

Charge as a Selection Criterion for Translocation through the Nuclear Pore Complex

Lucy J. Colwell¹, Michael P. Brenner^{1*}, Katharina Ribbeck^{2*}

1 School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, United States of America, **2** FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts, United States of America

Abstract

Nuclear pore complexes (NPCs) are highly selective filters that control the exchange of material between nucleus and cytoplasm. The principles that govern selective filtering by NPCs are not fully understood. Previous studies find that cellular proteins capable of fast translocation through NPCs (transport receptors) are characterized by a high proportion of hydrophobic surface regions. Our analysis finds that transport receptors and their complexes are also highly negatively charged. Moreover, NPC components that constitute the permeability barrier are positively charged. We estimate that electrostatic interactions between a transport receptor and the NPC result in an energy gain of several $k_B T$, which would enable significantly increased translocation rates of transport receptors relative to other cellular proteins. We suggest that negative charge is an essential criterion for selective passage through the NPC.

Citation: Colwell LJ, Brenner MP, Ribbeck K (2010) Charge as a Selection Criterion for Translocation through the Nuclear Pore Complex. *PLoS Comput Biol* 6(4): e1000747. doi:10.1371/journal.pcbi.1000747

Editor: Michael Gilson, University of Maryland, United States of America

Received: September 21, 2009; **Accepted:** March 18, 2010; **Published:** April 22, 2010

Copyright: © 2010 Colwell et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was funded by grants from Merck, the National Science Foundation Division of Mathematical Sciences, the Kavli Institute for Bionano Science and Technology at Harvard University and NIGMS grant GM068763 for National Centers of Systems Biology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: brenner@seas.harvard.edu (MPB); kribbeck@cgr.harvard.edu (KR)

Introduction

The defining feature of eukaryotic cells is the separation of nuclear and cytoplasmic compartments by the nuclear envelope. All nuclear proteins, for example polymerases and transcription factors, are made in the cytoplasm and imported into the nucleus. Conversely, RNAs that function in translation are made inside the nucleus and exported to the cytoplasm. The transport of material between the nucleus and cytoplasm occurs through nuclear pore complexes (NPCs), aqueous channels that are embedded in the nuclear envelope. Translocation through NPCs is fully reversible and uncoupled from NTP hydrolysis [1,2,3,4,5]. Kinetic measurements demonstrate that a single NPC can selectively translocate nearly one thousand molecules per second [6,7].

Proteins above 30–40 kDa typically only traverse the NPC at appreciable rates with the aid of dedicated nuclear transport receptors (for reviews see [8,9,10,11]). Nuclear transport receptors are soluble proteins that bind to their substrates and translocate together with them through the NPC channel. Importin β -like transport receptors mediate the vast majority of nuclear transport, and are typically classified as importins or exportins, depending on whether they mediate nuclear import or export. The importin β superfamily includes at least 21 members in the human proteome *Homo sapiens*, and 14 in the yeast *Saccharomyces cerevisiae* [8].

The NPC consists of approximately 30 proteins (termed nucleoporins) in *S. cerevisiae*, and roughly the same number in vertebrates [12,13,14,15,16]. Many nucleoporins contain a series of phenylalanine-rich repeats (FG-repeats), which typically occur within the amino acid motifs FxFG or GLFG (here “x” stands for a variable amino acid; [17], and references therein). The FG-

repeats are separated by largely unfolded and hydrophilic spacer sequences [17,18,19]. Immuno-electron microscopy data reveal that FG-repeat domains occupy the NPC channel as well as the cytoplasmic and nuclear rim of the NPC [19,20,21,22].

Several research groups have established that relatively specific interactions between transport receptors and FG-repeats within nucleoporins are necessary to facilitate selective translocation through the NPC barrier [23,24,25,26,27,28,29,30,31,32,33,34,35,36,37]. Transport receptor-FG-repeat interactions are mediated by specific hydrophobic regions on the surface of the transport receptors [6,24,25,26,28,29,32,33,38,39,40]. Indeed, the crystal structures of importin β [41], transportin [42], NTF2 [43,44], and TAP [45] reveal that their surfaces are characterized by a high proportion of hydrophobic regions.

Here, we observe that nuclear transport receptors carry more negative charge than the majority of cellular proteins. The influence of surface charge on the translocation reaction has not been addressed so far. We note that most components of the selectivity barrier within the NPC are characterized by net positive charge [17,19,34]. We calculate that electrostatic interactions between a negatively charged transport receptor and positively charged nucleoporins could result in an energy gain of multiple $k_B T$. This could help compensate for the energy barrier that transport receptors encounter on translocating through the spatially confined NPC channel. We propose that positively charged nucleoporin domains are an important component of the selective filter and promote the specific translocation of negatively charged transport receptors, while imposing a large energy barrier against translocation of positively charged cellular proteins.

