, plotted against molecular mass. According to the theory of sedimentation, S/D ~ molecular mass (Erickson, 2009). (D) S_{max}/S ratios, which are correlated with shape asymmetry plotted for reporter molecules of different molecular masses for the joint dataset as in B and C ± 95% confidence intervals. Enclosing envelopes (yellow surface representation) computed from either SAXS measurements for each substrate in our dataset, or modeled based on the S_{max}/S values in the Popken et al. (2015) datasets, using the crystallographic models of each of their GFP, PrA, and MBP subunits (Materials and methods). (E) Theoretical S_{max} and measured S sedimentation coefficients plotted as a function of τ in log-log scale. The fits of S_{max} and S to a power function of τ are shown in solid magenta and orange lines, respectively. (F) Same as inset of B, with a best fit to a theoretical model of an aqueous pore (Renkin, 1954) shown in solid magenta line.

inconsistent with the data (Fig. 3 F; complete analysis in Materials and methods). The aqueous channel model is also inconsistent with microscopy studies in permeabilized cells (Ma et al., 2012). Similarly, a sievelike hydrogel with a certain characteristic pore size (Ribbeck and Görlich, 2001; Frey and Görlich, 2007; Mohr et al., 2009) or a collection of narrow pores (Mohr et al., 2009; Eibauer et al., 2015) are expected to result in a more well-defined size threshold and be more sensitive to substrate shape, which would affect the likelihood of permeating narrow pores. Other models of passive transport are expected to depend more strongly on molecular mass (i.e., molecular volume; Fischer et al., 2004) than on substrate geometry. Specifically, several models assume a "virtual gate" that depends on entropic exclusion rather than on a geometric constriction (Rout et al., 2000, 2003; Lim et al., 2006; Hough et al., 2015; Sakiyama et al., 2016; Zahn et al., 2016). In such an entropy-driven barrier, substrates with larger molecular mass or volume cause a strain on the system by restricting the conformational freedom of the disordered FG domains that are anchored to the NPC scaffold, resulting in a size-dependent entropic penalty on their passive diffusion.

Comparison to Brownian dynamics simulations of passive diffusion

To test whether our observations can be explained by an entropy-driven model of transport, we simulated the passive transport system using Brownian dynamics simulations (Ermak and McCammon, 1978), as described in detail in the Materials and methods. We modeled the major system components: the NPC, the flexible FG repeats that are anchored to its scaffold, the passively diffusing substrates, the cytoplasmic volume, and the nuclear volume (Fig. 4 A). The chains of FG repeat domains were modeled as flexible strings of beads with each bead representing a single repeat, based on increasing evidence indicating that FG domains are highly disordered, lack secondary structure, and have broad distributions of radii of gyration and end-to-end distances; this representation is consistent with many recent findings (Denning et al., 2003; Lim et al., 2006; Peleg et al., 2011; Hough et al., 2015; Milles et al., 2015; Lemke, 2016; Raveh et al., 2016; Sakiyama et al., 2016; Vovk et al., 2016; Zahn et al., 2016) and the classic entropic spring model of polymers (van der and Maarel, 2008). To

Published October 3, 2016

gather sufficient statistics about each substrate, we generated 20,000 independent trajectories on a Linux computer cluster (Materials and methods). Video 2 shows a sample simulation trajectory with substrate macromolecules of various molecular masses and their occasional translocation between the cytoplasm and the nucleus.

To directly compare our simulations with the empirical measurements of $\hat{\tau}$, we also computed the simulated $\hat{\tau}$ values for each of the different substrate macromolecules directly from the simulations (Fig. 4 B). We used the microscopic definition of τ as the time interval between consecutive translocations across the NE (Fig. S2, A and B) and normalized this value to account for differences in the cytoplasmic volume, the nuclear volume, and the number of NPCs, as with the empirical $\hat{\tau}$ values (Materials and methods). As expected of a soft and stochastic filtering mechanism, the interval times between consecutive translocations were widely distributed (Fig. S2 B). Similar to the empirical data, the mean translocation time increased gradually with molecular mass, and remarkably, it could also be approximated by a cubic function of molecular mass, closely resembling the relationship that we observed experimentally after scaling the simulated $\hat{\tau}$ by a uniform constant factor to account for modeling inaccuracies (Fig. 4 B). Significant changes to the density of FG domains in the NPC (Fig. 4 C), in the cohesiveness of FG domains (Fig. 4 D), or to the dimensions of the NPC itself (Fig. S2 D) had gradual effects on size-dependent filtering in our simulations, with a stronger effect over the baseline levels of passive diffusion rates of all substrates (i.e., the plot is vertically shifted along the y axis) but a much more gradual tuning of the relative size selectivity of different substrates (i.e., the slope of the fitted power functions in Fig. 4, C and D, does not change significantly when plotted in log-log scale). Hence, although cohesive interactions and precise optimizations of the density of FG domains and the pore size could serve to fine-tune the passive barrier, they are not required for recapitulating the fundamental properties of its function. We conclude that a soft barrier to passive transport is consistent with an entropy-driven model of the NPC.

Relating the kinetics and the free energy barrier to passive diffusion

From the spatial density distributions of the different macromolecules in our simulations, we computed their free-energy landscapes as a function of their position along the NPC (Fig. 4, E and F; Materials and methods). As expected, the magnitude of the free-energy barrier ($\Delta G_{passive}$) increased slowly (sublinearly) with molecular mass (Fig. 4 F). In agreement with transition state theory, the simulated $\hat{\delta}$ values are exponentially related to ΔG (Fig. S2 C), fitting a slightly modified form of the Eyring-Polanyi equation with $R^2 = 0.999$ (Eyring, 1935). Hence, our simulations suggest that the kinetics of passive diffusion may be modeled as an elementary chemical reaction with a single reaction step and a single transition state.

Measurements in mutant strains reveal disproportionate role for different FG domains

To investigate empirically how the composition of the FG domains that comprise the permeability barrier affects the NPC's permeability, we have used a comprehensive combinatorial FG domain deletion catalog (Strawn et al., 2004; Terry and Wente, 2007; Adams et al., 2015). We selected a subset of 12 mutant strains that includes truncated FG domains from all 11 yeast FG nups at least once, in various combinations (Fig. 5 A and Table S2). As a positive control for elevated permeability, we used the previously published *Δnup170* strain (Shulga et al., 2000). For practicality, we first narrowed our investigation to just the GFP-2PrA and GFP-4PrA substrates, which bracket the often-stated ∼40-kD diffusion limit. For these strains, single-cell FRAP measurements were combined, as before. These permeability measurements of 674 single-cell measurements are summarized in Fig. 5, B and C. Curiously, the permeability of most strains was not greatly increased with respect to either substrate. This includes even the "maximally deleted" mutants, $N^{\Delta FG}C^{\Delta FG}nsp1^{\Delta FG}$ and $N^{\Delta FG}C^{\Delta FG}nup100^{\Delta FG}$-$nup145^{\Delta FG}$, which each lack roughly half of the FG mass of the WT NPC (Fig. 5 B; Strawn et al., 2004). In sharp contrast, the $nup116^{\Delta FG}$ strain was 2.7–3.8-fold leakier than the WT strain, even though Nup116 is a minimally deleted mutant for which FG repeat domain contributes less than ∼10% of the NPC's FG mass. Similarly, truncation of the FG domain of Nup100 (a Nup116 paralog), either individually or in combination with other FG nups, resulted in 1.5–2.0-fold increase in permeability to at least one of the substrates (Fig. 5 C). All other strains except the $nup49^{\Delta FG}nup57^{\Delta FG}$ strain were only 1.0–1.5-fold leakier than the WT strain to either the GFP-2PrA or the GFP-4PrA substrates. Hence, the NPC passive permeability barrier is remarkably resilient to some extensive deletions of FG domains, whereas deletions of certain barrier-forming FG domains, such as those of Nup116 and Nup100 (and to a lesser extent, Nup49 and Nup57) seem to play a disproportionate role on passive permeability.

Deletion of barrier-forming FG domains has a uniform effect on substrates of different sizes

As barrier-forming FG domains are deleted from the NPC, we may expect either that passive permeability would increase more significantly for larger substrates than for smaller substrates or, alternatively, that the deletion of these domains may increase the diffusion rates of both small and large substrates uniformly. To test the effect of FG domain deletion on substrates of different sizes, we chose a subset of four mutant strains that covered truncated FG domains from nine of the FG nups, including Nup116 and Nup100. We collected single-cell FRAP measurements (Fig. 6 A) over the entire set of six labeled substrates (covering the molecular mass range 27–67 for passive diffusion) and for each of these strains. This exploration resulted in 29 substrate–strain combinations, comprising 792 individual cell measurements. Again, the leakiest strains for the full dataset of substrates had either the $nup116^{\Delta FG}$ or $nup100^{\Delta FG}$ truncations, whereas the basal permeability of the maximally deleted $N^{\Delta FG}C^{\Delta FG}nsp1^{\Delta FG}$ strain remained statistically indistinguishable from the WT strain (Fig. 6 B). At any rate, deletions had a uniform effect on both small and large substrates compared with the WT strain, regardless of the general level of leakiness; that is, the slope of the fitted power function for each of the strains is attenuated only slightly or not at all in Fig. 6 B (within the statistical confidence intervals), and these slopes describe the extent of size-dependent filtering across substrates of different molecular masses. Notably, this empirical result is again consistent with our simulation model, in which massive deletions of FG repeats had a

Published October 3, 2016

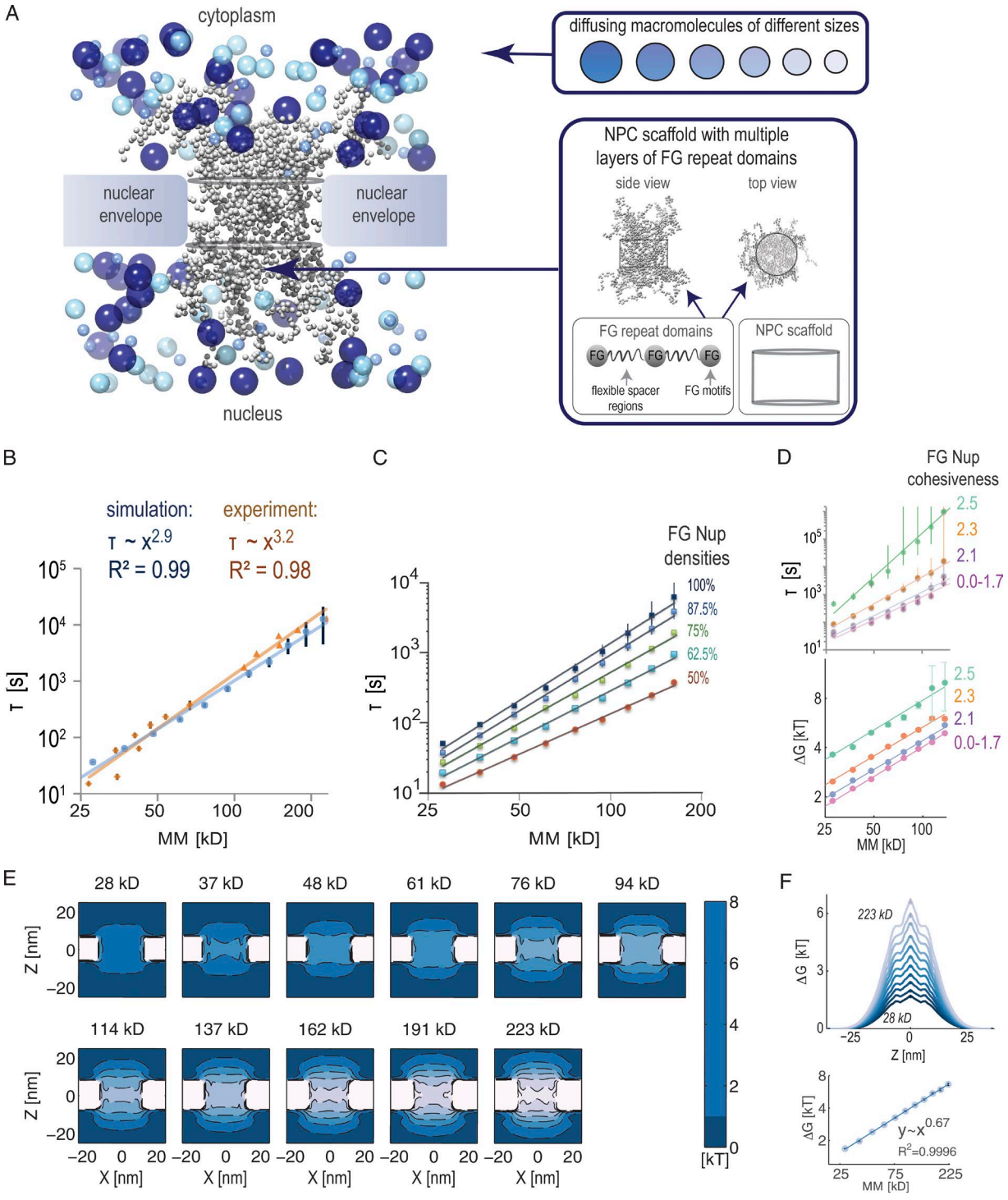


Figure 4. **Thermodynamics and kinetics analysis of simulated passive diffusion.** (A) Coarse-grained model of the passive diffusion through the NPC, with the NPC modeled as a cylindrical scaffold, with flexible FG repeats anchored to its walls. (B) Simulated and experimental exponential time constants of transport ± 95% confidence intervals, plotted as a function of molecular mass, with solid lines showing fits to power functions. The absolute values of the simulated time constants are scaled by a single constant value, to allow direct comparison with the experimental data, but without affecting the slope or the relative magnitude of the data points in any way. (C) As in B for simulated NPCs depleted in FG chains from 100% (top) to 50% (bottom). (D, top) As in B and C for FG domains with increasing cohesiveness. The legend indicates the interaction force k between pairs of FG motifs in units of kT/Å over a range of 6 Å. Results did not change significantly for k values in the range 0.0–1.7 kT/Å. (bottom) Energy for the same molecules and cohesiveness level as in the top panel; see also F. (E) The free energy barrier for diffusing macromolecules of different molecular masses in units of kT (color bar), along a cross section of the NPC going parallel to the NE through the NPC central axis (x axis) and from the nucleus to the cytoplasm (Z axis). (F) The free energy barrier projected along the z axis, by integrating over a tube of radius 12 nm about the NPC central axis. The top panel shows the free energy barrier as a function of Z for each substrate; the bottom panel shows the magnitude of the barrier as a function of molecular mass ± 95% confidence intervals.

Published October 3, 2016

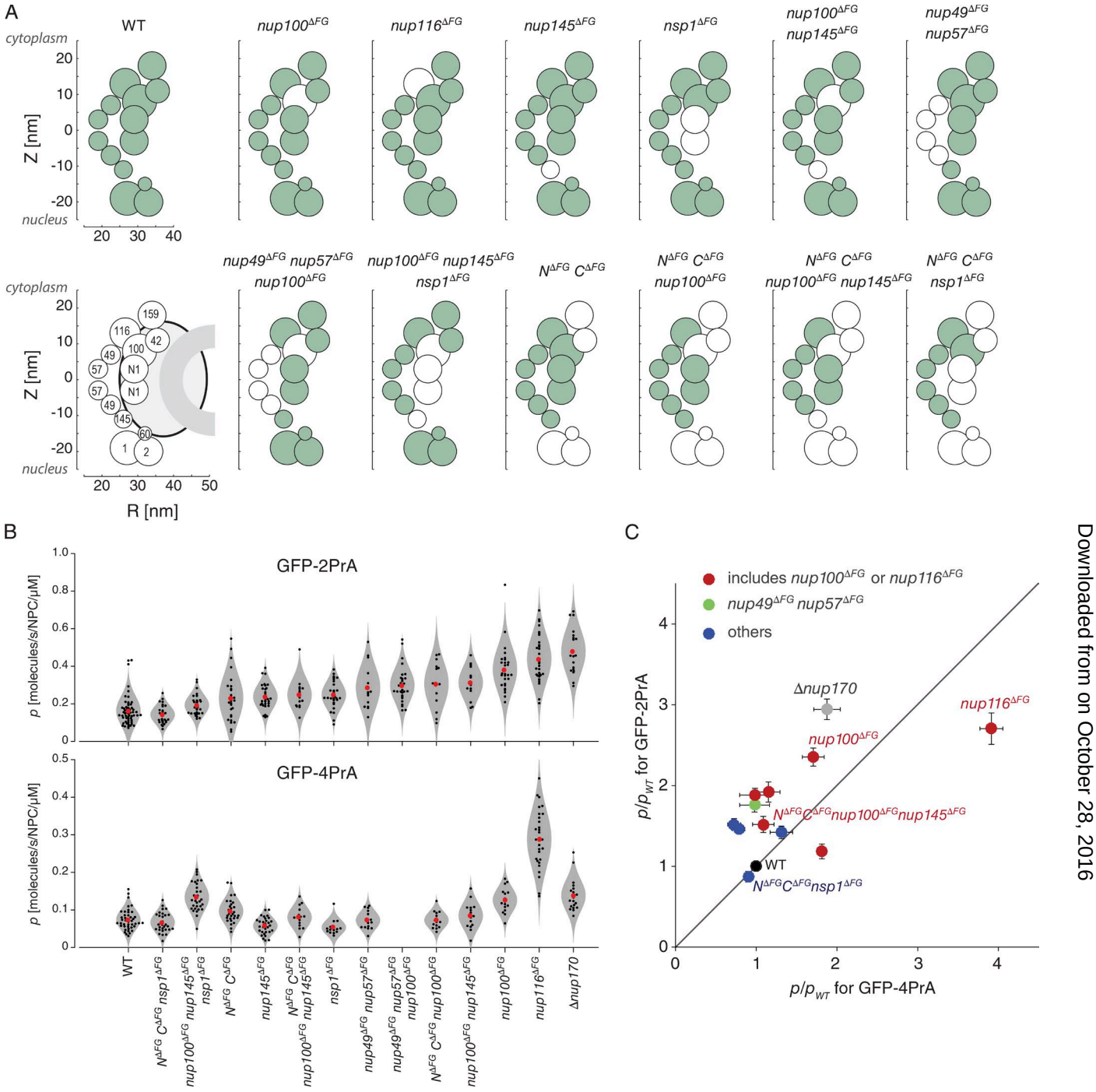


Figure 5. **Permeability of mutant strains.** (A) Combinatorially deleted strains used in this study; deleted nups in each strain are indicated by white circles. Bottom left illustrates that each diagram represents the right half of a cross section of a diagrammatic NPC, where *R* is the radial distance from the pore radial axis, and the z axis denotes vertical proximity to the cytoplasm (positive values) or the nucleus (negative values). These approximate FG nup positions are drawn according to the localization map for the FG nup C termini used in Table S2 of reference (Alber et al., 2007a). FG nup diameters are estimated for visualization purposes using Eq. 6 in reference (Marsh and Forman-Kay, 2010). (B) Measured NPC permeability coefficients (*p*) in different strains for GFP-2PrA (molecular mass 40.7 kD) and GFP-4PrA (53.6 kD). Each black dot is from one single-cell FRAP measurement. Red dots are the mean for the GFP-2PrA or GFP-4PrA substrate in that strain. Δ*nup170* was measured as positive control for a previously published "leaky strain" (Shulga et al., 2000). (C) Permeability of GFP-2PrA versus GFP-4PrA in the different strains relative to the permeability of the WT strain (±SEM). Strains lacking Nup100 or Nup116 GLFG domains are marked red, the *nup49*$^{\Delta FG}$*nup57*$^{\Delta FG}$ strain green, WT black, Δ*nup170* gray, and all others blue.

relatively mild effect on the size-dependent filtering and a stronger effect on their overall basal permeability (compare with changes in slope and in the baseline of the transport times in Fig. 5 F). This observation provides additional support for the model of entropic exclusion from the NPC and suggests that the density of certain FG repeats can be used to fine-tune the passive barrier. This result enables us to more formally define the basal permeability of each mutant strain as a strain-specific multiplier p_{strain}/p_{WT} relative to the WT, a quantity that applies uniformly to different substrates.