Author Summary

All proteins that move between the cytoplasm and the nucleus must pass through nuclear pore complexes, large aqueous channels around 40nm in diameter. In some cases the nuclear envelope is perforated with several thousand nuclear pore complexes, while in other cases they are few and far between. Macromolecular transport through nuclear pores is highly regulated; an elaborate system, involving the binding and unbinding of accessory proteins (transport receptors), allows regulation of which proteins can pass through the pores. The basic principles that govern this selective filtering are not fully understood. Some proteins pass through the pore without binding to transport receptors, while others require the binding of multiple transport receptors for efficient translocation. How does the pore select which proteins can pass through, and which cannot? This paper carries out a biophysical analysis of the properties of proteins that can translocate through the nuclear pore. We find that proteins capable of fast translocation are highly negatively charged, whereas proteins that cannot pass through the pore are positively charged. Moreover, proteins that constitute the interior of the pore channel itself are net positively charged. This suggests that electrostatic interactions between translocating proteins and the pore are an essential part of the selective filtering mechanism.

Results

Translocation competent particles are hydrophobic and negatively charged

To understand the biophysical properties that facilitate NPC translocation we first compared the amino acid composition of translocation-competent particles (nuclear transport receptors and cognate transport receptor-cargo complexes) with that of cargo proteins. We assembled a collection of nuclear transport receptors and their cognate cargos from both *S. cerevisiae* and *H. sapiens* (Table S1). In addition, we compiled a list of biophysical properties including measures of charge, polarity, and 27 different empirical metrics for hydrophobicity (Table S2). For every protein, the value of each property (with the exception of the isoelectric point) is obtained by summing the contribution from each amino acid in its sequence, and normalizing by sequence length. Note that these metrics do not take into account the solvent accessibility of each residue; this is not feasible due to the lack of structural information for many cargo proteins. The results of this analysis are displayed as a heat map in Figure 1A. Every property is normalized over the entire set of proteins to have mean zero. Bright red and green correspond to 3 standard deviations above and below this mean, respectively. The proteins are clustered according to the similarity of their properties. Each column corresponds to a different property and each row to a different protein or protein complex. The individual rows in the top panel of Fig. 1A reveal that the profiles of individual nuclear transport receptors resemble each other, but are visibly different to the profiles of individual cargo proteins (middle panel, Fig. 1A). This suggests that particles with translocation-promoting properties are distinct from the average cargo protein with regard to the properties measured here.

To eliminate redundancy in the set of physical properties, we applied principal component analysis (PCA) to the 27 hydrophobicity scales. PCA transforms a number of correlated variables into a smaller number of uncorrelated variables called principal

components, ordered according to the amount of variability that they explain.

For the 27 hydrophobicity scales, the first principal component captures 75% of the total variance, and thus serves as an “aggregate” hydrophobicity metric that correlates with each of the individual hydrophobicity scales. Moreover, we find that polarity correlates strongly with the aggregate hydrophobicity scale (correlation coefficient $r^2 = -0.93$, Fig. S1), and so can be eliminated as an independent property. In summary, the original 29 dimensional property space could be reduced to just two properties, the aggregate hydrophobicity scale and isoelectric point. Since the NPC is an aqueous channel that is freely permeable to ions, we assume that the pH within the NPC channel is comparable to that in the cytoplasm and so report the net charge at pH 7.2 of each protein or protein complex, instead of its isoelectric point hence forth. The charge Q of a protein at pH 7.2 can be predicted from its amino acid sequence according to

$$Q = \sum_X \frac{10^{7.2 - \text{pK}_X}}{10^{7.2 - \text{pK}_X} + 1}$$

where the sum is over all ionized amino acids in the protein, and the pK_X values were obtained from the *emboss iep* application [46]. We note that representations of protein properties in terms of hydrophobicity and charge have been employed elsewhere [19].

Fig. 1B depicts the hydrophobicity and charge at pH 7.2 of proteins from the yeast proteome. This plot reveals that transport receptors have extreme properties compared to the majority of yeast proteins: Transport receptors and receptor-cargo complexes (light and dark blue squares, and purple triangles, respectively) are collectively more negatively charged at pH 7.2. In contrast, cargo proteins (red circles) tend to be positively charged. The amount of positive charge on a typical cargo protein is far greater than the five to ten charges conferred by a monopartite or bipartite nuclear localization signal [47]. In addition, transport receptors are more hydrophobic than most cellular proteins; as mentioned in the introduction several experiments suggest that specific hydrophobic regions on the surface of nuclear transport receptors are critical for translocation. Note that the segregation of transport receptors from their cargo proteins on the basis of charge and hydrophobicity is also apparent in *H. sapiens* (Fig. 1C and S2). Together, these data suggest that net negative charge at pH 7.2 is an evolutionarily conserved property that distinguishes translocation competent particles from cargo proteins.

Many nucleoporins that constitute the selectivity barrier are positively charged

To address the role of charge in the translocation reaction we next analyzed the amino acid sequences of nucleoporins from *S. cerevisiae* (Fig. 2A; Table S3). Upon clustering according to the similarity of their properties, the nucleoporins segregate into two visually distinct categories, which correspond to their location within the NPC [16,22]. The top half of Fig. 2A contains structural nucleoporins that form the core scaffold of the NPC. These nucleoporins coat the surface of the nuclear membrane in which the NPC is embedded [16]. The bottom half of Fig. 2A (labeled N1–N13) comprises all nucleoporins that contain FG-repeats. These FG-nucleoporins are anchored to the NPC scaffold and protrude into the inner of the NPC channel, presumably forming the selective barrier [16]. Fig. 2B plots the hydrophobicity and charge at pH 7.2 of the nucleoporins from Fig. 2A. This plot reveals that the majority of FG-nucleoporins (red circles) are characterized by net positive charge and low hydrophobicity, while the scaffold nucleoporins (blue circles) are both more negatively charged and more hydrophobic. Note that mammalian

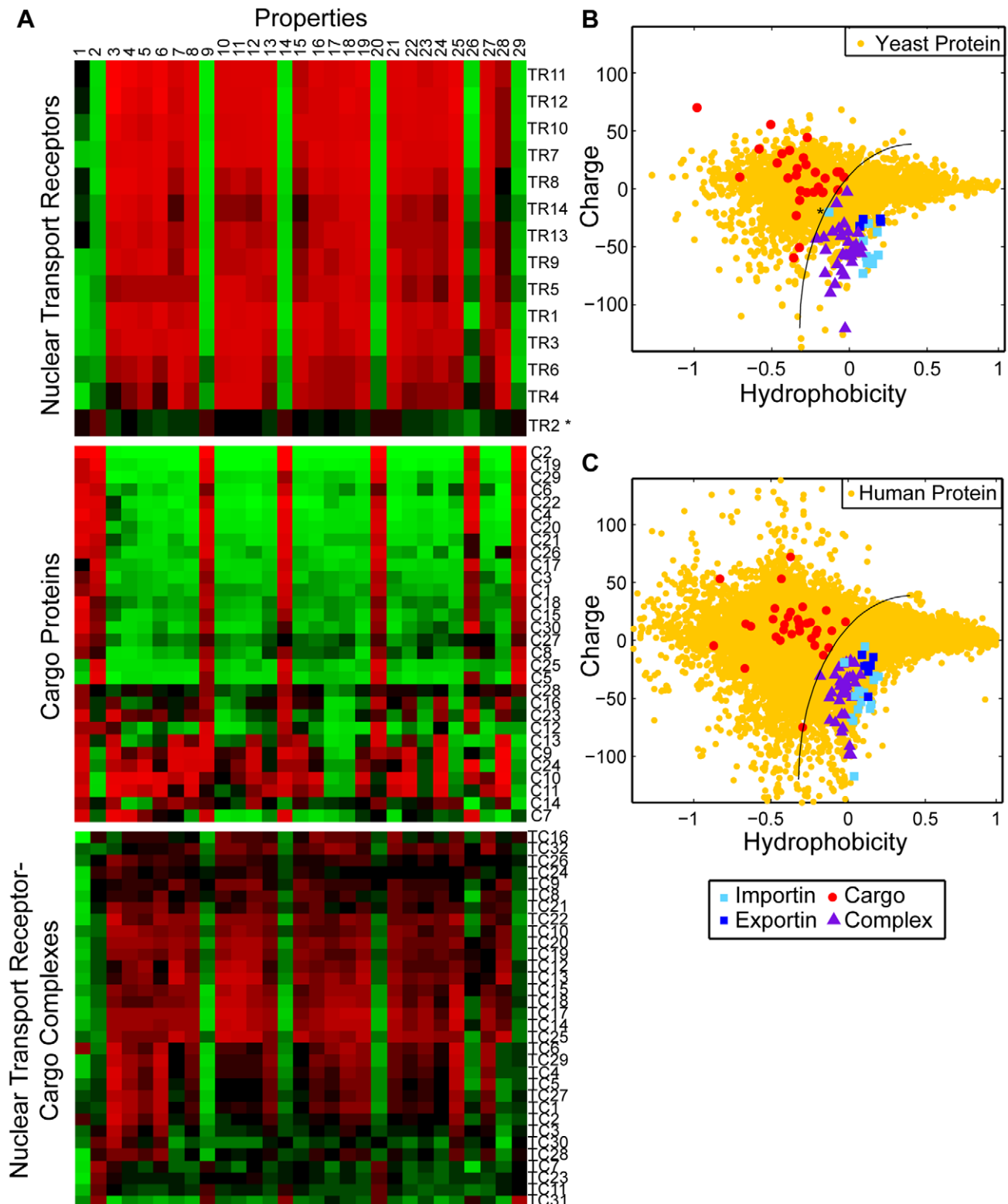


Figure 1. Nuclear transport receptors are more negatively charged than the majority of cellular proteins. A. Heat map of the physical properties of transport receptors, cargo proteins, and transport receptor-cargo complex from *S.c.* Each row corresponds to a different protein or complex (Table S1), and each column to a different property (Table S2). The value of each property was obtained as described in the main text. Similar proteins are clustered using a Euclidean distance metric. The profiles of transport receptors resemble each other, but differ from those of cargo proteins. B. Reduced representation using the first principal component of the 27 hydrophobicity scales, and net charge at intracellular pH. We compare importins (light blue squares), exportins (dark blue squares), cargo proteins (red circles), and transport receptor-cargo complexes (purple triangles) from *S.c.