Published October 3, 2016

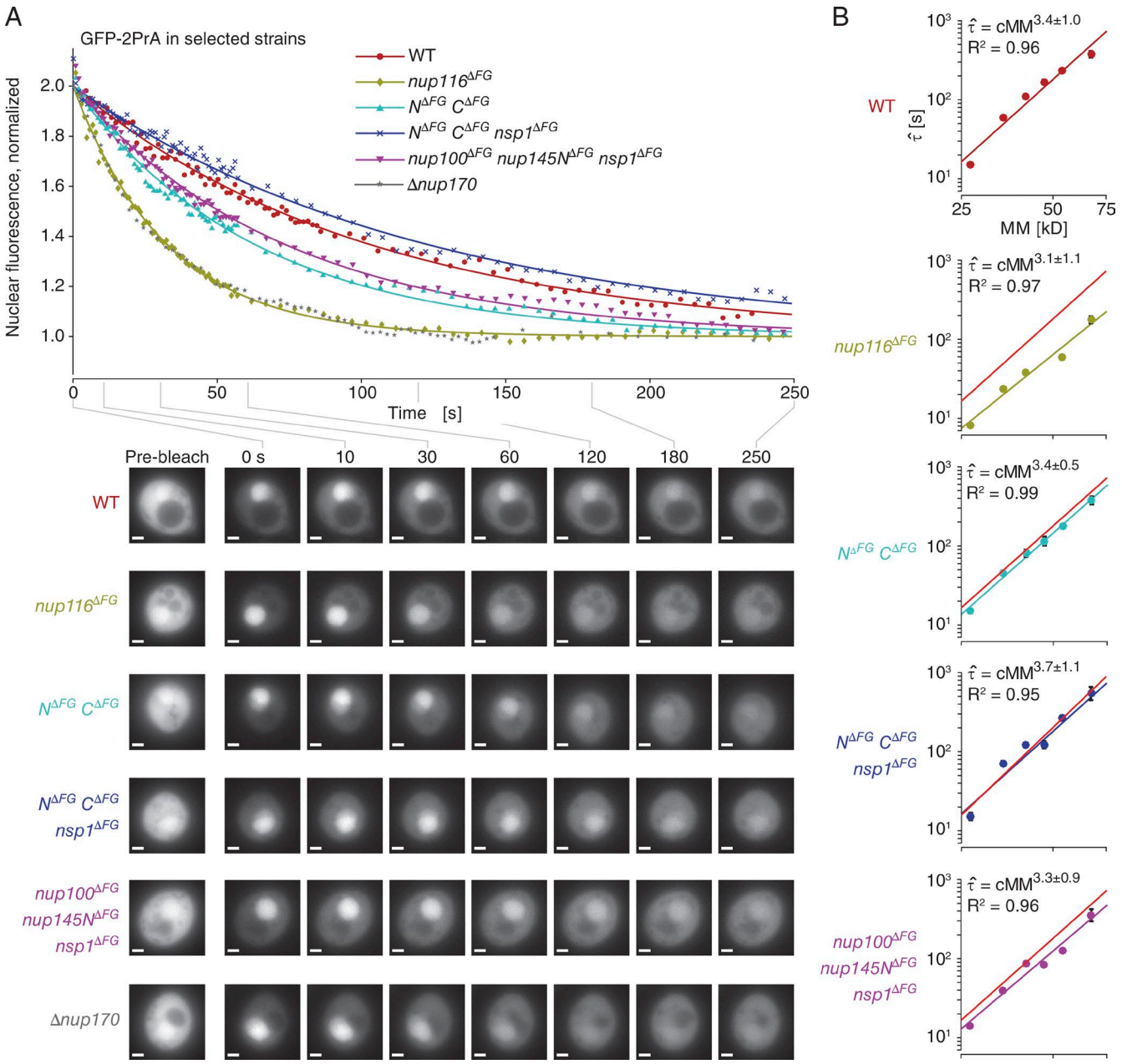


Figure 6. **Size-dependent permeability in four FG deletion strains.** (A) An example of NPC permeability assays from representative single cells expressing GFP-2PrA, from FG deletion and control strains. Bars, 1 µm. The permeability of Δnup170 was measured as positive control for a previously published leaky strain (Shulga et al., 2000). (B) Strains in which mean transport times ($\hat{\tau}$) were calculated for all six different-sized substrates are plotted ± 95% confidence intervals as a function of the reporter molecular masses in log-log scale, these being 29 (GFP), 34 (GFP-1PrA), 41 (GFP-2PrA), 47 (GFP-3PrA), 54 (GFP-4PrA), and 67 kD (GFP-6PrA). As GFP-xPrG data were only collected for WT cells, only GFP-xPrA data are included in this figure, for more direct comparisons. The solid line in each subplot indicates the fit to a power function. For the mutant strains, the WT fit is superimposed as a red line. The functional form of each power function is indicated along with the 95% confidence intervals for the power exponent.

Changes in passive permeability cannot be explained by indirect effects on the transport system

Many factors in the cellular biology of the mutant strains might have influenced our measurements of passive permeability. We first assessed whether passive permeability might be influenced by changes in the following general properties of the transport system: (1) NPC quantities; (2) inefficient targeting of truncated FG nups to the NPC; (3) increased pools of transport receptors at the NPC; and (4) changes in the rates of facilitated transport. We selected key strains that covered a range of phenotypes and FG flavors, in which we tagged key transport receptors and FG nups, to survey for global changes to the quantity or distribution of the machinery of nucleocytoplasmic transport. Namely, these strains were WT, the leaky *nup116*$^{\Delta FG}$ strain, and the massively deleted, nonleaky *nsp1*$^{\Delta FG}$ and $N^{\Delta FG}C^{\Delta FG}$*nsp1*$^{\Delta FG}$ strains. In these strains, we tagged the genomic copy of the transport receptors Kap95-mCherry, Crm1-GFP, Kap123-GFP, and Ntf2-GFP or the nups Nup116-GFP, Ndc1-GFP, and Nup170-GFP and acquired images of the steady-state distributions of each, in all four strains (Fig. S3, bottom). Using assisted automated image analysis, measurements of nuclear, cytoplasmic, and

Published October 3, 2016

(where possible) nuclear rim fluorescence of 9,621 such cells were obtained. The mean magnitude of change in each measurement, relative to WT, is summarized in Fig. S3. Importantly, we found that most strains changed minimally in their NPC abundance and that therefore these changes did not substantially contribute to our permeability measurements. We also found that Nup116-GFP targets efficiently to the NPC, even when truncated, suggesting that there are not overwhelming changes in NPC stoichiometry in most FG deletion strains. Finally, we found no substantial changes in quantities of fluorescent transport receptors at the nuclear rim in any strain (other than for Ntf2), and by surveying the most abundant transport receptors, we have sampled a significant fraction of facilitated transport.

How facilitated transport has been perturbed in these strains is a matter of particular interest, at least insofar as we can examine in this study, as well as how it contrasts with effects on passive transport. We did observe some changes in the nucleus/cytoplasm (N/C) distributions of soluble transport receptors, illustrating that the cycles of facilitated transport are sensitive to the loss of particular FG domains more than others. Inferring the meaning of some of these changes is beyond the scope of this study, considering the complexity of the system. Nonetheless, it is perhaps unsurprising that the distribution of Ntf2-GFP would be affected by the loss of Nsp1 FG domains, as the interaction between and importance of these two proteins is well characterized (Bayliss et al., 1999, 2002). Interestingly, the strain that showed the most notable effect on facilitated transport of the fluorescent transport receptors was the maximally deleted $N^{\Delta FG}C^{\Delta FG}nsp1^{\Delta FG}$ strain, in spite of having little effect on passive transport. Though technically more challenging (because of far faster equilibration rates), we therefore examined the dynamic rate of Kap123-GFP exchange between nucleus and cytoplasm, using FRAP, similar to the method we used to assay passive transport. We found that nucleocytoplasmic exchange of Kap123 decreased by 72% in $N^{\Delta FG}C^{\Delta FG}nsp1^{\Delta FG}$ (Figs. S3 and S4). In strong contrast to the permeability results, exchange of Kap123 decreased only modestly in $nup116^{\Delta FG}$ and $nsp1^{\Delta FG}$ and had little effect on nuclear rim fluorescence of transport receptors in these strains, even though $nup116^{\Delta FG}$ has the largest permeability to passive transport (compare Fig. 6 B). This discrepancy may be a result of loss of docking sites for Kap123, consistent with decreased rim intensity for Kap123 in that strain (Fig. S3) and with recent kinetic models that suggested that facilitated diffusion is rate-limited by the availability of docking sites rather than by transport efficiency (Kim and Elbaum, 2013b). Overall, these results show that different FG nups have a very different effect on passive and facilitated transport, indicating distinct functions for different FG nups in passive versus facilitated transport.

Direct effects on passive permeability as a result of FG nups positions, charge, and hydropathy

Because changes in passive permeability among strains (Figs. 5 and 6) could not be explained by indirect effects such as changes in NPC quantities or the rates of facilitated transport (Figs. S3 and S4), we set out to identify the biochemical and geometrical properties of specific FG domains that might directly affect passive permeability. For each strain, we computed its basal permeability to passive transport relative to the WT strain (p_{strain}/p_{WT}), averaging over all substrates for which measurements were taken. We then collated an exhaustive list of 33 FG domain features that may affect passive permeability (Table S3). We included such features as the total mass and number of deleted FG domains; sequence features that describe the flavor of deleted FG domains (GLFG/FXFG/other; i.e., their amino acid composition, hydrophobicity levels, fraction of charged residues, and net charge); and geometrical features that describe the approximate spatial localization of different FG nups relative to the cytoplasmic and nuclear end of the NPC, despite progress in mapping the vertebrate NPC (Eibauer et al., 2015; Kosinski et al., 2016), we cannot use these maps because of differences in scale and structure with the yeast NPC (Yang et al., 1998), and so instead, we used our previously reported map, which is at a level of resolution sufficient for the purposes of the current study (Alber et al., 2007b).

From the entire list of 33 features, we were unable to identify a single feature that correlated well with the observed variance in basal permeability among the different mutant strains (Fig. 7 A and Fig. S5). For example, consistent with previous studies (Strawn et al., 2004; Terry and Wente, 2007; Popken et al., 2015), the total FG mass deleted from the NPC did not correlate with the change in passive transport rates (Fig. 7 A). To identify combinations of features that might better explain the observed interstrain variability in basal permeability while minimizing the risk of data overfitting, we used the Least Absolute Shrinkage and Selection Operator (LASSO) method for multivariate statistical regression (Tibshirani, 1996), combined with principal component analysis (PCA; Jolliffe, 1982). To offset concerns that the regression analysis may produce a model that overfits the specific data used for training the model, we repeated the regression analysis over a smaller dataset, which included only measurements with the GFP-2PrA and GFP-4PrA constructs. Reassuringly, similar features and weights were selected by the model from the smaller dataset (Fig. S6 A), or when we removed either of the strains with the greatest effect on basal permeability, $nup116^{\Delta FG}$ and $nup100^{\Delta FG}$, from the dataset (Fig. S6 B), or lastly, to account for potential uncertainty in the copy number of certain FG nups (Mi et al., 2015), when we assumed a uniform stoichiometry of FG nups at the NPC (Fig. S6 C).

The combined PCA/LASSO model explains the observed variation in basal transport rates between strains with an R^2 value of 0.90 (compare Fig. 7 B with the fits of individual features in Fig. S5). Importantly, this statistical regression model has a straightforward and intuitive interpretation, pointing at two groups of features of FG Nups that in combination have the greatest effect on passive transport (Fig. 7 C). The first group consists of positional features of our medium-resolution model of the NPC (Alber et al., 2007b). The second group of features is related to the ratio of net charge to hydropathy score in the FG domains that were deleted from the NPC. Specifically, we conclude from our model that FG domains that are both anchored close to the cytoplasmic end of the NPC (Fig. S7 A) and have a high ratio of net charge to hydropathy (Fig. S7 B) contribute most to the passive barrier, most notably Nup100, Nup116, and possibly the adjacent Nup42, Nup49, and Nup57 (Fig. 7 D).

Discussion

NPCs facilitate the retention of nuclear and cytoplasmic factors in their appropriate subcellular compartments while tightly regulating their timely exchange. However, the precise relation

Published October 3, 2016

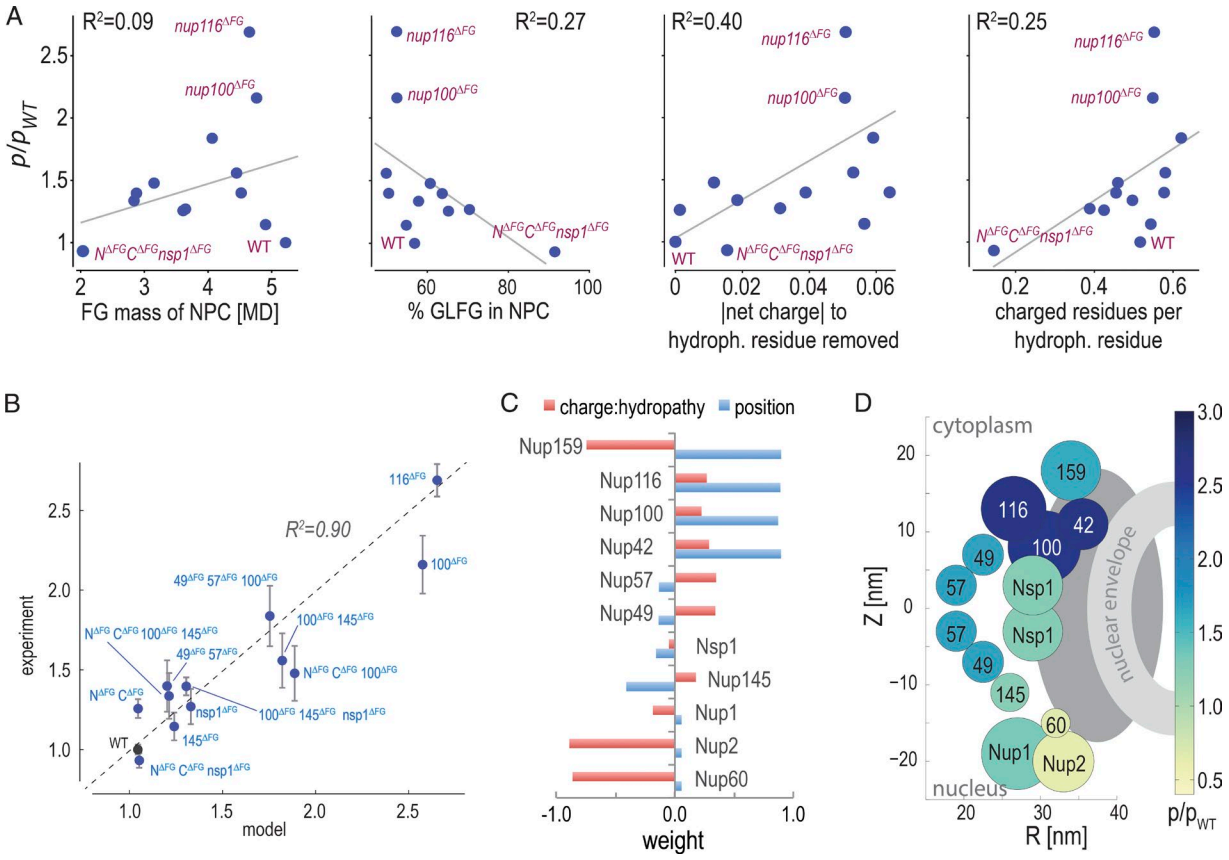


Figure 7. **No single factor is a good predictor of NPC permeability.** (A) Effect of select features on pore basal permeability (p/p_{WT}) relative to the WT strain. Shown are the four features with the highest absolute correlation. Plots for all 33 features are shown in Fig. S5. (B) Predicted basal permeability (p/p_{WT}) of combinatorial deletion strains (±SEM) relative to the WT strain according to the LASSO statistical regression model versus the experimental measurements for each strain. (C) Aggregate contribution of features related to either the ratio of net charge to hydropathy or position to the LASSO regression model. (D) p/p_{WT} of individual nups relative to the WT strain according to the regression model. The nups are positioned based on their proximity to the cytoplasm (z axis) and their distance from the NPC radial axis, as observed in immuno-EM maps (Alber et al., 2007b). The prediction for Nup159 should be assigned lower confidence because of the unusual negative charge and architecture of Nup159.

between the size and shape of macromolecules and their passive translocation rates, as well as the role of different FG nups in establishing the barrier, is still poorly understood. In this study, we provide the most rigorous quantitative characterization of in vivo size-dependent passive diffusion rates to date, based on measurements in both WT and mutant yeast strains (Figs. 1, 2, and 3). Our results clearly support the model of a soft barrier to passive diffusion (Fig. 1 A) and refute the prevailing paradigm of a rather firm barrier at any well-defined size threshold, including the oft-cited 30–60-kD threshold (Keminer and Peters, 1999; Mohr et al., 2009; Ma et al., 2012; Christie et al., 2016; Knockenhauer and Schwartz, 2016; Musser and Grünwald, 2016; Schmidt and Görlich, 2016). Our results provide mechanistic context to recent large-scale screens of nuclear compartmentalization, which indicated that relatively large proteins, larger than the 30–60-kD cutoff, are often equally partitioned between the nucleus and the cytoplasm, similar to smaller proteins (Wühr et al., 2015). Alternatively, such proteins are retained in just one compartment by either binding to even larger molecular complexes, for which passive diffusion is expected to be negligible even if we assume a soft barrier (Wühr et al., 2015) or are actively removed from the other compartment by facilitated diffusion (Kirli et al., 2015). These results indicate that the slow diffusion of much larger substrates across the NPC (Wang and Brattain, 2007; Popken et al., 2015) cannot be interpreted as residual leakage beyond a well-defined cutoff (Kirli et al., 2015) but rather as an intrinsic property of the NPC filtering mechanism.

Suitability of in vivo measurements and labeled probes

Our measurements were performed in live yeast cells, for which we checked for changes in NPC quantity, composition, and potential clogging with soluble transport receptors (Figs. S3 and S4). The measured permeability coefficient of the GFP substrate was 1.1 ± 0.04 molecules/s/NPC/μM (Fig. 2 E), which is comparable to results from in vitro studies of permeabilized HeLa cells (Ribbeck and Görlich, 2001; Mohr et al., 2009; Ma et al., 2012) and NPCs isolated from *Xenopus* oocytes (Kiskin et al., 2003). However, we prefer the use of in vivo measurements to quantify the operating parameters of the nucleocytoplasmic system, considering that a range of variables in in vitro systems, such as the lack of competition with other transport receptors (Timney et al., 2006; Zilman et al., 2010), RanGTP (Ma et al., 2012, 2016; Lowe et al., 2015), and nonspecific interactions with the cellular milieu (Tetenbaum-Novatt et al., 2012; Hough et al., 2015), were clearly demonstrated to affect the biophysical state of NPCs in general and passive diffusion in particular. These variables have different effects on small and large macromolecules (Lowe et al., 2015) and are therefore likely to

Published October 3, 2016

introduce a significant bias for characterizing the size threshold for passive diffusion in vitro (Keminer and Peters, 1999; Mohr et al., 2009). The labeled GFP-xPrA substrates are appropriate model substrates for passive diffusion in vivo because PrA and GFP are both understood to be functionally inert in yeast, lack nuclear localization and export signals, and do not interact with native yeast proteins (Gavin et al., 2002; Keilhauer et al., 2015). Moreover, they are more representative of cellular proteins than dextran molecules (Keminer and Peters, 1999), because FG repeats are likely to form weak nonspecific interactions even with mixtures of inert proteins like GFP and PrA (Tetenbaum-Novatt et al., 2012; Hough et al., 2015; Raveh et al., 2016).