* with the entire *S.c.* proteome (yellow circles). Transport receptors are characterized by high hydrophobicity and net negative charge, and reside at the edge of the *S.c.* proteome. In contrast, most cargos have comparably low hydrophobicity and net positive charge. Note that

the light blue square corresponding to importin α (marked by an asterisk) falls in the region occupied by transport receptor-cargo complexes: indeed, efficient translocation of importin α requires binding to the transport receptors importin β or CAS, respectively C. Reduced representation of proteins from *H.s.* We compare importins (light blue squares), exportins (dark blue squares), cargo proteins (red circles), and transport receptor-cargo complexes (purple triangles) with the entire *H.s.* proteome (yellow circles). doi:10.1371/journal.pcbi.1000747.g001

nucleoporins also segregate into distinct groups (Table S3 and Fig. S3).

FG-nucleoporins contain both structured and unstructured domains. The unstructured domains are comprised of numerous hydrophobic FG-repeats separated by hydrophilic spacers [17,19,34]. Part of the FG-domain from the yeast nucleoporin Nup100 is shown in Fig. 2C, note that positive charge is present within the hydrophilic spacers that separate the hydrophobic FG-repeats. In this case the positive charge is predominantly derived from the amino acid lysine.

It is presumably the unfolded FG-domains that interact with transport receptors and thus determine selectivity ([25,28,29,36,37,39]). In Fig. 2D we analyze the unfolded FG-domains of all 13 FG-nucleoporins (Fig. 2A) as described by [19,36]. The data suggest that these FG-nucleoporin domains (red circles) are positively charged and hence complementary to the transport receptors (blue squares). This indicates a potential role for electrostatic interactions between the transport receptors and the unfolded domains of the FG-nucleoporins in NPC selectivity.

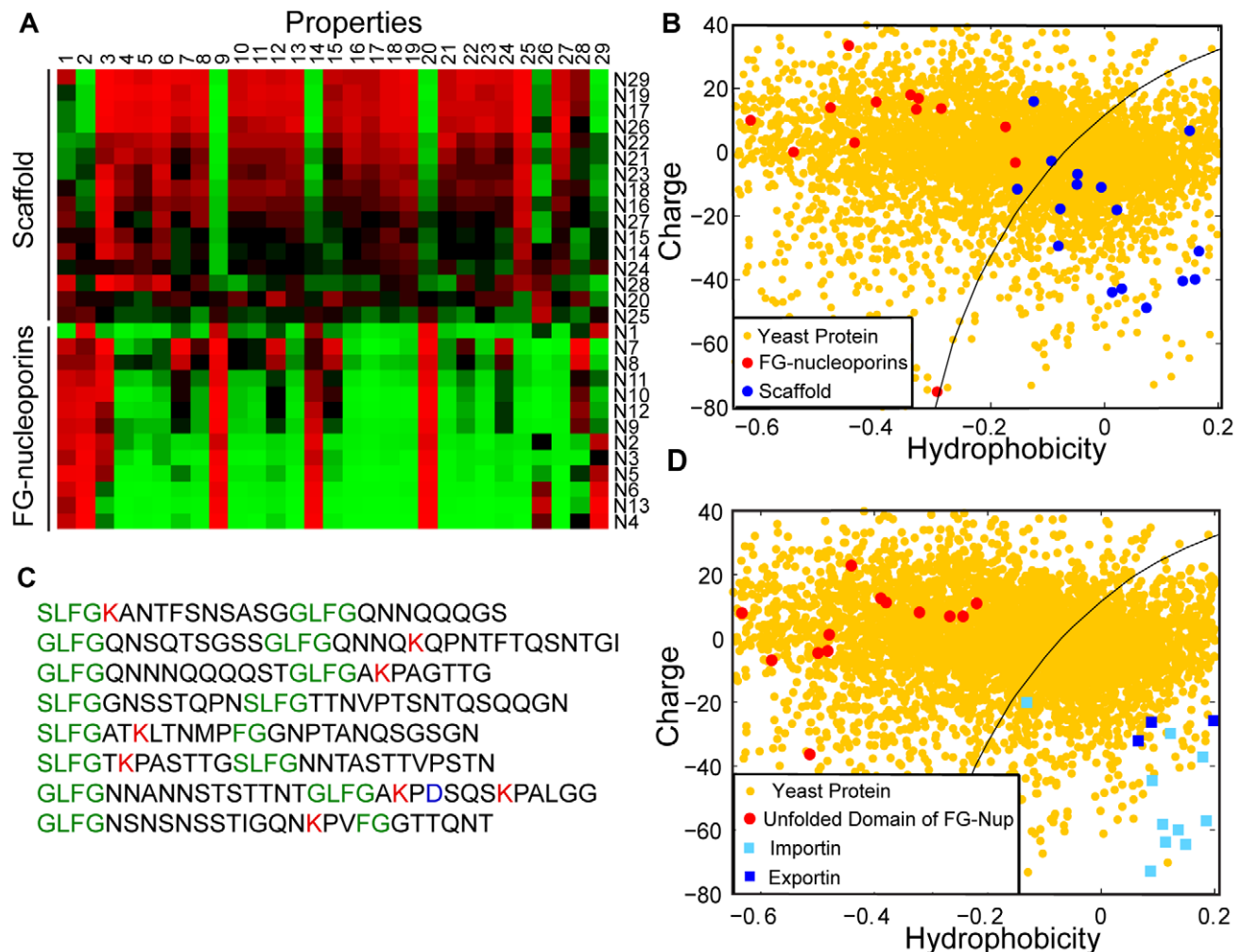


Figure 2. Most components of the NPC selectivity barrier are characterized by net positive charge. A. The physical properties of *S.c.* nucleoporins displayed as a heat map. Each column represents a different property (Table S2), and each row a nucleoporin (Table S3). The value of each property was obtained as described in the main text. Clustering the proteins using a Euclidean metric separates them into two groups, corresponding to scaffold and FG-nucleoporins [17,22]. B. 2D property space representation of *S.c.* nucleoporins. The hydrophobicity index is plotted against net charge at pH 7.2. The scaffold nucleoporins (blue circles) are relatively hydrophobic, while the FG-nucleoporins (red circles) are characterized by low hydrophobicity and net positive charge. Yellow circles represent the *S.c.* proteome. C. An FG-rich domain from the nucleoporin Nup100. Like other FG-nucleoporins, Nup100 contains largely unstructured FG-domains that contain numerous short hydrophobic FG-repeats separated by hydrophilic, positively-charged linkers. D. Comparison of the unstructured FG-domains [19] (red circles) with transport receptors (light and dark blue squares), and the *S.c.* proteome (yellow circles). The unstructured FG-domains and the transport receptors have complementary properties; the former have low hydrophobicity and mostly net positive charge, the latter have high hydrophobicity and net negative charge at pH 7.2. This suggests a role for electrostatic interactions between transport receptors and the unstructured FG-domains in the translocation reaction. doi:10.1371/journal.pcbi.1000747.g002

Model

Electrostatic interactions between transport receptors and FG-nucleoporins could result in an energy gain of multiple $k_B T$

The rate at which a protein (or protein complex) translocates through the NPC depends upon the energy barrier that it must overcome to enter the NPC [48,49]. The size of this energy barrier is given by $\Delta G = \Delta H - T\Delta S$, where the enthalpy change (ΔH) describes the binding energy of a protein to NPC components, T is the temperature, and ΔS is the change in entropy as a protein enters the NPC. When the charged protein enters the NPC and becomes spatially confined, the entropy of the system decreases, increasing ΔG and disfavoring translocation. However, ΔG can be lowered if specific interactions between a translocating protein and the NPC decrease enthalpy and thereby compensate for the decrease in entropy.