Substrate determinants of passive diffusion

Our results indicate that passive diffusion rates depend most strongly on the substrate molecular mass, that is, its volume (Fischer et al., 2004; Fig. 3, A and B), and not so much on molecular shape asymmetry (Fig. 3 E). Various other substrate features have been previously proposed to affect passive diffusion (Naim et al., 2009; Colwell et al., 2010; Tagliazucchi et al., 2013; Goryaynov and Yang, 2014). Indeed, substrates of similar size and shape in our dataset did show relatively small but statistically significant variation in their passive transport rates (compare GFP-xPrA and GFP-xPrG substrates in Fig. 3 A and Table S3). Nonetheless, except in relatively extreme cases (such as the incorporation of many surface hydrophobic residues into diffusion substrates [Naim et al., 2009]), our data suggest that molecular volume is the most prominent determinant of passive diffusion.

Simple biophysical principles underlying size-selective passive diffusion

Several of the proposed mechanistic models of transport posit that passively diffusing substrates are restricted by a well-defined geometric barrier, such as a narrow aqueous pore (Keminer and Peters, 1999) or a sievelike saturated hydrogel (Ribbeck and Görlich, 2001; Frey and Görlich, 2007; Mohr et al., 2009). Therefore, passive diffusion is expected to be sensitive to substrate geometry and plummet at a well-defined size threshold as the substrate reaches the characteristic pore size, which is inconsistent with our data (Fig. 3, E and F). In contrast, if the barrier to passive transport is formed because of entropic exclusion (Rout et al., 2000, 2003; Lim et al., 2006; Sakiyama et al., 2016), the NPC is expected to sense the excluded volume occupied by the substrate, which is a simple function of molecular mass/molecular volume (Fischer et al., 2004) rather than its specific geometry. If passive diffusion depends on overall excluded volume, it is also expected to decay more gradually compared with a geometric barrier. Our simulation results confirm that such a soft virtual gate to passive transport is formed by the anchoring of flexible polymers to a cylindrical pore (Rout et al., 2000, 2003; Fig. 4 and Video 2), in agreement with the overall size-dependent filtering properties of the NPC observed in living cells (Fig. 4 B). Our experimental measurements of passive diffusion in mutant strains are remarkably consistent with our simulation results, showing that the deletion of FG domains may increase the overall leakiness of a given strain, but it affects substrates of different sizes in a similar way, thereby maintaining the relative size selectivity of the NPC (Fig. 6 B). Simply put, the anchored FG domains form a highly crowded environment in the NPC, but because of their disordered and dynamic nature, this environment, the NPC's central channel, is still permeable to macromolecules. For these macromolecules to pass through the NPC, they must displace these domains, competing for the restricted available volume and restraining the FG domains' conformations. This generates entropic repulsive forces shaping a free energy barrier to passive diffusion (Fig. 4, E and F); because these forces are caused by displacement, they therefore scale with the volume of the diffusing macromolecules.

Simulations of passive diffusion are robust to changes in the model details

The NPC is a remarkably resilient machine that can withstand extensive deletions and maintain a similar function in a range of divergent eukaryotes. Although the NPCs of different organisms vary in their spatial architecture and in FG nup composition, we incorporated within our model only the basic properties that are sufficient to recapitulate the in vivo properties of the NPC, as observed in all eukaryotes. To avoid bias from uncertain model details, we have attempted to minimize the complexity of our model. Our current model does not attempt to include a high-resolution physical representation of the NPC, but it includes all the major system components that are required to form an entropically driven gate to passive diffusion—namely, a cylindrical pore filled with highly flexible polymers. To verify that this simplification did not skew our results, we tested the resilience of our model over a wide range of model parameters and showed that such changes, including substantial changes to the density of the FG domains (Fig. 4 C), their cohesiveness (Fig. 4 D), and the dimensions of the NPC (Fig. S2 D) had a limited effect on our results. Similarly, competition with transport receptors or the functional arrangement of different types of FG repeats may contribute to modulating this core function (Vovk et al., 2016; Zahn et al., 2016) or contribute to other functions such as facilitated diffusion, but are secondary to this entropic effect in forming the passive barrier itself.

Other approaches have been used to simulate the transport system using a generally higher level of model complexity (Mincer and Simon, 2011; Moussavi-Baygi et al., 2011; Ghavami et al., 2016) than our present simulations. In these studies, passive diffusion rates have been quantified only for a small number of relevant substrate sizes (Mincer and Simon, 2011; Moussavi-Baygi et al., 2011) or they have focused on the system thermodynamics rather than kinetics (Ghavami et al., 2016), prohibiting direct comparisons with our kinetic experimental and computational data. Nonetheless, the free energy barrier to passive diffusion inferred from our simulations increases to the power of 0.67 of molecular mass (Fig. 4 D), which is consistent with the latter thermodynamic study, in spite of different model details (Ghavami et al., 2014, 2016). This robustness to changes in the model details is consistent both with the robustness of the NPC to substantial genetic deletions (Strawn et al., 2004; Terry and Wente, 2007) and with how biomimetic nanopores containing end-grafted FG domains recapitulate major properties of the transport system (Jovanovic-Talisman et al., 2009).

Properties of functional region for passive permeability

Our results indicate that the passive barrier may be mainly formed by a subset of FG nups (Fig. 7 D)—Nup116, Nup100, Nup42, Nup49, and Nup57 (Fig. 7 D)—whereas facilitated transport may be more affected by deletions of another subset

Published October 3, 2016

of FG nups (Figs. S3 and S4). These results strongly support the concept that the NPC has a higher-level arrangement of particular FG repeats with distinct functionalities distributed by a complex scaffold (Rout and Wente, 1994; Alber et al., 2007b; Patel et al., 2007; Yamada et al., 2010; Tu et al., 2013; Popken et al., 2015). Indeed, truncating both Nup116 and Nup100 together is lethal in yeast (Iovine and Wente, 1997), and these two Nups were suggested to contribute to the passive barrier in the past (Hülsmann et al., 2012; Popken et al., 2015). It has been recently suggested, based on simulation studies, that Nup116, Nup100, Nup42, Nup49, and Nup57 participate in forming a dense region of FG domains within the central channel (Ghavami et al., 2014) and that the truncation of their FG domains may affect the passive transport by disrupting this region, but not the spatial organization and the functions of other FG nups that are projected out of the pore in these simulations (Popken et al., 2015). By applying statistical regression methods to our results, we concluded that the FG nups contributing to passive permeability share two distinct properties: they are colocalized at the cytoplasmic end of the NPC central channel, and they also all exhibit a high ratio of net charge to hydropathy score (Figs. 7, S6, and S7). Interestingly, all of these FG nups are also canonically GLFG nups, which were previously suggested to be functionally distinct from other FG nups (Yamada et al., 2010). However, our results indicate that their position near the cytoplasmic end of the central channel also seems obligatory, as Nup145N, which also carries many GLFG motifs, has little contribution to the passive permeability barrier in our measurements. We note that despite this colocalization at the cytoplasmic end, the effect on entry and exit rates from the nucleus seems to be identical, as the equilibrium ratio between nuclear and cytoplasmic fluorescence levels is insignificant in leaky strains (Fig. S8).

GLFG domains are depleted of charged residues (Patel et al., 2007; Yamada et al., 2010; Adams et al., 2015; Schmidt and Görlich, 2015), yet their overall net charge is in fact strongly positive and their hydropathy score low compared with other FG repeat flavors (i.e., they are overall exceptionally hydrophilic; Fig. S7, B and C). Moreover, although all FG repeats appear to be disordered (Denning et al., 2003; Yamada et al., 2010; Hough et al., 2015), high net charge and low hydropathy are hallmarks of extreme protein disorder (Tcherkasskaya et al., 2003; Dunker et al., 2008; Fig. S4 B). This is surprising because GLFG domains are assumed to be rather sticky and compact compared with other FG domains (Patel et al., 2007; Yamada et al., 2010; Schmidt and Görlich, 2015). It has also been suggested that the high net charge of FG nups may contribute to interaction with the mostly acidic transport receptors (Tagliazucchi et al., 2013), perhaps helping to inhibit passive diffusion through competition. Therefore, the role of highly charged regions at the cytoplasmic end of the NPC in forming the passive barrier needs to be further characterized in future studies; for instance, by mutating charged residues or modulating the hydrophobicity levels of these domains and by nuclear magnetic resonance spectroscopy or SAXS.

Implications of soft barrier on facilitated diffusion

In contrast to the relatively straightforward Eyring-Polanyi relation shown here between $\hat{\delta}$ and $\Delta G_{passive}$ (Eyring, 1935; Fig. S2 C), accurately relating the energy landscape for diffusing through the NPC to transport kinetics may require more elaborate approaches that account for the increased interaction time of NTRs within the NPC, competition effects, availability of docking sites and differences between import and export (Kopito and Elbaum, 2007; Zilman et al., 2007; Kim and Elbaum, 2013a; Tu et al., 2013). Facilitated diffusion also depends on unique properties of the interactions between NTRs and FG Nups, which were suggested to exchange rapidly (Hough et al., 2015; Milles et al., 2015) through unique interaction mechanisms (Raveh et al., 2016). Nonetheless, if we assume that passive and facilitated diffusion proceed through the same physical conduit, we expect the free energy gain from interactions between NTRs and FG Nups ($\Delta G_{interaction}$) to roughly counter $\Delta G_{passive}$. Because the latter increases only slowly (sublinearly) with molecular mass (Fig. 4 F; see also Ghavami et al., 2016), a relatively small number of interactions are expected to facilitate the fast diffusion of relatively large cargo molecules. For example, according to the approximately cubic relation between molecular mass and $\hat{\delta}$ (Fig. 3 B), we estimate the mean timespan between consecutive passive translocations of 4-MD mRNP complexes (Grünwald and Singer, 2010) to be 10^7 times slower than for a 27-kD GFP molecule, an energy difference of 12–16 kT according to the Eyring-Polanyi relation (Fig. S2 C; Eyring, 1935). A barrier of this magnitude is not insignificant, but it could be readily overcome by the binding energy from a relatively small number of NTRs. Indeed, 4- to 5-MD mRNP complexes translocate through the NPC in merely 200 ms (Grünwald and Singer, 2010; Lowe et al., 2015), and the facilitated diffusion rates of ∼10-MD complexes of quantum dots and NTRs crossed the NPC within a few seconds when bound to 20 NTR molecules and in less than a second when bound to 40 NTR molecules, respectively (Lowe et al., 2010). These data are consistent with the gradual overcoming of a surmountable energy barrier, as expected from both our empirical and computational results. At the other end of the size spectrum, the rather gradual decrease in the energy barrier with molecular mass also explains why specialized NTRs are used even for some small macromolecules that are transported at high quantities, such as the 25-kD Ran (Smith et al., 1998; Ribbeck and Görlich, 2001), because the energy penalty for passive diffusion scales with the total volume of passive diffusers according to our model.

Understanding how the NPC forms such a soft barrier while facilitating active transport over this wide range of molecular sizes will be advanced significantly by the ongoing structural characterization of the NPC and its FG domains, particularly those within critical locations of the NPC. Additional measurements and simulations that use such refined data will illuminate the functional arrangement of the NPC that enables the cell to maintain the proper flow of nucleocytoplasmic traffic.

Materials and methods

Plasmids

Plasmids were constructed by standard molecular biological methods. DNA constructs were made for expression of eGFP C-terminally tagged with one, two, three, four, or six repeats of the PrA domain from *Staphylococcus aureus*. These GFP-xPrA genes were subcloned either into the yeast constitutive expression plasmid pYX242 (EMD Millipore, discontinued) or as 6xHIS fusion genes into the bacterial expression plasmid pET28a (EMD Millipore). GFP-1PrG and -2PrG constructs were made similarly, as yeast and bacterial expression constructs, by fusing to GFP either the C2 domain (for GFP-1PrG) or both

Published October 3, 2016

the C2 and C3 domains (for GFP-2PrG) from *Streptococcus* protein G (Goward et al., 1990). For details of each plasmid, see Table S1.

Yeast strains

FG deletion strains used in this study are derivatives of those described in Strawn et al. (2004) and Terry and Wente (2007). For NPC permeability assays with these strains, their genomic copy of HTB2 was tagged with mCherry. In a subset of these strains, genomic copies of CRM1, KAP95, KAP123, NTF2, NUP116, NDC1, and NUP170 were C-terminally tagged with either GFP(S65T) or mCherry. Genomic tagging cassettes were generated by PCR of either pKT355 (Sheff and Thorn, 2004), in the case of mCherry, or pFA6-GFP(S65T)-HISMX6, for tagging with GFP (Longtine et al., 1998). Details of these strains can be found in Table S2.

For all experiments, yeast strains were grown to early log phase in synthetic complete media containing 2% glucose. For live-cell microscopy, cells were attached to Con A–coated Matek dishes (Zenklusen et al., 2007), in which they were allowed to grow for up to 2 h before preparing a fresh plate.

FRAP assay of NPC permeability

Single yeast cell measurements of permeability were obtained using a substantially more quantitative adaption of a FRAP-based method described by us (Lord et al., 2015).

FRAP experiments were conducted on a DeltaVision deconvolution epifluorescence microscope (Applied Precision Ltd.) fitted with a spot-bleaching laser using a 100× 1.4 NA objective (Olympus). For each assay, a point within the cell was targeted with the bleaching laser, as far from the nucleus as possible. This point was illuminated for 1 s, during which time much of the cytoplasmic pool diffuses through that illuminated area and is bleached, leaving a darkened cytoplasm and a brighter nucleus. Occasionally, a second bleach would be necessary. Over the next 1–10 min, a time course of images of GFP and the Htb2-mCherry nuclear marker were acquired, during which time the visible nuclear and cytoplasmic GFP pools returned to a steady state of approximately equal fluorescence intensities. Subsequently, a stack of GFP and Htb2-mCherry images were acquired in the z axis to obtain the subcellular size and shape measurements necessary for quantitative analysis of the time course.

Images of the 3D image stack were deconvolved for a pseudo-confocal image series, and then the most focused image planes were combined by maximum intensity projection. Nuclear, cytoplasmic, and vacuolar regions were segmented by an automated process based on edge detection of this projection, with these subcellular objects fitted to ellipses. Measurements of nuclear and cytoplasmic volumes and surface areas were obtained from these 2D ellipses by approximating each object as a 3D ellipsoid with a height derived from the mean of the minor and major radii (or just using the minor radii for the height approximation of the cell body, because they tend to sit flat on the slide, like a rugby ball). NPC count of each cell was estimated assuming 11.2 NPCs/μm^2 (Winey et al., 1997). More direct 3D rendering of the cells was not possible using this epifluorescence system, because deconvolution algorithms generate artifacts near the top and bottom of the yeast cells, which tend to bend light like small lenses. In spite of the assumptions required in this approach, the morphological values obtained by this method were within 10% of those obtained by direct measurement of a complete 3D rendering using confocal microscopy.

Time course images were processed and segmented using an automated but user-directed software. Images were corrected for flat-field illumination, and then a four-dimensional watershed split was performed to excise the data of the targeted cell. Time course images were segmented into nuclear and cytoplasmic regions using an algorithm based on edge detection of optimized images of the GFP channel (cytoplasm) and the mCherry channel (nucleus). Background-corrected, mean fluorescence pixel intensities of the nucleus and cytoplasm over time were obtained from the segmented regions, normalized to constant total cellular fluorescence to correct for photobleaching during image acquisition. As we were using a nonconfocal imaging system, cytoplasmic pixel measurements excluded the area near to the brighter nucleus to minimize the effect of out-of-focus light. Conversely, nuclear measurements have negligible effect from cytoplasmic light intrusion, which is considerably darker than the nucleus during the most important part of the time course (and thus contributes very little signal). Note that for the reverse experiment, in which the nucleus is bleached, it is technically more challenging to target the smaller nucleus, and the nuclear pixel measurements are more affected by light haze from the bright surrounding cytoplasm.

Data integration for computing NPC permeability coefficients

In the experiments, we directly measured the nuclear fluorescence over time, the volume of the cytoplasm and the nucleus, and the surface area of the NE. We show how these measurements can be combined to compute the permeability coefficient particular to each strain–substrate pair in units of molecules per second per NPC per micromole (or 1.86×10^{-6} cm/s; explained below in this section). In the analysis, we assumed that the diffusion of GFP-xPrA within the nucleus and cytoplasm is rapid, compared with its exchange across the NE, and that the rates of mass transport across the envelope follow Fick's first law. This approach is similar to that described by Peters and colleagues for their ex vivo single-channel experiments (Keminer and Peters, 1999; Kiskin et al., 2003). The first assumption is supported by the empirical observation that equilibration times in each compartment are less than a second, whereas equilibration across the membrane is approximately several seconds to minutes. The second assumption is validated by the decay in nuclear fluorescence following single exponential kinetics in the FRAP experiments.

Based on these assumptions, the flux of GFP (J) at any time, namely the net amount of substance flowing from the nucleus to the cytoplasm through NPCs per time unit per membrane surface area unit, is simply related to the nuclear (N) and cytoplasmic (C) concentrations of GFP by Fick's first law applied to porous membranes:

$$J = (C - N) \cdot p, \tag{1}$$

where C and N are given in micromoles, p is a coefficient of permeability for the NE given in units of molecules per second per NPC per micromole, particular to this strain and reporter. Note that these units are equivalent to 1.86×10^{-6} cm/s, using the fact that NPCs are regularly spaced throughout the NE at a conveniently constant density of 11.2 NPCs/μm^2 surface area throughout the cell cycle (Winey et al., 1997), consistently so among strains according to measurements of rim intensity of nups (Fig. S3). We favor the use of molecules per second per NPC per micromole because these units are more intuitive in the context of the current study. The flux J, in molecules per second per NPC per micromole, can be also stated in terms of the net change in N over time t in seconds, the nuclear volume Vn in liters, and the Avogadro constant Na, specified in μM^{-1}:

$$J = \frac{dN}{dt} \cdot \frac{Na \cdot Vn}{NPCs}. \tag{2}$$

By rearranging Eqs. 1 and 2, we get:

$$\frac{dN}{dt} = \frac{NPCs \cdot p}{Na \cdot Vn} \cdot (C - N).$$

Published October 3, 2016

The cytoplasmic concentration C can be eliminated entirely if we assume that each cell has a constant amount of GFP substrate, T, and use the cytoplasmic volume Vc.

$$\frac{dN}{dt} = \frac{NPCs \cdot p}{Na \cdot Vn} \cdot \left(\frac{T - N \cdot Vn}{Vc} - N\right).$$

After some simple rearrangements:

$$\frac{dN}{dt} = \frac{NPCs \cdot p}{Na \cdot Vn} \cdot \frac{[T - N \cdot (Vc + Vn)]}{Vc}.$$

This solution for the resulting simple ordinary differential equation (ODE) for N over time (t) is:

$$N = \frac{T}{Vc + Vn} + \alpha\, e^{\frac{NPCs \cdot p \cdot t \cdot (Vc + Vn)}{Na \cdot Vc \cdot Vn}}.$$

Time course traces of nuclear GFP fluorescence (Nf) were therefore fitted to a single exponential,

$$Nf = a + c\, e^{-b\,t}.$$

The time constant of exponential decay τ is the inverse of b, whereas p can be derived from τ, which we measure directly, provided morphological measurements of the cell, thus:

$$p = \frac{Na \cdot Vc \cdot Vn}{\tau \cdot NPCs \cdot (Vc + Vn)}.$$

We note that p is independent of the morphological measurements of the cell, which were only used to compute it from the raw measurements of τ. Thus, p is derived as the permeability of a particular strain's NPCs toward a particular GFP-xPrA variant, in units of net molecules diffusing through a single NPC per second per 1 µM concentration differential across the NPC. For each strain–substrate condition, 14–58 single-cell measurements were made, and the mean value stated ± SEM. All image processing and analyses were handled by MatLab scripts, coordinated by custom graphical user interfaces, which facilitated the evaluation and refinement of multidimensional image segmentation, side-by-side with the resultant fluorescence traces and derived measurements.