We next estimate whether electrostatic interactions between a transport receptor and the NPC interior could in principle be large enough to compensate for the loss of entropy. Within the cell, proteins are surrounded by counter ions that screen their charge; the screening length within the cytoplasm is estimated to be $\kappa^{-1} \sim 1\text{nm}$ [50]. We calculate the size of the interaction between each charge on the surface of the receptor, (where many charged amino acids reside; Fig. S4) and those nucleoporin charges within a hemisphere of radius 1nm (Q_{NPC} ; Fig. 3). We estimate the total

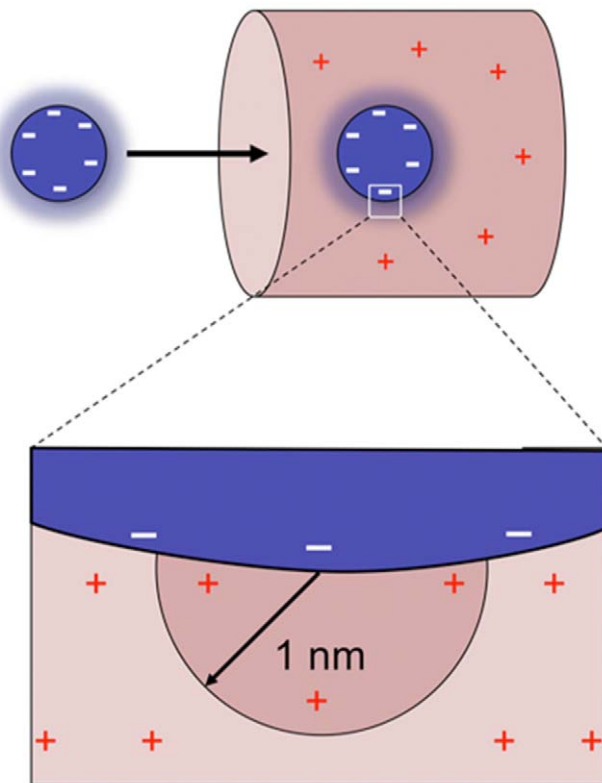


Figure 3. Strategy to estimate the electrostatic interactions between transport receptors and selective barrier components. The interaction between one charge on the surface of the receptor with those nucleoporin charges within a screening length of 1nm (Q_{NPC}) is determined. This interaction is then scaled up to include the transport receptor's total charge (Q_{NTR}), given by the sum of its charged residues. See main text for details. doi:10.1371/journal.pcbi.1000747.g003

charge of the transport receptor (Q_{NTR}) as the sum of its charged residues.

With this approach the electrostatic interaction energy between a transport receptor and the NPC can be approximated by the Coulombic interaction between Q_{NTR} and Q_{NPC} , separated by a screening length $\ell \approx \kappa^{-1}$:

$$\Delta H_{Electrostatic} \approx \frac{Q_{NTR} Q_{NPC}}{\epsilon \cdot \ell}, \quad (1)$$

where Q_{NTR} is the charge derived from the transport receptor, Q_{NPC} the charge derived from the NPC and ϵ is the dielectric constant of water. A direct derivation of equation (1) follows from assuming the electrostatic interactions are governed by Debye-Huckel theory, and then computing the interaction energy between the transport receptor and a hemisphere of size $\ell \approx \kappa^{-1}$ of uniform background charge density filling the nuclear pore. Sophisticated treatments of electrostatic interaction energy are possible [51,52], and such analyses will likely modify the $\Delta H_{Electrostatic}$ predicted by (1), however they should not change the order of magnitude. For human transport receptors at pH 7.2, we find that $-120 \cdot e \leq Q_{NTR} \leq -20 \cdot e$, with median $-54 \cdot e$, where e is the electron charge. To calculate Q_{NPC} , we approximate the yeast NPC as a cylinder with a radius of $\sim 19\text{nm}$ and a height of $\sim 37\text{nm}$ [16], and thus a pore volume of $\pi(19\text{ nm})^2 \times 37\text{ nm} = 41962\text{ nm}^3$. The yeast NPC is thought to contain 13 different types of FG-nucleoporins, with either eight or 16 copies of each [16,19]. A conservative estimate of the average number of nucleoporins within a hemisphere of radius 1nm is therefore $(8 \times 13)/41962\text{ nm}^3 \times 2\pi/3(1\text{ nm})^3 \approx 0.005$. The charge of an individual interior (yeast) nucleoporin ranges from $-3 \cdot e$ to $33 \cdot e$, with median $14 \cdot e$, resulting in $Q_{NPC} \sim 0.005 \times 14e = 0.07e$.