Similar measurements of the dynamics of Kap123-GFP were taken using this FRAP method. These measurements also fit well to the basic form of the same single exponential relationship, as expected (Kopito and Elbaum, 2007; Kim and Elbaum, 2013b). However, in this case, the value of the exponential time constant τ includes the transport of Kap123 alone and in an unknown quantity of cargo and/or NTR complexes, each for which the time constant may be dominated by docking probabilities to the NPC rather than transport efficiency (Kim and Elbaum, 2013b).

Estimation of transport time constant for proteins in the Popken et al. dataset

Popken et al. (2015) reported the ratio of nuclear to cytoplasmic fluorescence of inert reporter macromolecules after 3 h of expression and 1 h of equilibration. By assuming exponential decay of the normalized cytoplasmic pool (as observed in all of our FRAP experiments) and that their yeast cells had similar mean NPC counts, cytoplasmic volumes, and nuclear volumes to our own, we estimated the transport time constant (the exponential decay coefficient) $\hat{\tau}$ that would lead to the same nuclear to cytoplasmic fluorescence ratio within an hour. We were able to apply this analysis only to the bulkier macromolecules in the size range 109–230 kD, which transport slowly enough to prevent full equilibration within this time scale.

Sedimentation constant measurement by sucrose gradient ultracentrifugation

WT yeast, expressing each of the GFP-xPrA substrates were harvested then lysed by cryogenic grinding. Cell lysates were clarified by low- and then medium-speed centrifugation, first at 14,000 rpm in a benchtop mini-centrifuge for 15 min and then at 55,000 rpm in a benchtop ultracentrifuge in a TLA55 rotor (Beckman Coulter) for 30 min. Lysates for all substrates were combined, and then 100 µl combined lysate was overlaid on 5–15 or 5–20% sucrose gradients. Overlaid on the same gradients was a commercial mixture of low molecular mass globular protein standards, containing 100–200 µg each protein standards, RNase, albumin, ovalbumin, and chymotrypsinogen A. Sucrose gradients were centrifuged for 19–24 h in an SW55 rotor (Beckman Coulter) at 55,000 rpm. Gradients were unloaded into 1.5-mm fractions using a piston gradient fractionator. All methods were performed at 4°C in solutions buffered with transport buffer with Tween (20 mM Hepes, pH 7.4, 110 mM KOAc, 0.1% Tween-20, and 2 mM $MgCl_2$) spiked with a 1:200 mixture of both a commercial protease inhibitor cocktail (Sigma-Aldrich) and Solution P (0.5 mg/ml Pepstatin A and 20 mg/ml PMSF, in ethanol). Morris buffer concentrate (6×) was added to the gradient fractions for separation by SDS-PAGE. One set of gels was Western blotted and then probed with anti-GFP (Roche), whereas another set of gels was stained with Coomassie to detect the protein standards. Relative quantities and migration profiles of the GFP-xPrA substrates and the globular standard proteins were measured from images of the Western blots or Coomassie-stained gels, respectively. Sedimentation coefficients were calculated from the peak migration distance of each protein using standard calculations (Griffith, 1983). Stated sedimentation values are a mean of three to six such experiments. S values of GFP-xPrA are compared with those of the globular commercial standards in Fig. S1 A.

Diffusion constant measurement by DLS

Recombinant GFP-xPrA fusion proteins were expressed as HIS-tagged versions from pET28 plasmids in BL21-DE3 *Escherichia coli*. Cells were lysed using a microfluidizer, clarified by centrifugation, and the fusion proteins were purified from the lysate via their 6xHIS tags with TALON resin using standard protein purification methods. Purifications were performed in PBS in the presence of 0.1 mM EDTA and a protease inhibitor cocktail. Diffusion constants of 1-mg/ml samples of the recombinantly purified fusion proteins were measured using a DLS instrument (Wyatt Technology) using the manufacturer's Dynamics software. Reported values are the averages of 5–10 measurements.

SAXS

SAXS profiles of five constructs of GFP-PrAs and two constructs of GFP-PrGs (Fig. 3 D and Table S4) were collected at concentrations ranging from 0.5 to 2.0 mg/ml at temperatures of 15°C (beamline 4–2 of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory). All purified, His-tagged SAXS samples were prepared as for the DLS experiments and directly dialyzed overnight at 4°C against 20 mM Hepes, 300 mM NaCl, and 5% glycerol buffer. The 5% glycerol protects proteins from radiation damage during x-ray exposure (Kuwamoto et al., 2004). Up to 20 1-s exposures at SSRL beamline 4–2 were used for each sample. The SAXS profile of each buffer was obtained in the same manner and subtracted from a protein SAXS profile. The buffer subtracted SAXS profiles were initially analyzed using the ATSAS package (Petoukhov et al., 2012) to calculate radius of gyration (R_g) and maximum particle size (D_{max}; Table S4). The mean of the smaller scattering angle regions ($q < 0.15$ Å^{-1}) of the lower concentration profiles (~0.5 mg/ml) and the mean of the wider scattering

Published October 3, 2016

angle regions ($q > 0.12$ Å^{-1}) of the higher concentration (~2.0 mg/ml) profiles were merged to obtain the final experimental SAXS profiles.

Shape modeling based on SAXS and ultracentrifugation data

The ab initio shape of each of the substrates in our dataset of GFP-PrA macromolecules (Fig. 3 D, substrates ≤68 kD) were generated from their corresponding SAXS profiles by running DAMMIF (Franke and Svergun, 2009) 20 times and then refined through an additional 40 DAMMIN (Svergun, 1999) runs followed by superposition and averaging with DAMAVER (Volkov and Svergun, 2003). Because S_{max}/S were highly correlated with the shape asymmetry measured by SAXS (Fig. 3 C, top) and consistent with diffusion coefficient measurements by DLS (Fig. 3 D, bottom), we modeled the shapes for the larger substrates in the Popken et al. (2015) dataset based on their S_{max}/S ratios (Fig. 3 D). For illustration purposes, the crystallographic models of GFP (accession no. 1EMA), PrA (accession no. 1ZXG), PrA (accession no. 1GB1), and maltose binding protein (accession no. 1ANF) subunits were fitted to the enclosing envelopes.

Comparison of aqueous channel model to experimental measurements

The empirically tested theory of diffusion through porous aqueous membranes predicts that the permeability coefficient p of porous membranes depends on the particle probability to enter the pore (Eq. 3) multiplied by the particle friction with the pore walls (Eq. 4; Renkin, 1954):

$$F_1 = \left(1 - \frac{r}{r_0}\right)^2, \quad (3)$$

$$F_2 = 1 - 2.1\frac{r}{r_0} + 2.1\left(\frac{r}{r_0}\right)^3 - 0.95\left(\frac{r}{r_0}\right)^5, \quad (4)$$

and $p = F_1 \cdot F_2$, where r is the particle radius and r_0 is the pore radius. We show that this model of a narrow aqueous channel results in a rigid rather than soft barrier to passive transport that is incompatible with the data (Fig. 3 F; using the approximation $r \sim MM^{1/3}$ for globular proteins [Erickson, 2009]). More formally, when $r << r_0$, then $F_1 \approx 1$ and so F_2 dominates p, and in this case, it is easy to show by taking the first derivative of F_2 that p decreases extremely slowly with respect to either radius or molecular mass. In particular, in this radius range, it decreases more slowly than the cubic relation observed experimentally (Fig. 3 F). Conversely, as r/r_0 reaches 1.0, p is dominated by F_1, which sharply drops to zero at that limit, as can be shown by taking the derivative of F_1. Specifically, it cannot be approximated by an inverse-cubic relation between p and the molecular mass, except over an extremely narrow range of radii or molecular masses: the derivative of an inverse-cubic function takes the value −3 on a log-log scale, so we can take the derivative of $\ln(F_1)$ with respect to $ln(MW)$ and compare to −3, again assuming $r \sim MM^{1/3}$:

$$\frac{dLn(F_1)}{dLn(MM)} = \frac{-2}{3}\frac{r}{r_0 - r} = \frac{-2}{3}\frac{MM^{1/3}}{MM_0^{1/3} - MM^{1/3}}. \quad (5)$$

Eq. 5 takes the value −3 ± 0.5 (approximately cubic) only when r/r_0 is between 79 and 84%, or stated in terms of molecular masses, when MM/MM_0 is between 49 and 59% of MM_0, the molecular mass of a spherical molecule of radius r_0. Hence, a simple aqueous pore model cannot be used to approximate the cubic relation that we observed over the wide range 27 to 230 kD, or for that matter, with any other soft gradually decaying (long-tailed) size-dependent barrier (compare Fig. 1 A and Fig. 3 F). We conclude that passive transport cannot correspond to a translocation through a simple aqueous channel that obeys the classic models of porous membranes and that a more elaborate model is needed to explain the observed transport rates.

Brownian dynamics simulations

We modeled the transport system using our open-source Integrative Modeling Platform package (http://integrativemodeling.org; Russel et al., 2012). The simulation model consists of the NE, the NPC, and the diffusing substrates, all enclosed within a bounding box of 100 × 100 × 100 nm (Fig. 4 A). The NPC central channel was modeled as a 15-nm–high cylinder with the inner diameter of 27 nm, forming a cavity within the NE (control simulations on a slightly larger NPC that better reflects the known NPC geometry are described at the end of this section). 32 chains of FG repeats were anchored to the NPC scaffold in 4 layers of 8 chains each. Each chain was modeled as a flexible string of 45 beads of radius 0.6 nm each (Fig. 4 A, inset). Consecutive beads were connected an equilibrium length of 1.8 nm between consecutive bead centers and using a linear potential, approximating the springlike nature of flexible polymers in general (van der and Maarel, 2008) and FG repeat domains in particular (Denning et al., 2003; Lim et al., 2006; Hough et al., 2015; Milles et al., 2015; Lemke, 2016; Raveh et al., 2016; Sakiyama et al., 2016), but using a constant force of 1.0 kcal/M/A as a first approximation. Substrate molecules were modeled as rigid spheres of molecular masses from 28 to 223 (radii of 2.0 to 4.0 nm, assuming constant protein density of 1.38 g/cm^2; Fig. 4 A, inset). 100 substrate molecules were included, resulting in the substrate concentration of ~0.2 mM. Excluded volume interactions were approximated using a constant repulsive force of 20.0 kcal/M/A, which was exerted on overlapping spheres or on spheres that penetrated either the NE or the bounding box. To approximate the differences in the diffusion rates of molecules of different sizes, we assume for simplicity that the Stokes radius of each sphere equals its actual radius R in angstroms and assign each sphere a translational diffusion coefficient using the Stokes-Einstein equation:

$$D = \frac{kT}{6\pi\eta R}$$

in A^2/fs units, where $\eta = 0.92$ is the dynamic viscosity of water at 297.15 K in mPa*s units and kT is given in femtojoules. Brownian dynamics simulations were implemented in Integrative Modeling Platform (Russel et al., 2012) in standard form (Ermak and McCammon, 1978) at 297.15 K. The force vector $\vec{f}_i(t)$ acting on the i'th particle at time step t is computed from the gradient of the energy function. The particle coordinates X_i are then updated using the discrete integration equation:

$$X_i(t+1) = X_i(t) + \frac{\Delta t}{kT} D\vec{f}_i(t) + R(\sqrt{D\Delta t}),$$

where $R(\sqrt{D\Delta t})$ is a normally distributed random variable with zero mean and variance $D\Delta t$. Hydrodynamic interactions were neglected.

The dynamics of the entire system were simulated for 20.0 µs with a time step of 470 fs. The first 10.0 µs of each simulation were considered as equilibration time and discarded from the subsequent statistical analysis. The entire process was repeated 20,000 times independently. By collecting statistics from each of the 100 copies of each substrate type, we obtained data from 10.0 s of simulation per particle type using a total computation time of ~9,000 central processing unit days on our local Linux computer cluster. The simulations with modified FG repeat densities were conducted as described in the first paragraph of this section, but the number of chains of FG repeats was decreased to either 28, 24, 20, or 16 chains, arranged in four layers as before. Control simulations were also performed on a larger system that more closely approximates the known dimensions of the yeast NPC and the NE (Yang et al., 1998; Alber et al., 2007a; inner diameter of 38 nm and NE width of 30 nm), with seven layers of eight chains of FG repeats each to maintain the same grafting density. These control simulations produced the same overall trends as those of the smaller

Published October 3, 2016

systems (Fig. S2 D). Another control simulation used an eightfold larger box size (8.7-fold larger cytoplasmic and nuclear volumes) to verify that we model the effect of changes in these volumes correctly, resulting in a ninefold decrease in simulated transport rates. The smaller systems were used for the rest of the analysis to obtain better-converged statistics. Control simulations of cohesive FG domains were conducted by introducing a binding site on a random surface point on each of the FG beads with an attractive range of 6 Å and an attractive force constant of 2.5 kcal/M/Å (potential binding energy of 15 kcal/M per site), acting between pairs of FG beads. Each bead was also assigned a random rotational diffusion coefficient using the Einstein-Stokes equation. Torques between binding sites were integrated similarly to the integration of translational forces described earlier in this section for $X_i(t + 1)$. The configuration files for all simulations are available at https://zenodo.org/record/154746#.V-WW-pMrJE6.

Glossary

τ is the exponential time constant in seconds for the measured decay in nuclear fluorescence after photobleaching in individual cells. At the single-molecule level, τ is the mean timespan between two consecutive translocations through the NE. τ is a function of p, the cytoplasmic volume Vc, the nuclear volume Vn, and the number of NPCs in the cell:

$$\tau = \frac{Na \cdot Vc \cdot Vn}{\mathrm{p} \cdot NPCs \cdot (Vc + Vn)},$$

where Na is Avogadro's number in μmol^{-1} (Materials and methods). $\hat{\tau}$ is the hypothetical value of τ in a typical yeast cell with mean values of cytoplasmic volume, nuclear volume, and NE surface area. p is the permeability coefficient of the NE in molecules per second per NPC per micromole (flux per concentration difference), or equivalently, 1.86×10^{-6} cm/s, assuming a density of 11.2 NPCs/μm^2; p is inversely proportional to $\hat{\tau}$. $\Delta G_{passive}$ is the free energy penalty for the translocation of an inert macromolecule from the cytoplasm or the nucleus into the NPC.

Computation of free energy landscape from the simulations

The positional distribution of the simulated particles was computed over a 100 × 100 × 100 grid with voxel size of 1 × 1 × 1 nm in time intervals of 10 ns over all simulations. The 2D energy landscape (Fig. 5 E) spanned by the z axis (projected position along the NPC central axis) and the x axis (arbitrary direction on the orthogonal plane parallel with the NE) was computed by averaging over all planes that lie on the central axis of the NPC and transformed to an energy landscape with units of kT based on the Boltzmann distribution by taking the log of these densities. Similarly, the energy landscape across the NPC (Fig. 5 F) was computed by obtaining the marginal distribution along the z axis (integrating over the x and y coordinates; Fig. 5 F).

Transport factor and nup distribution analysis

The steady-state distributions of genomically tagged Kap95-mCherry, Kap123-GFP, Ntf2-GFP, Crm1-GFP, Nup116-GFP, Nup170-GFP, and Ndc1-GFP were observed in each of the representative yeast strains, $nup116^{\Delta FG}$, $nsp1^{\Delta FG}$, $N^{\Delta FG}$ $C^{\Delta FG}$ $nsp1^{\Delta FG}$, and WT. All images for each fluorescently tagged protein were acquired on the same day for each strain to avoid daily fluctuations in fluorescence measurements. For each combination of fusion protein and strain, images of fields of live cells were acquired at 100× magnification on an Axiovert 200 microscope (ZEISS) fitted with an UltraView spinning disk confocal imaging head (PerkinElmer), using an iXon EMCCD camera (Andor Technology), controlled by MetaMorph imaging software. A series of images through the complete z axis of each field was acquired.

Semiautomated image analysis was performed using a series of image segmentation, measurement and data-review scripts, and graphical user interfaces created in MatLab for this expressed purpose. Each image was corrected for flat field illumination, background subtracted, and then all images in the z axis were collapsed by maximum projection down to a single image for each field of cells. Using image segmentation scripts, cell bodies were detected and then segmented into nuclear, cytoplasmic, and nuclear rim regions. These segmentation algorithms followed the basic algorithmic structure of edge detection, then edge completion, followed by automatic extraneous object exclusion. Nuclear fluorescence measurements were derived simply as mean nuclear pixel intensity, whereas cytoplasmic fluorescence was obtained from a k-means analysis of the extranuclear pixels (this analysis was calibrated to exclude dark vacuolar regions and the bleed-through light at perinuclear regions). Where possible, the scripts calculated nuclear rim fluorescence by tracing the brightest pixels around the periphery of the nuclear regions and then averaged this summed pixel intensity over perimeter distance. This analysis was performed to derive N/C fluorescence ratios and nuclear rim intensity measurements for 98–1,401 individual cells from each of the 27 strain-reporter combinations; a total of 9,621 cells were thus measured using these methods.

Statistical regression analysis

Statistical regression analysis using LASSO was done using the MatLab implementation of LASSO. LASSO is particularly suitable for finding statistical correlations in small datasets, because it deploys a regularization scheme that effectively penalizes solutions with a larger number of variables, thereby reducing the risk of overfitting (Tibshirani, 1996).

Preprocessing. For each of 13 strains (WT plus 12 mutant strains with FG repeat deletions), 33 features (Fig. S5 and Tables S3 and S4) that quantitatively describe the NPC composition, geometry, and biophysical properties of that strain were gathered. A 13 × 33 matrix M was constructed, with M_{ij} being the value of the j'th feature in the i'th strain. The regression analysis aimed to create a predictor from the rows of M (the vectors of features for each strain). The columns of M (the vectors of each feature) were centered by subtractions of their mean and normalized by division with their standard errors. In addition, a 13-dimensional response vector Y was constructed with the WT relative permeability of each strain, formally defined as the ratio between the permeability of each strain and the WT permeability, averaged across all cargo types for which a measurement was taken in that strain. The regression analysis aimed to create a predictor from the rows of M (the feature vector of each strain) to Y. Conversely, each feature was represented as a 13-dimensional feature vector f, in which each vector coordinate f_j represents the value of that feature in strain j. The coordinates of f and s were normalized by subtracting the feature mean and dividing by the feature SD across all mutant strains.

To remove strong statistical redundancies that obscure the regression analysis, we first clustered the 33 feature vectors using their mutual Pearson's correlation values and randomly selected representative features from cliques of features in which all pairs of features had Pearson's correlation >0.9. We then applied PCA to further eliminate redundant degrees of freedom while maximizing the information content in each selected degree of freedom (Jolliffe, 1982). We applied the PCA/LASSO method iteratively, removing features that were selected with extremely low weights until the process converged on a final list of eight features from two highly correlated subgroups that are predicted to have the greatest effect on passive transport rates (features related to charge to hydrophobicity and features related to proximity to the cytoplasmic end of the pore).

Online supplemental material

Fig. S1 shows supporting analyses of shape effects in relation to Fig. 3. Fig. S2 shows the thermodynamics and kinetics of simulated

passive diffusion in relation to Fig. 4. Fig. S3 shows the general effect on the transport machinery in FG deletion strains, providing controls for our measurements of passive diffusion. Fig. S4 shows Kap123 dynamics, which seem independent of passive NPC permeability, but are greatly reduced upon deletion of particular FG regions. Fig. S5 shows relative transport times plotted as a function of individual features of the mutant strains and the corresponding R^2 correlation coefficients. Fig. S6 shows an assessment of the robustness of positive and negative contribution of various FG Nup features in the combined PCA/LASSO regression model of passive transport in different settings. Fig. S7 shows a sequence and spatial analysis of the different FG Nups and mutant strains. Fig. S8 shows the change in the ratio between nuclear and cytoplasmic fluorescence at equilibrium relative to the WT strain for the different GFP-xPrA substrates in the nup116$^{\Delta FG}$ strain. Video 1 shows FRAP permeability assays of four representative WT cells, each expressing a different-sized GFP-xPrA substrate. Video 2 shows sample simulation trajectory of passive diffusion of macromolecules of different molecular masses/radii through the NPC, using the model described in the text and in Fig. 4 and Fig. S2. Table S1 shows the plasmids used in this study. Table S2 shows the strains used in this study. Table S3 shows the features used to describe the NPCs of mutant strains and for the regression analysis. Table S4 shows the SAXS measurements of maximal end-to-end distances and radii of gyration for GFP-xPrA and GFP-xPrG molecules.

Acknowledgments

We thank Liesbeth Veenhoff and Patrick Onck for helpful discussion, sharing of unpublished data, and critical reading of the manuscript; and Laura Terry, who assisted in the assembly of strains for these studies.

This work was supported by National Institutes of Health grants P41 GM109824 (to M.P. Rout and A. Sali), R01 GM112108 (to M.P. Rout and A. Sali), U01 GM098256 (to M.P. Rout and A. Sali), and R01 GM071329 (to M.P. Rout).

The authors declare no competing financial interests.

Submitted: 4 January 2016
Accepted: 6 September 2016

References

Adams, R.L., L.J. Terry, and S.R. Wente. 2015. A novel *Saccharomyces cerevisiae* FG nucleoporin mutant collection for use in nuclear pore complex functional experiments. *G3 (Bethesda).* 6:51–58. http://dx.doi.org/10.1534/g3.115.023002

Alber, F., S. Dokudovskaya, L.M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B.T. Chait, et al. 2007a. Determining the architectures of macromolecular assemblies. *Nature.* 450:683–694. http://dx.doi.org/10.1038/nature06404

Alber, F., S. Dokudovskaya, L.M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B.T. Chait, et al. 2007b. The molecular architecture of the nuclear pore complex. *Nature.* 450:695–701. http://dx.doi.org/10.1038/nature06405

Atkinson, C.E., A.L. Mattheyses, M. Kampmann, and S.M. Simon. 2013. Conserved spatial organization of FG domains in the nuclear pore complex. *Biophys. J.* 104:37–50. http://dx.doi.org/10.1016/j.bpj.2012.11.3823

Bayliss, R., K. Ribbeck, D. Akin, H.M. Kent, C.M. Feldherr, D. Görlich, and M. Stewart. 1999. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J. Mol. Biol.* 293:579–593. http://dx.doi.org/10.1006/jmbi.1999.3166

Bayliss, R., S.W. Leung, R.P. Baker, B.B. Quimby, A.H. Corbett, and M. Stewart. 2002. Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J.* 21:2843–2853. http://dx.doi.org/10.1093/emboj/cdf305

Christie, M., C.W. Chang, G. Rona, K.M. Smith, A.G. Stewart, A.A. Takeda, M.R. Fontes, M. Stewart, B.G. Vertessy, J.K. Forwood, and B. Kobe. 2016. Structural biology and regulation of protein import into the nucleus. *J. Mol. Biol.* 428:2060–2090. http://dx.doi.org/10.1016/j.jmb.2015.10.023

Colwell, L.J., M.P. Brenner, and K. Ribbeck. 2010. Charge as a selection criterion for translocation through the nuclear pore complex. *PLOS Comput. Biol.* 6:e1000747. http://dx.doi.org/10.1371/journal.pcbi.1000747

Denning, D.P., S.S. Patel, V. Uversky, A.L. Fink, and M. Rexach. 2003. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc. Natl. Acad. Sci. USA.* 100:2450–2455. http://dx.doi.org/10.1073/pnas.0437902100

Dunker, A.K., I. Silman, V.N. Uversky, and J.L. Sussman. 2008. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* 18:756–764. http://dx.doi.org/10.1016/j.sbi.2008.10.002

Eibauer, M., M. Pellanda, Y. Turgay, A. Dubrovsky, A. Wild, and O. Medalia. 2015. Structure and gating of the nuclear pore complex. *Nat. Commun.* 6:7532. http://dx.doi.org/10.1038/ncomms8532

Erickson, H.P. 2009. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proced. Online.* 11:32–51. http://dx.doi.org/10.1007/s12575-009-9008-x

Ermak, D.L., and J.A. McCammon. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 69:1352–1360. http://dx.doi.org/10.1063/1.436761

Eyring, H. 1935. The activated complex in chemical reactions. *J. Chem. Phys.* 3:107–115. http://dx.doi.org/10.1063/1.1749604

Fischer, H., I. Polikarpov, and A.F. Craievich. 2004. Average protein density is a molecular-weight-dependent function. *Protein Sci.* 13:2825–2828. http://dx.doi.org/10.1110/ps.04688204

Franke, D., and D.I. Svergun. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* 42:342–346. http://dx.doi.org/10.1107/S0021889809000338

Frey, S., and D. Görlich. 2007. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell.* 130:512–523. http://dx.doi.org/10.1016/j.cell.2007.06.024

Gavin, A.C., M. Bösche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature.* 415:141–147. http://dx.doi.org/10.1038/415141a

Ghavami, A., L.M. Veenhoff, E. van der Giessen, and P.R. Onck. 2014. Probing the disordered domain of the nuclear pore complex through coarse-grained molecular dynamics simulations. *Biophys. J.* 107:1393–1402. http://dx.doi.org/10.1016/j.bpj.2014.07.060

Ghavami, A., E. van der Giessen, and P.R. Onck. 2016. Energetics of transport through the nuclear pore complex. *PLoS One.* 11:e0148876. http://dx.doi.org/10.1371/journal.pone.0148876

Goryaynov, A., and W. Yang. 2014. Role of molecular charge in nucleocytoplasmic transport. *PLoS One.* 9:e88792. http://dx.doi.org/10.1371/journal.pone.0088792

Goward, C.R., J.P. Murphy, T. Atkinson, and D.A. Barstow. 1990. Expression and purification of a truncated recombinant streptococcal protein G. *Biochem. J.* 267:171–177. http://dx.doi.org/10.1042/bj2670171

Griffith, O.M. 1983. Techniques of preparative, zonal, and continuous flow ultracentrifugation. Beckman Intruments Inc. Spinco Division, Palo Alto, CA. 50 pp.

Grünwald, D., and R.H. Singer. 2010. In vivo imaging of labelled endogenous β-actin mRNA during nucleocytoplasmic transport. *Nature.* 467:604–607. http://dx.doi.org/10.1038/nature09438

Hough, L.E., K. Dutta, S. Sparks, D.B. Temel, A. Kamal, J. Tetenbaum-Novatt, M.P. Rout, and D. Cowburn. 2015. The molecular mechanism of nuclear transport revealed by atomic-scale measurements. *eLife.* 4:4. http://dx.doi.org/10.7554/eLife.10027

Hülsmann, B.B., A.A. Labokha, and D. Görlich. 2012. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell.* 150:738–751. http://dx.doi.org/10.1016/j.cell.2012.07.019

Iovine, M.K., and S.R. Wente. 1997. A nuclear export signal in Kap95p is required for both recycling the import factor and interaction with the nucleoporin GLFG repeat regions of Nup116p and Nup100p. *J. Cell Biol.* 137:797–811. http://dx.doi.org/10.1083/jcb.137.4.797

Jolliffe, I.T. 1982. A note on the use of principal components in regression. *Appl. Stat.* 31:300–303. http://dx.doi.org/10.2307/2348005

Jovanovic-Talisman, T., J. Tetenbaum-Novatt, A.S. McKenney, A. Zilman, R. Peters, M.P. Rout, and B.T. Chait. 2009. Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. *Nature.* 457:1023–1027. http://dx.doi.org/10.1038/nature07600

Kapinos, L.E., R.L. Schoch, R.S. Wagner, K.D. Schleicher, and R.Y. Lim. 2014. Karyopherin-centric control of nuclear pores based on molecular occupancy and kinetic analysis of multivalent binding with FG nucleoporins. *Biophys. J.* 106:1751–1762. http://dx.doi.org/10.1016/j.bpj.2014.02.021

Keilhauer, E.C., M.Y. Hein, and M. Mann. 2015. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol. Cell. Proteomics.* 14:120–135. http://dx.doi.org/10.1074/mcp.M114.041012

Keminer, O., and R. Peters. 1999. Permeability of single nuclear pores. *Biophys. J.* 77:217–228. http://dx.doi.org/10.1016/S0006-3495(99)76883-9

Kim, S., and M. Elbaum. 2013a. Enzymatically driven transport: a kinetic theory for nuclear export. *Biophys. J.* 105:1997–2005. http://dx.doi.org/10.1016/j.bpj.2013.09.011

Kim, S., and M. Elbaum. 2013b. A simple kinetic model with explicit predictions for nuclear transport. *Biophys. J.* 105:565–569. http://dx.doi.org/10.1016/j.bpj.2013.04.025

Kirli, K., S. Karaca, H.J. Dehne, M. Samwer, K.T. Pan, C. Lenz, H. Urlaub, and D. Görlich. 2015. A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *eLife.* 4:4. http://dx.doi.org/10.7554/eLife.11466

Kiskin, N.I., J.P. Siebrasse, and R. Peters. 2003. Optical microwell assay of membrane transport kinetics. *Biophys. J.* 85:2311–2322. http://dx.doi.org/10.1016/S0006-3495(03)74655-4

Knockenhauer, K.E., and T.U. Schwartz. 2016. The nuclear pore complex as a flexible and dynamic gate. *Cell.* 164:1162–1171. http://dx.doi.org/10.1016/j.cell.2016.01.034

Kopito, R.B., and M. Elbaum. 2007. Reversibility in nucleocytoplasmic transport. *Proc. Natl. Acad. Sci. USA.* 104:12743–12748. http://dx.doi.org/10.1073/pnas.0702690104

Kosinski, J., S. Mosalaganti, A. von Appen, R. Teimer, A.L. DiGuilio, W. Wan, K.H. Bui, W.J. Hagen, J.A. Briggs, J.S. Glavy, et al. 2016. Molecular architecture of the inner ring scaffold of the human nuclear pore complex. *Science.* 352:363–365. http://dx.doi.org/10.1126/science.aaf0643

Kuwamoto, S., S. Akiyama, and T. Fujisawa. 2004. Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants. *J. Synchrotron Radiat.* 11:462–468. http://dx.doi.org/10.1107/S0909049504019272

Lemke, E.A. 2016. The multiple faces of disordered nucleoporins. *J. Mol. Biol.* 428:2011–2024. http://dx.doi.org/10.1016/j.jmb.2016.01.002

Lim, R.Y., N.P. Huang, J. Köser, J. Deng, K.H. Lau, K. Schwarz-Herion, B. Fahrenkrog, and U. Aebi. 2006. Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. *Proc. Natl. Acad. Sci. USA.* 103:9512–9517. http://dx.doi.org/10.1073/pnas.0603521103

Longtine, M.S., A. McKenzie III, D.J. Demarini, N.G. Shah, A. Wach, A. Brachat, P. Philippsen, and J.R. Pringle. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast.* 14:953–961. http://dx.doi.org/10.1002/(SICI)1097-0061(199807)14:10<953::AID-YEA293>3.0.CO;2-U

Lord, C.L., B.L. Timney, M.P. Rout, and S.R. Wente. 2015. Altering nuclear pore complex function impacts longevity and mitochondrial function in *S. cerevisiae*. *J. Cell Biol.* 208:729–744. http://dx.doi.org/10.1083/jcb.201412024

Lowe, A.R., J.J. Siegel, P. Kalab, M. Siu, K. Weis, and J.T. Liphardt. 2010. Selectivity mechanism of the nuclear pore complex characterized by single cargo tracking. *Nature.* 467:600–603. http://dx.doi.org/10.1038/nature09285

Lowe, A.R., J.H. Tang, J. Yassif, M. Graf, W.Y. Huang, J.T. Groves, K. Weis, and J.T. Liphardt. 2015. Importin-β modulates the permeability of the nuclear pore complex in a Ran-dependent manner. *eLife.* 4:e04052. http://dx.doi.org/10.7554/eLife.04052

Ma, J., A. Goryaynov, A. Sarma, and W. Yang. 2012. Self-regulated viscous channel in the nuclear pore complex. *Proc. Natl. Acad. Sci. USA.* 109:7326–7331. http://dx.doi.org/10.1073/pnas.1201724109

Ma, J., A. Goryaynov, and W. Yang. 2016. Super-resolution 3D tomography of interactions and competition in the nuclear pore complex. *Nat. Struct. Mol. Biol.* 23:239–247. http://dx.doi.org/10.1038/nsmb.3174

Marsh, J.A., and J.D. Forman-Kay. 2010. Sequence determinants of compaction in intrinsically disordered proteins. *Biophys. J.* 98:2383–2390. http://dx.doi.org/10.1016/j.bpj.2010.02.006

Mettenleiter, T.C. 2016. Breaching the barrier—The nuclear envelope in virus infection. *J. Mol. Biol.* 428:1949–1961. http://dx.doi.org/10.1016/j.jmb.2015.10.001

Mi, L., A. Goryaynov, A. Lindquist, M. Rexach, and W. Yang. 2015. Quantifying nucleoporin stoichiometry inside single nuclear pore complexes in vivo. *Sci. Rep.* 5:9372. http://dx.doi.org/10.1038/srep09372

Milles, S., D. Mercadante, I.V. Aramburu, M.R. Jensen, N. Banterle, C. Koehler, S. Tyagi, J. Clarke, S.L. Shammas, M. Blackledge, et al. 2015. Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell.* 163:734–745. http://dx.doi.org/10.1016/j.cell.2015.09.047

Mincer, J.S., and S.M. Simon. 2011. Simulations of nuclear pore transport yield mechanistic insights and quantitative predictions. *Proc. Natl. Acad. Sci. USA.* 108:E351–E358. http://dx.doi.org/10.1073/pnas.1104521108

Mohr, D., S. Frey, T. Fischer, T. Güttler, and D. Görlich. 2009. Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J.* 28:2541–2553. http://dx.doi.org/10.1038/emboj.2009.200

Moussavi-Baygi, R., Y. Jamali, R. Karimi, and M.R. Mofrad. 2011. Brownian dynamics simulation of nucleocytoplasmic transport: a coarse-grained model for the functional state of the nuclear pore complex. *PLOS Comput. Biol.* 7:e1002049. http://dx.doi.org/10.1371/journal.pcbi.1002049

Musser, S.M., and D. Grünwald. 2016. Deciphering the structure and function of nuclear pores using single-molecule fluorescence approaches. *J. Mol. Biol.* 428:2091–2119. http://dx.doi.org/10.1016/j.jmb.2016.02.023

Naim, B., D. Zbaida, S. Dagan, R. Kapon, and Z. Reich. 2009. Cargo surface hydrophobicity is sufficient to overcome the nuclear pore complex selectivity barrier. *EMBO J.* 28:2697–2705. http://dx.doi.org/10.1038/emboj.2009.225

Paine, P.L., and C.M. Feldherr. 1972. Nucleocytoplasmic exchange of macromolecules. *Exp. Cell Res.* 74:81–98. http://dx.doi.org/10.1016/0014-4827(72)90483-1

Paine, P.L., L.C. Moore, and S.B. Horowitz. 1975. Nuclear envelope permeability. *Nature.* 254:109–114. http://dx.doi.org/10.1038/254109a0

Patel, S.S., B.J. Belmont, J.M. Sante, and M.F. Rexach. 2007. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell.* 129:83–96. http://dx.doi.org/10.1016/j.cell.2007.01.044

Peleg, O., M. Tagliazucchi, M. Kröger, Y. Rabin, and I. Szleifer. 2011. Morphology control of hairy nanopores. *ACS Nano.* 5:4737–4747. http://dx.doi.org/10.1021/nn200702u

Peters, R. 2005. Translocation through the nuclear pore complex: selectivity and speed by reduction-of-dimensionality. *Traffic.* 6:421–427. http://dx.doi.org/10.1111/j.1600-0854.2005.00287.x

Petoukhov, M.V., D. Franke, A.V. Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, C. Gorba, H.D.T. Mertens, P.V. Konarev, and D.I. Svergun. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Cryst.* 45:342–350. http://dx.doi.org/10.1107/S0021889812007662

Popken, P., A. Ghavami, P.R. Onck, B. Poolman, and L.M. Veenhoff. 2015. Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex. *Mol. Biol. Cell.* 26:1386–1394. http://dx.doi.org/10.1091/mbc.E14-07-1175

Quimby, B.B., and M. Dasso. 2003. The small GTPase Ran: interpreting the signs. *Curr. Opin. Cell Biol.* 15:338–344. http://dx.doi.org/10.1016/S0955-0674(03)00046-2

Raveh, B., J.M. Karp, S. Sparks, K. Dutta, M.P. Rout, A. Sali, and D. Cowburn. 2016. Slide-and-exchange mechanism for rapid and selective transport through the nuclear pore complex. *Proc. Natl. Acad. Sci. USA.* 113:E2489–E2497. http://dx.doi.org/10.1073/pnas.1522663113

Reits, E.A., and J.J. Neefjes. 2001. From fixed to FRAP: measuring protein mobility and activity in living cells. *Nat. Cell Biol.* 3:E145–E147. http://dx.doi.org/10.1038/35078615

Renkin, E.M. 1954. Filtration, diffusion, and molecular sieving through porous cellulose membranes. *J. Gen. Physiol.* 38:225–243.

Ribbeck, K., and D. Görlich. 2001. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* 20:1320–1330. http://dx.doi.org/10.1093/emboj/20.6.1320

Ribbeck, K., and D. Görlich. 2002. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* 21:2664–2671. http://dx.doi.org/10.1093/emboj/21.11.2664

Rout, M.P., and S.R. Wente. 1994. Pores for thought: nuclear pore complex proteins. *Trends Cell Biol.* 4:357–365. http://dx.doi.org/10.1016/0962-8924(94)90085-X

Rout, M.P., J.D. Aitchison, A. Suprapto, K. Hjertaas, Y. Zhao, and B.T. Chait. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* 148:635–651. http://dx.doi.org/10.1083/jcb.148.4.635

Rout, M.P., J.D. Aitchison, M.O. Magnasco, and B.T. Chait. 2003. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* 13:622–628. http://dx.doi.org/10.1016/j.tcb.2003.10.007

Russel, D., K. Lasker, B. Webb, J. Velázquez-Muriel, E. Tjioe, D. Schneidman-Duhovny, B. Peterson, and A. Sali. 2012. Putting the pieces together: integrative modeling platform software for structure determination of

macromolecular assemblies. *PLoS Biol.* 10:e1001244. http://dx.doi.org/10.1371/journal.pbio.1001244

Sakiyama, Y., A. Mazur, L.E. Kapinos, and R.Y. Lim. 2016. Spatiotemporal dynamics of the nuclear pore complex transport barrier resolved by high-speed atomic force microscopy. *Nat. Nanotechnol.* 11:719–723. http://dx.doi.org/10.1038/nnano.2016.62

Schellhaus, A.K., P. De Magistris, and W. Antonin. 2016. Nuclear reformation at the end of mitosis. *J. Mol. Biol.* 428:1962–1985. http://dx.doi.org/10.1016/j.jmb.2015.09.016

Schmidt, H.B., and D. Görlich. 2015. Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *eLife.* 4:4. http://dx.doi.org/10.7554/eLife.04251

Schmidt, H.B., and D. Görlich. 2016. Transport selectivity of nuclear pores, phase separation, and membraneless organelles. *Trends Biochem. Sci.* 41:46–61. http://dx.doi.org/10.1016/j.tibs.2015.11.001

Schmitz, K.S. 1990. An Introduction to Dynamic Light Scattering by Macromolecules. Elsevier Science Publishing Co. Inc., New York. 470 pp.