Equation (1) implies that $\Delta H_{Electrostatic} = 55.3 \frac{Q_{NPC} Q_{NTR}}{\epsilon} k_B T$, where Q_{NPC} and Q_{NTR} are expressed in units of the elementary charge. Using $\epsilon = 80$ we have that

$$\Delta H_{Electrostatic} \approx 0.05 \frac{Q_{NTR}}{e} k_B T \quad (2)$$

Thus, for a transport receptor with negative charge $Q = -50 \cdot e$, the energy gain due to direct interaction with the NPC interior is $\Delta H_{Electrostatic} \sim -2.5 k_B T$.

We emphasize that the model underlying Equation (2) is simplistic, for example: (i) Our treatment of the electrostatic interactions assumes Debye Huckel interactions, which are quantitatively modified when surface charge densities are sufficiently high; (ii) Detailed structural information for the charge distribution in the NPC channel is not available; (iii) We ignore potential entropic contributions to the electrostatic energy.

Equation (2) thus represents an order of magnitude estimate. Calculation of the entropic cost ΔS for a particle to enter the pore is complicated [12], requiring detailed knowledge of the environment both inside and outside the pore [53]. However, it is significant that the $\Delta H_{Electrostatic}$ predicted by Eqn. 2 is the same order of magnitude as ΔS [12]. Note that the entropic gain upon leaving the pore will compensate for the ensuing reduction in (electrostatic or other) binding energy [48]. The $\Delta H_{Electrostatic}$ magnitude calculated here would decrease the energy barrier for a transport receptor, increasing translocation efficiency. We therefore propose that electrostatic interactions between negatively charged particles and the positively charged selective barrier components provide a substantial part of the binding energy needed to mediate entry of a particle into the pore.

Discussion

Our study finds that transport receptors are more negatively charged than the majority of cellular proteins. Existing crystal structures of transport receptors such as importin β reveal that negative charge is distributed over the surface of the protein (Fig. S4; [40,45]). High sequence homology of importin β -like transport receptors, both within species and between species, suggests that net negative surface charge is a conserved property of this protein family.

How could negative surface charge promote the translocation of transport receptors through the NPC channel? Current models of translocation through the NPC postulate that hydrophobic FG-repeats within the selective barrier principally determine NPC selectivity [6,12,32,36,39,48,54,55]. We suggest that charge within the NPC channel may be a second feature relevant for translocation. The unfolded domains that separate FG-repeats are characterized by net positive charge (Fig. 2C, D), and we suggest that they represent a critical element of the selective barrier. Indeed, the positive charge of the spacers is conserved across multiple species [56], suggesting a functional constraint on the design of the spacer elements.

We propose that their negative surface charge allows transport receptors to adsorb to the positively charged nucleoporin domains via electrostatic interactions, facilitating selective partitioning of transport receptors and transport receptor-cargo complexes into the NPC (Fig. 4). Those soluble cellular proteins that are positively charged (Fig. 1B, C) should fail to enter the NPC efficiently because the corresponding energy barrier is too high. Note that according to this model, a translocating particle could become trapped within the NPC if its charge is too negative. The translocation rate is maximal only if the charge of a particle compensates for the decrease in entropy, rendering the total free energy barrier flat. We also note that repulsive electrostatic interactions between the patches of positive charge on the FG-domains may compete with the meshwork forming inter-FG linkages.

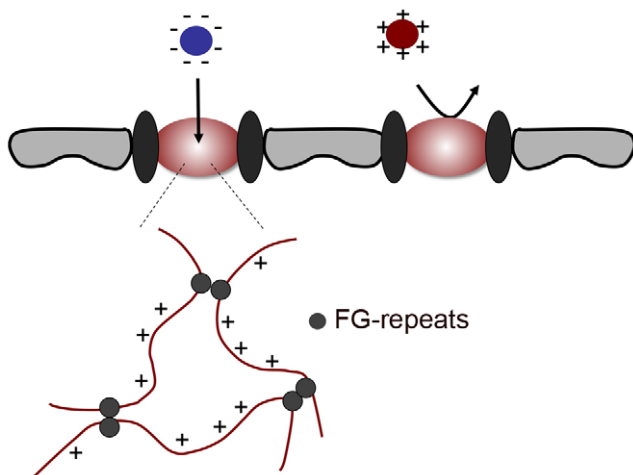


Figure 4. Charge as a selection criterion for nuclear transport. The unfolded domains that constitute the permeability barrier are positively charged (red lines) and generate a positive charge density within the barrier. Particles with negative surface charge (e.g. transport receptors) adsorb to the positively charged nucleoporin domains and thereby selectively partition into the permeability barrier. In contrast, most soluble cellular proteins lack net negative charge; their entry into the NPC is energetically highly unfavorable. doi:10.1371/journal.pcbi.1000747.g004