Sheff, M.A., and K.S. Thorn. 2004. Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast.* 21:661–670. http://dx.doi.org/10.1002/yea.1130

Shulga, N., N. Mosammaparast, R. Wozniak, and D.S. Goldfarb. 2000. Yeast nucleoporins involved in passive nuclear envelope permeability. *J. Cell Biol.* 149:1027–1038. http://dx.doi.org/10.1083/jcb.149.5.1027

Simon, D.N., and M.P. Rout. 2014. Cancer and the nuclear pore complex. *Adv. Exp. Med. Biol.* 773:285–307. http://dx.doi.org/10.1007/978-1-4899-8032-8_13

Smith, A., A. Brownawell, and I.G. Macara. 1998. Nuclear import of Ran is mediated by the transport factor NTF2. *Curr. Biol.* 8:1403–1406. http://dx.doi.org/10.1016/S0960-9822(98)00023-2

Strambio-De-Castillia, C., M. Niepel, and M.P. Rout. 2010. The nuclear pore complex: bridging nuclear transport and gene regulation. *Nat. Rev. Mol. Cell Biol.* 11:490–501. http://dx.doi.org/10.1038/nrm2928

Strawn, L.A., T. Shen, N. Shulga, D.S. Goldfarb, and S.R. Wente. 2004. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat. Cell Biol.* 6:197–206. http://dx.doi.org/10.1038/ncb1097

Svergun, D.I. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* 76:2879–2886. http://dx.doi.org/10.1016/S0006-3495(99)77443-6

Tagliazucchi, M., O. Peleg, M. Kröger, Y. Rabin, and I. Szleifer. 2013. Effect of charge, hydrophobicity, and sequence of nucleoporins on the translocation of model particles through the nuclear pore complex. *Proc. Natl. Acad. Sci. USA.* 110:3363–3368. (published erratum appears in *Proc. Natl. Acad. Sci. USA*. 2013. 110:10336–10337) http://dx.doi.org/10.1073/pnas.1212909110

Tcherkasskaya, O., E.A. Davidson, and V.N. Uversky. 2003. Biophysical constraints for protein structure prediction. *J. Proteome Res.* 2:37–42. http://dx.doi.org/10.1021/pr025552q

Terry, L.J., and S.R. Wente. 2007. Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. *J. Cell Biol.* 178:1121–1132. http://dx.doi.org/10.1083/jcb.200704174

Tetenbaum-Novatt, J., and M.P. Rout. 2010. The mechanism of nucleocytoplasmic transport through the nuclear pore complex. *Cold Spring Harb. Symp. Quant. Biol.* 75:567–584. http://dx.doi.org/10.1101/sqb.2010.75.033

Tetenbaum-Novatt, J., L.E. Hough, R. Mironska, A.S. McKenney, and M.P. Rout. 2012. Nucleocytoplasmic transport: a role for nonspecific competition in karyopherin-nucleoporin interactions. *Mol. Cell. Proteomics*. 11:31–46. http://dx.doi.org/10.1074/mcp.M111.013656

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B.* 58:267–288.

Timney, B.L., J. Tetenbaum-Novatt, D.S. Agate, R. Williams, W. Zhang, B.T. Chait, and M.P. Rout. 2006. Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo. *J. Cell Biol.* 175:579–593. http://dx.doi.org/10.1083/jcb.200608141

Tran, E.J., M.C. King, and A.H. Corbett. 2014. Macromolecular transport between the nucleus and the cytoplasm: Advances in mechanism and emerging links to disease. *Biochim. Biophys. Acta.* 1843:2784–2795. http://dx.doi.org/10.1016/j.bbamcr.2014.08.003

Tu, L.C., G. Fu, A. Zilman, and S.M. Musser. 2013. Large cargo transport by nuclear pores: implications for the spatial organization of FG-nucleoporins. *EMBO J.* 32:3220–3230. http://dx.doi.org/10.1038/emboj.2013.239

van der Maarel, J.R.C. 2008. Introduction to Biopolymer Physics. World Scientific, Singapore. 247 pp.

Volkov, V.V., and D.I. Svergun. 2003. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Cryst.* 36:860–864. http://dx.doi.org/10.1107/S0021889803000268

Vovk, A., C. Gu, M.G. Opferman, L.E. Kapinos, R.Y. Lim, R.D. Coalson, D. Jasnow, and A. Zilman. 2016. Simple biophysics underpins collective conformations of the intrinsically disordered proteins of the Nuclear Pore Complex. *eLife.* 5:e10785. http://dx.doi.org/10.7554/eLife.10785

Wang, R., and M.G. Brattain. 2007. The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa. *FEBS Lett.* 581:3164–3170. http://dx.doi.org/10.1016/j.febslet.2007.05.082

Winey, M., D. Yarar, T.H. Giddings Jr., and D.N. Mastronarde. 1997. Nuclear pore complex number and distribution throughout the *Saccharomyces cerevisiae* cell cycle by three-dimensional reconstruction from electron micrographs of nuclear envelopes. *Mol. Biol. Cell.* 8:2119–2132. http://dx.doi.org/10.1091/mbc.8.11.2119

Wühr, M., T. Güttler, L. Peshkin, G.C. McAlister, M. Sonnett, K. Ishihara, A.C. Groen, M. Presler, B.K. Erickson, T.J. Mitchison, et al. 2015. The nuclear proteome of a vertebrate. *Curr. Biol.* 25:2663–2671. http://dx.doi.org/10.1016/j.cub.2015.08.047

Yamada, J., J.L. Phillips, S. Patel, G. Goldfien, A. Calestagne-Morelli, H. Huang, R. Reza, J. Acheson, V.V. Krishnan, S. Newsam, et al. 2010. A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol. Cell. Proteomics.* 9:2205–2224. http://dx.doi.org/10.1074/mcp.M000035-MCP201

Yang, Q., M.P. Rout, and C.W. Akey. 1998. Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Mol. Cell.* 1:223–234. http://dx.doi.org/10.1016/S1097-2765(00)80023-4

Zahn, R., D. Osmanović, S. Ehret, C. Araya Callis, S. Frey, M. Stewart, C. You, D. Görlich, B.W. Hoogenboom, and R.P. Richter. 2016. A physical model describing the interaction of nuclear transport receptors with FG nucleoporin domain assemblies. *eLife.* 5:5. http://dx.doi.org/10.7554/eLife.14119

Zenklusen, D., A.L. Wells, J.S. Condeelis, and R.H. Singer. 2007. Imaging real-time gene expression in living yeast. CSH *Prot.* 2007:pdb.prot4870.

Zilman, A., S. Di Talia, B.T. Chait, M.P. Rout, and M.O. Magnasco. 2007. Efficiency, selectivity, and robustness of nucleocytoplasmic transport. *PLOS Comput. Biol.* 3:e125. http://dx.doi.org/10.1371/journal.pcbi.0030125

Zilman, A., S. Di Talia, T. Jovanovic-Talisman, B.T. Chait, M.P. Rout, and M.O. Magnasco. 2010. Enhancement of transport selectivity through nano-channels by non-specific competition. *PLOS Comput. Biol.* 6:e1000804. http://dx.doi.org/10.1371/journal.pcbi.1000804

Timney et al., http://dx.doi.org/10.1083/jcb.201601004

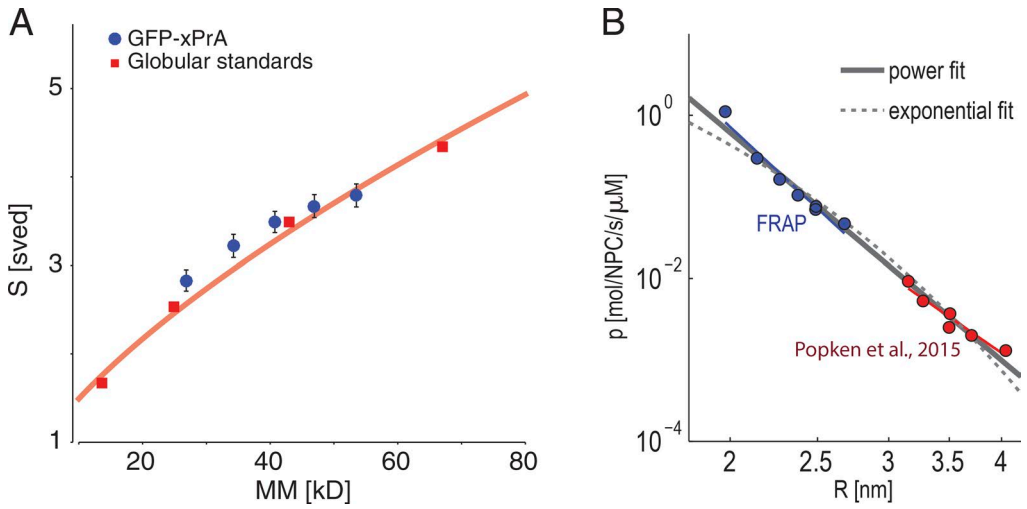


Figure S1. **Supporting analyses of shape effects.** (A) Sedimentation values for commercial globular protein standards measured on the same sucrose gradients as GFP-xPrA (± SEM). A fit of the S values of the standards to $c \cdot MW^{2/3}$ is shown in light red, to illustrate the trend in S versus molecular mass. According to the theory of sedimentation, $S \sim MW^{2/3}$ for ideal globular proteins (Erickson, 2009). (B) Passive permeability plotted as a function of theoretical minimal molecular radius of each substrate molecule in log-linear scale for our FRAP dataset (blue circles) and the dataset of the extrapolated measurements based on Popken et al. (2015), as described in Materials and methods (red circles). Note that R is not the actual radius, but a direct function proportional to the cubed square root of molecular mass. The gray solid line and gray dashed line indicate a fit to a power function and an exponential function, respectively, over the joint dataset. The blue and red solid lines indicate the fits to power functions for either our FRAP dataset or the Popken et al. (2015) dataset.

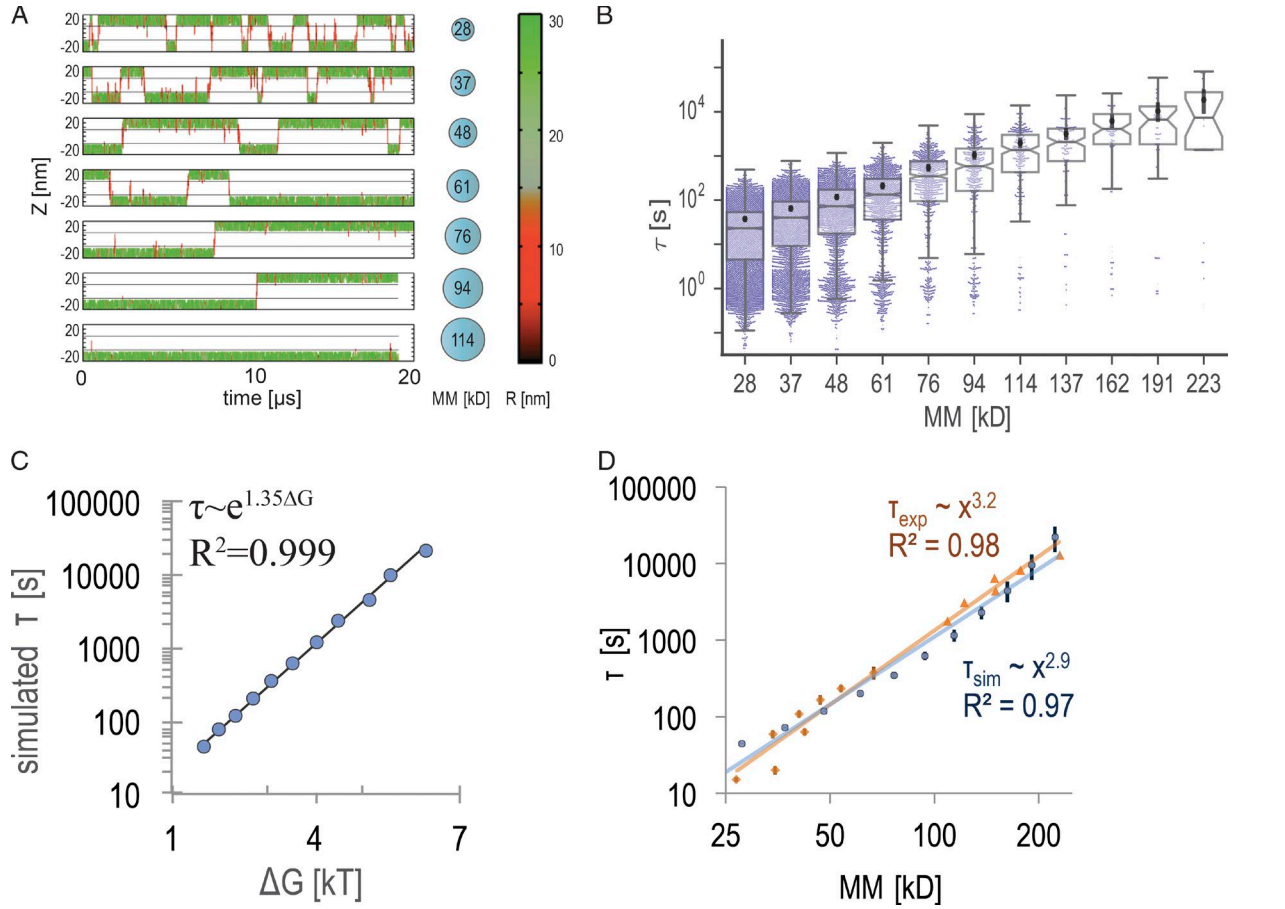


Figure S2. **Thermodynamics and kinetics of simulated passive diffusion (related to Fig. 4).** (A) Tracking of translocation between the nucleus and the cytoplasm (z axis) over time for individual macromolecules of different molecular masses in a sample simulation trajectory. The line color indicates the distance R between the substrate and the NPC central axis. (B) Distribution of normalized transport times. The swarm of purple dots indicates individual transport events observed during the simulations. Mean transport times and 95% confidence intervals around the mean are indicated by black filled circles and black lines. The boxes indicate 25th, 50th, and 75th percentiles of transport times; notches indicate 95% confidence intervals around the median; box whiskers cover data points within up to 1.5 times the interquartile range. (C) Simulated transport times as in Fig. 4 B as a function of the magnitude of the energy barrier shown in Fig. 4 F (bottom) follows a slightly modified form of the Eyring-Polanyi equation, with energy units scaled by an empirical 1.35 factor. (D) Simulated (blue) and experimental (orange) mean transport times plotted ± 95% confidence intervals as a function of molecular mass as in Fig. 4 B, for a larger channel of radius 20 nm and height 30 nm. The y axis is normalized by a single constant, which is applied uniformly to all substrates.

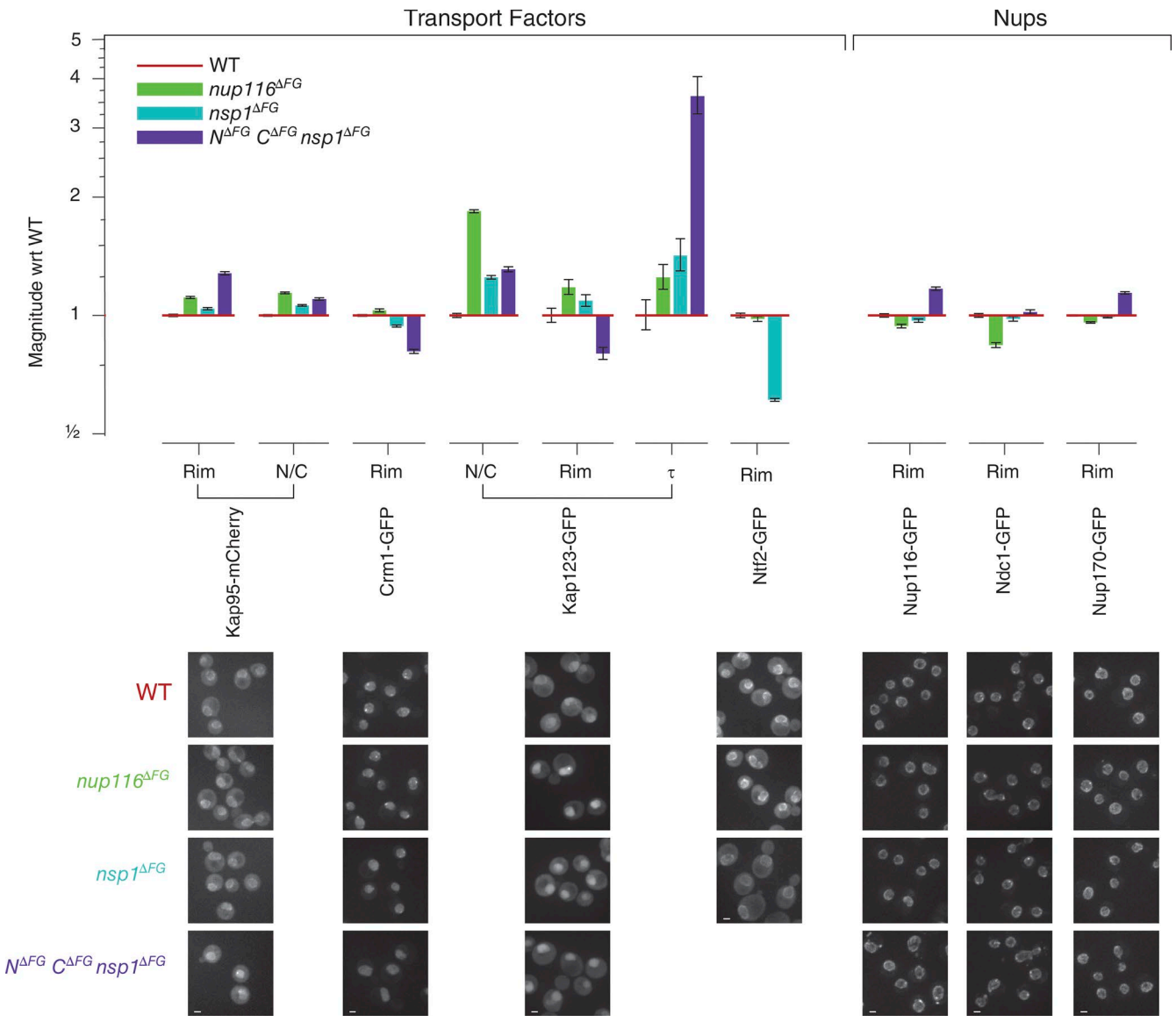


Figure S3. **Effect on the transport machinery in FG deletion strains.** Genomic copies of several transport receptors and nups were tagged with fluorescent proteins (GFP or mCherry) in WT and three FG deletion strains. Representative fields of live cells from each strain expressing each reporter are shown in the bottom panel, as maximum projections of a complete z axis image series. For each column of images, acquired within minutes of each other, exposure and contrast settings are identical. Bars, 1 μm. Measurements of the steady-state N/C fluorescence ratio or nuclear rim fluorescence (Rim) were computed from many cells and plotted in the top panel as the ratio of each measurement to the WT value (shown by the red line) in log scale. Error bars show SEM for each measurement (the error in WT can be seen as the error bars on the zero-height bar that precedes each series of three mutant measurements). Also shown is the time constant for the exponential recovery of Kap123-GFP fluorescence measured in FRAP experiments (see also Fig. S4). wrt, with regard to.

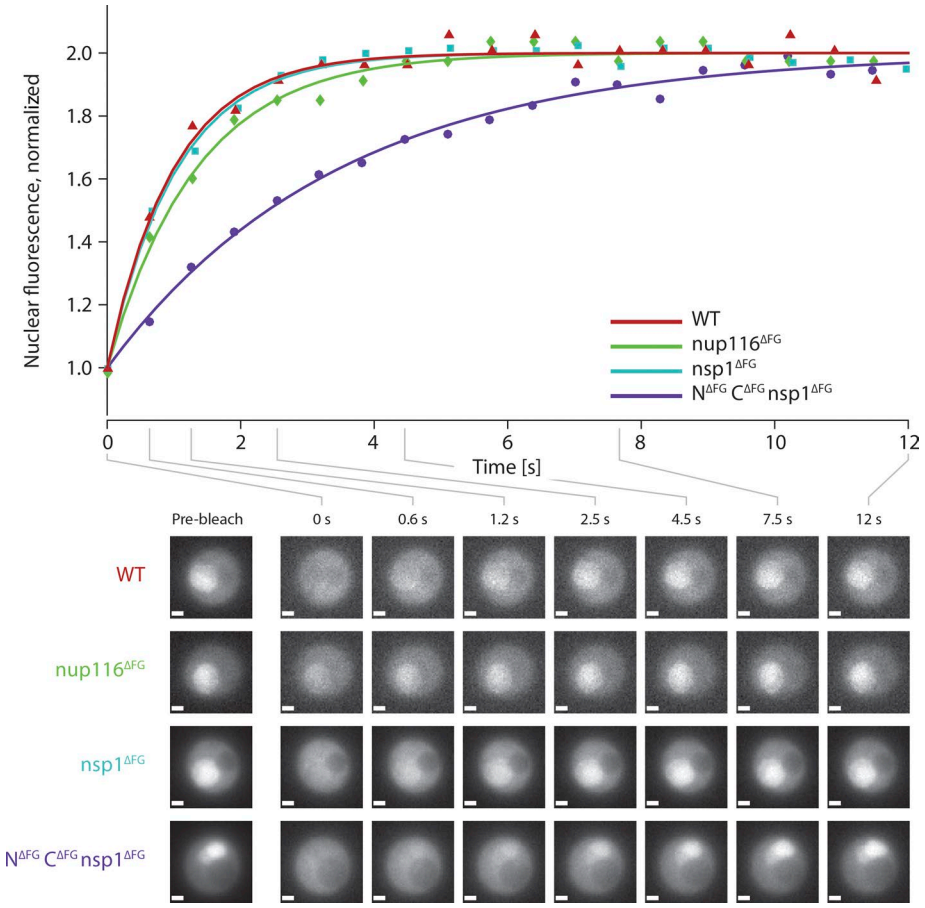


Figure S4. **Kap123 dynamics seem independent of passive NPC permeability, but are greatly reduced upon deletion of particular FG regions.** Dynamic exchange of Kap123-GFP in WT and three mutant strains was measured using FRAP, using identical methodology as for our measurements of passive GFP-xPrA transport. Nuclear fluorescence images of representative single-cell assays are shown (bottom). Bars, 1 µm. Nuclear fluorescence from these images was measured (top), fit to a single exponential relationship, and is plotted normalized between 1 ($t = 0$) and 2 ($t = \infty$).

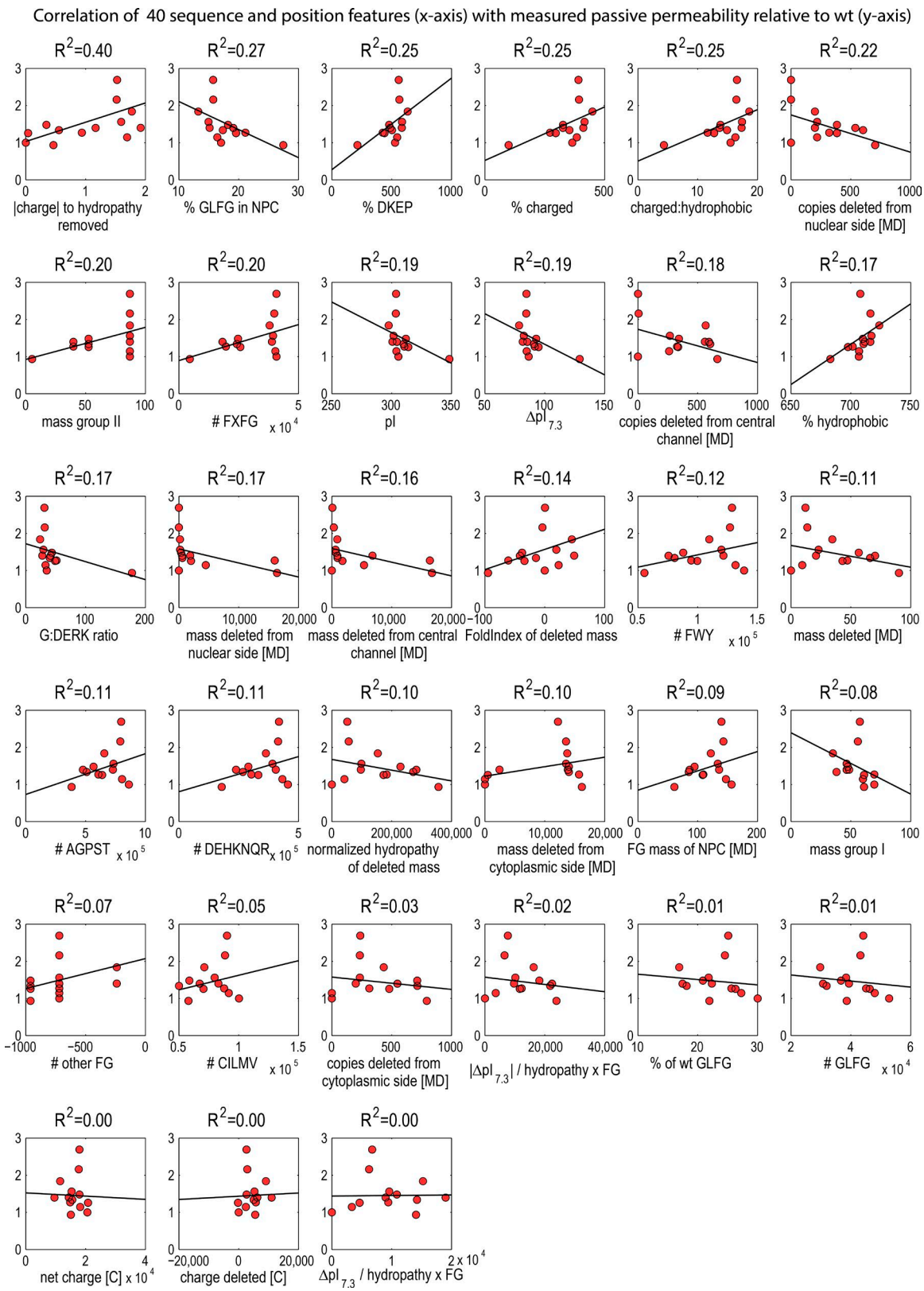


Figure S5. **Relative transport times plotted as a function of individual features of the mutant strains and the corresponding R^2 correlation coefficients.**

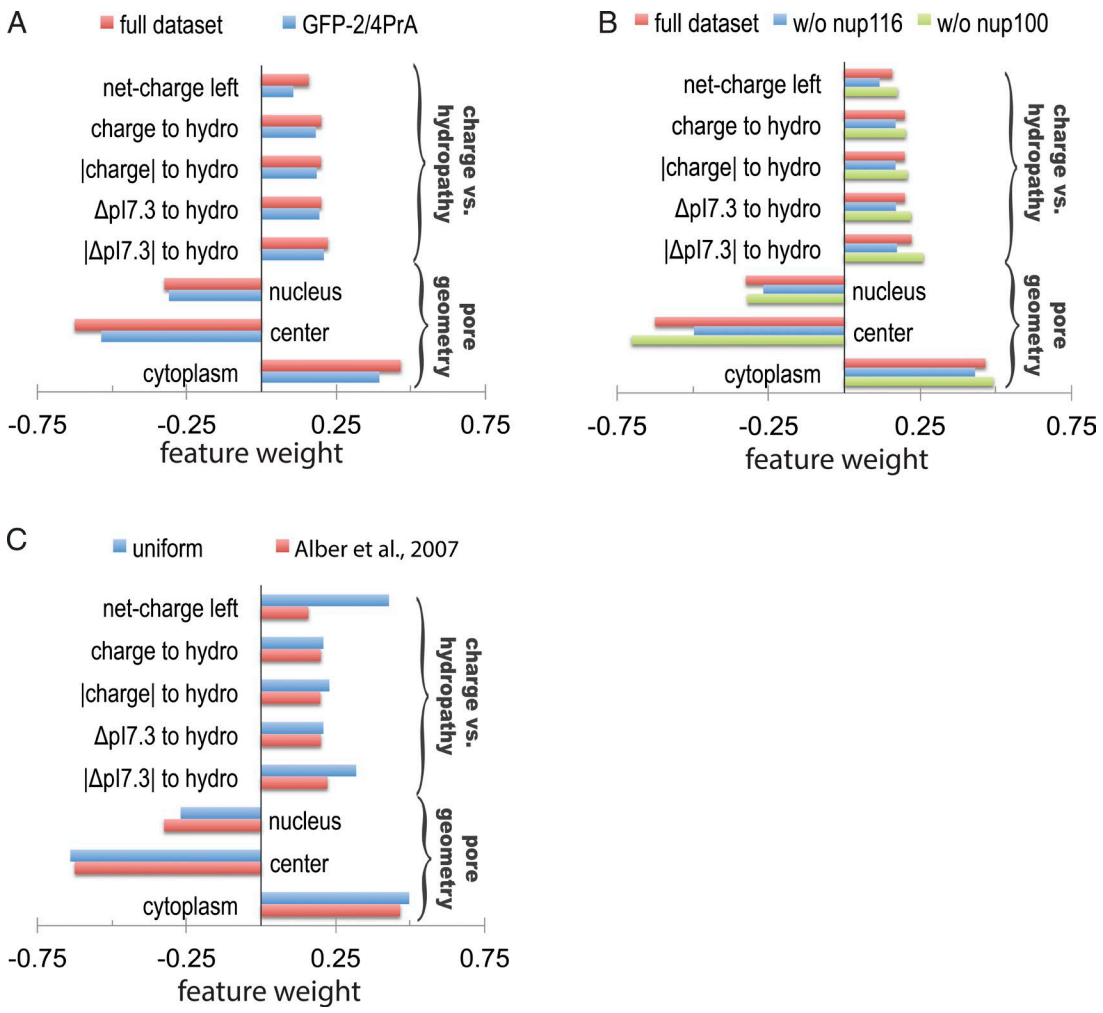


Figure S6. **Robustness of positive and negative contribution of various FG nup features in the combined PCA/LASSO regression model of passive transport in different settings.** (A) Feature contributions in regression model that was trained using either the full dataset of a partial dataset that includes only experiments with GFP-2PrA and GFP-4PrA cargos. (B) Feature contributions in regression model that was trained using either the full dataset of strains or all strains excluding the Nup116 deletion strain or excluding the Nup100 strain (green). (C) Feature contributions when FG nups are assume to follow either the uniform stoichiometry where each FG nup is assigned an equal copy number or the Alber et al. (2007a,b) stoichiometry of FG nups, which was used in all other analyses in this manuscript.

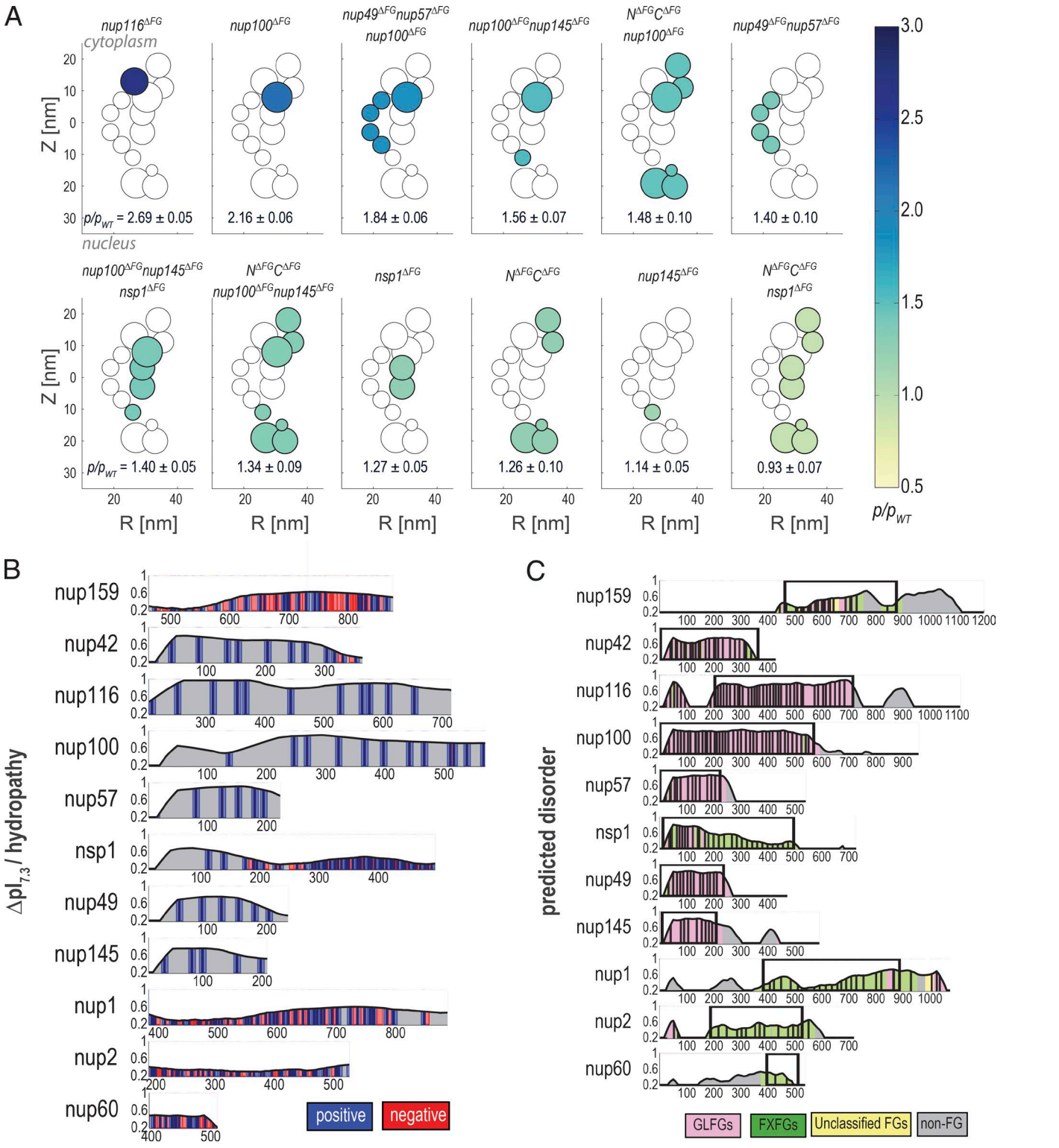


Figure S7. **Sequence and spatial analysis of FG nups and mutant strains.** (A) The relation between spatial position of FG nups and the effect of their deletion on permeability. The spatial positions of the various FG nups are shown as in Fig. 5 A. The deleted FG nups are shaded and color-coded according to their measured permeability relative to the WT strain, from yellow (p/p_{WT} = 0.5) to dark blue (p/p_{WT} = 3.0). The mutations that cause the most leakiness are relatively close to the cytoplasmic side, although on its own, this factor is not sufficient to predict the effect on passive permeability (see Fig. 7 A). (B) The ratio between the deviation of the pI from 7.3 (a proxy for the local net charge) and the hydropathy level for the deleted domains of the different FG nup sequences. All values are smoothed using a Gaussian kernel with window size of 100 residues and normalized between 0 and 1. (C) Disorder prediction for different FG nup sequences according to the PONDR-FIT disorder prediction tool. Deleted FG repeat domains are indicated by a rectangle. Different colors indicate FG motifs of different types, according to the color code at the bottom. All values are smoothed as in B values are normalized such that 0 mean equals probability for order and disorder (equivalent to 0.5 in raw PONDR-FIT score).

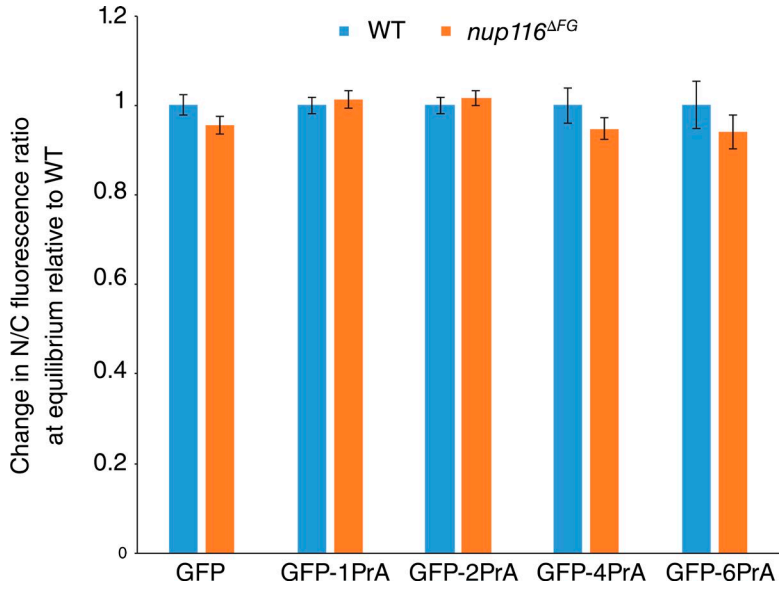
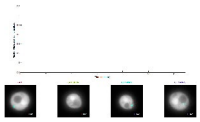
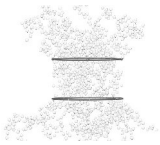


Figure S8. **The change in the ratio between nuclear and cytoplasmic fluorescence at equilibrium relative to the WT strain for the different GFP-xPrA substrates in the *nup116*$^{\Delta FG}$ strain.** 95% confidence intervals indicated for both the WT and the *nup116*$^{\Delta FG}$ strains. Note that the absolute N/C fluorescence ratio is ~1.3 for all strains and cargos. The slightly depressed cytoplasmic fluorescence is likely a result of a significant volume of small cytoplasmic vacuoles and vesicles that are nonpermeable to the labeled substrate, but that are smaller than the optic resolution of our microscope. This may result in 20–25% reduction in the measured cytoplasmic fluorescence relative to the nucleus.



Video 1. **FRAP permeability assays of four representative WT cells, each expressing a different-sized GFP-xPrA substrate, were combined and synchronized to create this video, together with a plot of their normalized nuclear fluorescence decay curves.**



Video 2. **Sample simulation trajectory of passive diffusion of macromolecules of different molecular mass/radii through the NPC, using the model described in the text and in Fig. 4 and Fig. S2.** The FG nups are shown as strings of gray spheres.

Table S1. **Plasmids used in this study**

Descriptive name	Formal name[a]	Details	Source
pYX242	pYX242	Continuous high-expression yeast plasmid, TPI promotor, LEU2 selectable	EMD Millipore (discontinued)
GFP	pBT055	eGFP in pYX242	This study
GFP-1PrA	pBT056	GFP with one repeat of PrA in pYX242	This study
GFP-2PrA	pBT057	GFP with two repeats of PrA in pYX242	This study
GFP-3PrA	pBT058	GFP with three repeats of PrA in pYX242	This study
GFP-4PrA	pBT059	GFP with four repeats of PrA in pYX242	This study
GFP-6PrA	pBT060	GFP with six repeats of PrA in pYX242	This study
GFP-1PrG	pBT067	GFP with C2 PrG repeat in pYX242	This study
GFP-2PrG	pBT068	GFP with C2 and C3 PrG repeats in pYX242	This study
pET28a	pET28a	Bacterial N-terminal 6xHIS expression plasmid	EMD Millipore
HIS-GFP	HIS-GFP	GFP subcloned into pET28a	This study
HIS-GFP-1PrA	pBT061	GFP-1PrA subcloned into pET28a	This study
HIS-GFP-2PrA	pBT062	GFP-2PrA subcloned into pET28a	This study
HIS-GFP-3PrA	pBT063	GFP-3PrA subcloned into pET28a	This study
HIS-GFP-4PrA	pBT064	GFP-4PrA subcloned into pET28a	This study
HIS-GFP-6PrA	pBT065	GFP-6PrA subcloned into pET28a	This study
HIS-GFP-1PrG	pBT069	GFP-1PrG subcloned into pET28a	This study
HIS-GFP-2PRG	pBT070	GFP-2PrG subcloned into pET28a	This study

[a]For requests.

Table S2. **Strains used in this study**

Descriptive name	Formal name[a]	Description	Genotype
WT	BTY2283	Htb2-mCherry–tagged version of SWY2283[b] (FG deletion WT background strain)	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HTB2-mCherry::HIS5
nup100$^{\Delta FG}$	BTY2762	Htb2-mCherry–tagged version of SWY2762[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-Loxp-Nup100$^{\Delta}$GLFG HTB2-mCherry::HIS5
nup116$^{\Delta FG}$	BTY2790	Htb2-mCherry–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG HTB2-mCherry::HIS5
nup145$^{\Delta FG}$	BTY2869	Htb2-mCherry–tagged version of SWY2869[b]	MATa ade2-1::ADE2 trp1-1 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 myc-loxP-nup145$^{\Delta}$GLFG HTB2-mCherry::HIS5
nup49$^{\Delta FG}$ nup57$^{\Delta FG}$	BTY2883	Htb2-mCherry–tagged version of SWY2883[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 myc-LoxP-nup57$^{\Delta}$GLFG T7-LoxP-nup49$^{\Delta}$GLFG HTB2-mCherry::HIS5
nup49$^{\Delta FG}$ nup57$^{\Delta FG}$ nup100$^{\Delta FG}$	BTY2871	Htb2-mCherry–tagged version of SWY2871[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 myc-LoxP-nup57$^{\Delta}$GLFG T7-LoxP-nup49$^{\Delta}$GLFG HA-LoxP-nup100$^{\Delta}$GLFG HTB2-mCherry::HIS5
nsp1$^{\Delta FG}$	BTY2920	Htb2-mCherry–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG HTB2-mCherry::HIS5
nup100$^{\Delta FG}$ nup145$^{\Delta FG}$	BTY2982	Htb2-mCherry–tagged version of SWY2982[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 myc-LoxP-nup145$^{\Delta}$GLFG HA-Loxp-Nup100$^{\Delta}$GLFG HTB2-mCherry::HIS5
nup100$^{\Delta FG}$ nup145$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY2981	Htb2-mCherry–tagged version of SWY2981[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 myc-LoxP-nup145$^{\Delta}$GLFG HA-Loxp-Nup100$^{\Delta}$GLFG flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG HTB2-mCherry::HIS5
N$^{\Delta FG}$ C$^{\Delta FG}$	BTY3041	Htb2-mCherry–tagged version of SWY3041[b] (All peripheral nuclear and cytoplasmic nups truncated - Nup1, Nup60, Nup2, Nup42, and Nup159)	MATa ade2-1::ADE2 trp1-1 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60$^{\Delta}$FxF HTB2-mCherry::HIS5
N$^{\Delta FG}$ C$^{\Delta FG}$ nup100$^{\Delta FG}$	BTY3042	Htb2-mCherry–tagged version of SWY3042[b]	MATa ade2-1::ADE2 trp1-1 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60$^{\Delta}$FxF HA-loxP-nup100$^{\Delta}$GLFG + pRS314 (TRP1 CEN) HTB2-mCherry::HIS5
N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY3062	Htb2-mCherry–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG HTB2-mCherry::HIS5
N$^{\Delta FG}$ C$^{\Delta FG}$ nup100$^{\Delta FG}$ nup145$^{\Delta FG}$	BTY3292	Htb2-mCherry–tagged version of SWY3292[c]	MATa ade2-1::ADE2 trp1-1 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60$^{\Delta}$FxF HA-loxP-nup100$^{\Delta}$GLFG myc-loxP-nup145$^{\Delta}$GLFG HTB2-mCherry::HIS5
Kap95-mCherry WT	BTY038	Kap95-mCherry–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 Kap95-mCherry::HIS3
Kap95-mCherry nup116$^{\Delta FG}$	BTY039	Kap95-mCherry–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG Kap95-mCherry::HIS3
Kap95-mCherry nsp1$^{\Delta FG}$	BTY040	Kap95-mCherry–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Kap95-mCherry::HIS3
Kap95-mCherry N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY041	Kap95-mCherry–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Kap95-mCherry::HIS3
Crm1-GFP WT	BTY042	Crm1-GFP–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 Crm1-GFP::HIS3MX6
Crm1-GFP nup116$^{\Delta FG}$	BTY043	Crm1-GFP–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG Crm1-GFP::HIS3MX6
Crm1-GFP nsp1$^{\Delta FG}$	BTY044	Crm1-GFP–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Crm1-GFP::HIS3MX6
Crm1-GFP N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY045	Crm1-GFP–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Crm1-GFP::HIS3MX6
Kap123-GFP WT	BTY046	Kap123-mCherry–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 KAP123-GFP::HIS3MX6
Kap123-GFP nup116$^{\Delta FG}$	BTY047	Kap123-mCherry–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG KAP123-GFP::HIS3MX6
Kap123-GFP nsp1$^{\Delta FG}$	BTY048	Kap123-mCherry–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG KAP123-GFP::HIS3MX6
Kap123-GFP N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY049	Kap123-mCherry–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Kap123-GFP::HIS3MX6

Table S2. **Strains used in this study** *(Continued)*

Descriptive name	Formal name[a]	Description	Genotype
Ntf2-GFP WT	BTY050	Ntf2-GFP–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 Ntf2-GFP::HIS3MX6
Ntf2-GFP nup116$^{\Delta FG}$	BTY051	Ntf2-GFP–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG NTF2-GFP::HIS3MX6
Ntf2-GFP nsp1$^{\Delta FG}$	BTY052	Ntf2-GFP–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Ntf2-GFP::HIS3MX6
Nup116-GFP WT	BTY053	Nup116-GFP–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 NUP116-GFP::HIS3MX6
nup116$^{\Delta FG}$-GFP	BTY054	Truncated version of nup116 in SWY2790[b] was GFP tagged	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG-GFP::HIS3MX6
Nup116-GFP nsp1$^{\Delta FG}$	BTY055	Nup116-GFP–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG NUP116-GFP::HIS3MX6
Nup116-GFP N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY056	Nup116-GFP–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Nup116-GFP::HIS3MX6
Ndc1-GFP WT	BTY057	Ndc1-GFP–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 Ndc1-GFP::HIS3MX6
Ndc1-GFP nup116$^{\Delta FG}$	BTY058	Ndc1-GFP–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG Ndc1-GFP::HIS3MX6
Ndc1-GFP nsp1$^{\Delta FG}$	BTY059	Ndc1-GFP–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Ndc1-GFP::HIS3MX6
Ndc1-GFP N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY060	Ndc1-GFP–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Ndc1-GFP::HIS3MX6
Nup170-GFP WT	BTY061	Nup170-GFP–tagged version of SWY2283[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 Nup170-GFP::HIS3MX6
Nup170-GFP nup116$^{\Delta FG}$	BTY062	Nup170-GFP–tagged version of SWY2790[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 T7-LoxP-Nup116$^{\Delta}$GLFG Nup170-GFP::HIS3MX6
Nup170-GFP nsp1$^{\Delta FG}$	BTY063	Nup170-GFP–tagged version of SWY2920[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Nup170-GFP::HIS3MX6
Nup170-GFP N$^{\Delta FG}$ C$^{\Delta FG}$ nsp1$^{\Delta FG}$	BTY064	Nup170-GFP–tagged version of SWY3062[b]	MATa ade2-1::ADE2 ura3-1 leu2-3,112 his3-11,15 lys2 can1-100 HA-LoxP-nup42$^{\Delta}$FG myc-LoxP-nup159$^{\Delta}$FG T7-LoxP-nup1$^{\Delta}$FxFG myc-LoxP-nup2$^{\Delta}$FxFG myc-LoxP-nup60FxF flag-LoxP-nsp1$^{\Delta}$FG$^{\Delta}$FxFG Nup170-GFP::HIS3MX6

[a]For requests.
[b]Strawn et al., 2004.
[c]Terry et al., 2007.

Table S3. .**Features used to describe the NPCs of mutant strains and for the regression analysis**

No.	Feature name	Description
	Mass features	
1	FG mass of NPC (MD)	Combined mass of all domains of FG repeats in the NPC
2	Mass deleted (MD)	Total mass of FG repeats deleted from the NPC
	NPC geometry	
3	Mass deleted from nuclear side (MD)	Total mass of FG repeat domains deleted from the nuclear end of the NPC, considering the stoichiometry of the NPC, and based on the positions of the anchor domains of FG nups reported in the architectural map of the NPC (Alber et al., 2007b)
4	Mass deleted from central channel (MD)	Total mass of FG repeat domains deleted from the central channel of the NPC, considering the stoichiometry of the NPC, and based on the positions of the anchor domains of FG nups reported in the architectural map of the NPC (Alber et al., 2007b)
5	Mass deleted from cytoplasmic side (MD)	Total mass of FG repeat domains deleted from the cytoplasmic end of the NPC, considering the stoichiometry of the NPC, and based on the positions of the anchor domains of FG nups reported in the architectural map of the NPC (Alber et al., 2007b)
6	Copies deleted from nuclear side	Total number of copies of FG nups deleted from the nuclear end of the NPC, considering the stoichiometry of the NPC, and based on the positions of the anchor domains of FG nups reported in the architectural map of the NPC (Alber et al., 2007b)
7	Copies deleted from central channel	Total number of copies of FG nups deleted from the central channel of the NPC, considering the stoichiometry of the NPC, and based on the positions of the anchor domains of FG nups reported in the architectural map of the NPC (Alber et al., 2007b)
8	Copies deleted from cytoplasmic side	Total number of copies of FG nups deleted from the cytoplasmic end of the NPC, considering the stoichiometry of the NPC, and based on the positions of the anchor domains of FG nups reported in the architectural map of the NPC (Alber et al., 2007b)
	Charge and hydrophobicity	
9	Net charge (C)	The net charge of all domains of FG repeats at the NPC, computed as the sum of the charges of each deleted domain as calculated by the MatLab isoelectric function, and weighted by the copy number of each domain
10	Charge deleted (C)	The total charge deleted from the NPC, computed as the sum of the charges of each deleted domain as calculated by the MatLab isoelectric function, and weighted by the copy number of each domain
11	pI	The pI of the FG repeat domains at the NPC, averaged over the individual pI of each domain as computed by the MatLab isoelectric() function and weighted by their estimated copy numbers
12	Delta-pI7.3	The difference between the average pI of the NPC (as defined above) and estimated pH of 7.3 in the cytoplasm and nucleus of yeasts. The actual pH value within yeast cells fluctuates, but is likely close to this value (Orij et al., 2009, 2011). Positive values are correlated with a positive net charge.
13	[Delta pI7.3]/hydropathyxFG	The sum of the ratio between [pI-7.3] to the hydropathy score (Kyte and Doolittle, 1982) for each FG motif that remains in the NPC. For each FG motif, the pI and the hydropathy scores are computed over a window of 150 residues, truncated at the termini of the sequence if needed,
14	|Delta pI7.3|/hydropathyxFG	The sum of the ratio between |pI-7.3| to the hydropathy score (Kyte and Doolittle, 1982) for each FG motif that remains in the NPC. For each FG motif, the pI and the hydropathy scores are computed over a window of 150 residues, truncated at the termini of the sequence if needed,
15	|Charge| to hydropathy removed	The ratio between the absolute net charge and the sum of hydropathy scores of all residues removed from the NPC
16	Percent charged	Percent of amino acids that are charged (D/E/R/K) in the FG repeat domains that remain at the NPC
17	Percent hydrophobic	Percent of hydrophobic residues (A/I/L/F/W/V) in the FG repeat domains that remain at the NPC
18	Charged/hydrophobic	Number of charged residues per hydrophobic residues (A/I/L/F/W/V) in the FG repeat domains that remain at the NPC
19	Normalized hydropathy of deleted mass	The sum of per-residue hydropathy scores (Kyte and Doolittle, 1982) of all residues deleted from the WT NPC, weighted by their estimated stoichiometry
	Predicted protein disorder	
20	Fold Index of deleted mass	The Fold Index score (Prilusky et al., 2005) of the residues of deleted FG repeats, weighted by their estimated stoichiometry, computed as 2.785 × hydro-abs(charge) − 1.15 where hydro is the mean Kyte-Doolitle hydropathy score (Kyte and Doolittle, 1982) and charge is the net-charge of all residues divided by the number of residues.

Table S3. .**Features used to describe the NPCs of mutant strains and for the regression analysis** *(Continued)*

No.	Feature name	Description
21	No. of AGPST	Number of A/G/P/S/T residues in the NPC, suggested to affect protein disorder (Weathers et al., 2004)
22	No. of CILMV	Number of C/I/L/M/V residues in all domains of FG repeats in the NPC, suggested to affect protein disorder (Weathers et al., 2004)
23	No. of DEHKNQR	Number of D/E/H/K/N/Q/R residues in the NPC, suggested to affect protein disorder (Weathers et al., 2004)
24	No. of FWY	Number of aromatic residues in the NPC, suggested to affect protein disorder (Weathers et al., 2004)
	FG motifs flavors	
25	No. of GLFG	Number of distinct GLFG-flavored motifs, matching the regular expression "[AGK MNQST][AGILMPSTV]FG" in all domains of FG repeats in the NPC
26	No. of FXFG	Number of distinct FXFG-flavored motifs, fitting the regular expression "[LPFI].FG" or "F.F[^G]" in all domains of FG repeats in the NPC
27	No. of other FG	Number of distinct non-GLFG and non-FXFG flavored motifs in all domains of FG repeats in the NPC
28	Percent GLFG in NPC	Percent of GLFG-flavored FG motifs from all FG motifs in all domains of FG repeats in the NPC
29	Percent of WT GLFG	GLFG-flavored FG motifs remaining in all domains of FG repeats in the NPC as a percentage of the WT
	General sequence features of FG repeat domains	
30	G/DERK ratio	Ratio of G to D/E/R/K amino acids in all domains of FG repeats in the NPC. This ratio was suggested to separate flavors of FG domains (DeGrasse et al., 2009)
31	Mass–group I of G/DERK ratio (MD)	Mass of Fall domains of FG repeats in the NPC with G/DERK ratio of group I in (DeGrasse et al., 2009)
32	Mass–group II of G/DERK ratio (MD)	Mass of Fall domains of FG repeats in the NPC with G/DERK ratio of group II in (DeGrasse et al., 2009)
33	Percent DKEP	Percent of D/K/E/P residues in the FG repeat domains that remain at the NPC; these were indicated to affect transport by increasing the radius of gyration of certain FG repeat domains (Yamada et al., 2010)

Table S4. **SAXS maximal end-to-end distances and radii of gyration for GFP-xPrA and GFP-xPrG molecules**

	Molecular mass	*Dmax*	*Rg*
	kD	*nm*	*nm*
GFP-HIS	26.8	8.1	2.5
GFP-1PrA	34.2	10.3	3.1
GFP-2PrA	40.7	11	3.3
GFP-3PrA	46.8	12.3	3.7
GFP-4PrA	53.7	12.9	3.9
GFP-6PrA	66.8	14.1	4.3
GFP-1PrG	34.7	10.6	3.2
GFP-2PrG	42.2	11.5	3.5

References

Alber, F., S. Dokudovskaya, L.M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B.T. Chait, et al. 2007b. The molecular architecture of the nuclear pore complex. *Nature.* 450:695–701. http://dx.doi.org/10.1038/nature06405

Alber, F., S. Dokudovskaya, L.M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B.T. Chait, et al. 2007a. Determining the architectures of macromolecular assemblies. *Nature*. 450:683–694. http://dx.doi.org/10.1038/nature06404

DeGrasse, J.A., K.N. DuBois, D. Devos, T.N. Siegel, A. Sali, M.C. Field, M.P. Rout, and B.T. Chait. 2009. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Mol. Cell. Proteomics.* 8:2119–2130. http://dx.doi.org/10.1074/mcp.M900038-MCP200

Erickson, H.P. 2009. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proced. Online.* 11:32–51. http://dx.doi.org/10.1007/s12575-009-9008-x

Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132. http://dx.doi.org/10.1016/0022-2836(82)90515-0

Orij, R., J. Postmus, A. Ter Beek, S. Brul, and G.J. Smits. 2009. In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in *Saccharomyces cerevisiae* reveals a relation between intracellular pH and growth. *Microbiology.* 155:268–278. http://dx.doi.org/10.1099/mic.0.022038-0

Orij, R., S. Brul, and G.J. Smits. 2011. Intracellular pH is a tightly controlled signal in yeast. *Biochim. Biophys. Acta.* 1810:933–944. http://dx.doi.org/10.1016/j.bbagen.2011.03.011

Popken, P., A. Ghavami, P.R. Onck, B. Poolman, and L.M. Veenhoff. 2015. Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex. *Mol. Biol. Cell.* 26:1386–1394. http://dx.doi.org/10.1091/mbc.E14-07-1175

Prilusky, J., C.E. Felder, T. Zeev-Ben-Mordehai, E.H. Rydberg, O. Man, J.S. Beckmann, I. Silman, and J.L. Sussman. 2005. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics.* 21:3435–3438. http://dx.doi.org/10.1093/bioinformatics/bti537

Strawn, L.A., T. Shen, N. Shulga, D.S. Goldfarb, and S.R. Wente. 2004. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat. Cell Biol.* 6:197–206. http://dx.doi.org/10.1038/ncb1097

Terry, L.J., and S.R. Wente. 2007. Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. *J. Cell Biol.* 178:1121–1132. http://dx.doi.org/10.1083/jcb.200704174

Weathers, E.A., M.E. Paulaitis, T.B. Woolf, and J.H. Hoh. 2004. Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS Lett.* 576:348–352. http://dx.doi.org/10.1016/j.febslet.2004.09.036

Yamada, J., J.L. Phillips, S. Patel, G. Goldfien, A. Calestagne-Morelli, H. Huang, R. Reza, J. Acheson, V.V. Krishnan, S. Newsam, et al. 2010. A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol. Cell. Proteomics.* 9:2205–2224. http://dx.doi.org/10.1074/mcp.M000035-MCP201