Electrostatic interactions as proposed here would help the NPC control entry of particles according to their surface charge, independently of their size, and thereby efficiently hinder passive diffusion of positively charged proteins. This is illustrated by histones, relatively small proteins that do not diffuse efficiently through the pore channel by themselves. Using the simple model in (2) we estimate the contribution of electrostatic interactions to the free energy of translocation for histone1 (H1). H1 has a molecular weight of 21kD and a charge $Q=53\cdot e$ at pH 7.2. Thus it would encounter an energy barrier of $\sim 2.6 k_B T$ at the NPC, making spontaneous diffusion highly inefficient. By binding to its two transport receptors, importin β and importin7 [57], H1 acquires a net charge of $-64\cdot e$, resulting in an energy barrier of $\sim -3.2 k_B T$. Combining this electrostatic energy gain with the entropic cost of entering the pore, and possible interactions with FG-repeats, could flatten the energy barrier, thus maximizing the H1 translocation rate.

Numerous signal transduction molecules are capable of rapidly shuttling between the nucleus and cytoplasm. How their entry into or exit from the nucleus is regulated is a matter of intense investigation. Fig. 5 shows the charge of a small number of well-characterized signaling proteins (listed in Table S4) plotted against their hydrophobicity. Two conclusions can be drawn from this plot. First, a number of shuttling proteins fall in the translocation competent regime, so are predicted to efficiently self-translocate through the NPC without association to transport receptors. One example is beta-catenin (marked with asterisk), which has a molecular weight of 84kDa and is above the passive diffusion limit of the NPC. Indeed, beta-catenin was discovered to translocate through the NPC without associating to a transport receptor [58]. Secondly, a number of signaling proteins, for example STAT3 and

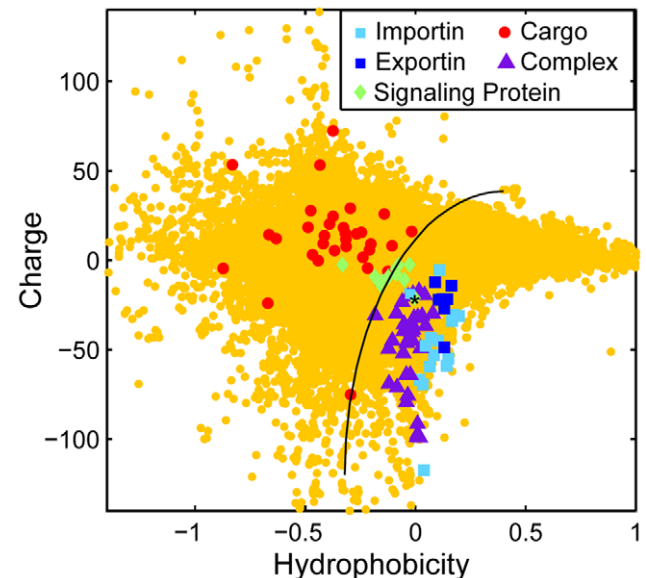


Figure 5. Charge and hydrophobicity of a selection of signaling proteins from *H.s.* (listed in Table S4) (green diamonds) in comparison with importins (light blue squares), exportins (dark blue squares), cargo proteins (red circles), transport receptor-cargo complexes (purple triangles), and the *H.s.* proteome (yellow circles). Signaling proteins that fall into the transport receptor-cargo complex regime (bottom right quadrant) are predicted by this analysis to efficiently self-translocate through the NPC without specific association to transport receptors. Note that beta-catenin (marked with asterisk), one example of a large protein that translocates without binding to a transport receptor, falls in this regime. doi:10.1371/journal.pcbi.1000747.g005

SMAD2, fall on the edge of the translocation competence regime. Our analysis predicts that addition of negative charge, as occurs through phosphorylation, could contribute to regulating the translocation rate of such particles. Using Eqn. (2) we estimate that the addition of a single negative charge (corresponding to phosphorylation of a single site) would lead to a decrease in the energy barrier by $\sim 0.05 k_B T$. Thus, phosphorylation of a small number of residues may tune the rate of translocation, even in the absence of a transport receptor.

Specific hydrophobic regions on the surface of nuclear transport receptors are critical for translocation. One important future challenge is to dissect the contribution of both a particle's negative charge and its hydrophobicity to the translocation reaction. This question could be addressed with direct experimental tests that measure a particle's translocation rate as a function of its charge, size and hydrophobicity. The charge and size of particles can be independently varied, and doing so would disentangle their relative contributions to translocation through the pore. The results of such experiments will enable the correlation of a protein's translocation rate through the NPC with its charge and hydrophobicity properties. Moreover, this information could facilitate a novel prediction tool for transport receptor-cargo matching.

Supporting Information

Figure S1 Correlation between the polarity scale (Grantham) and the aggregate hydrophobicity scale developed in the main text for transport receptors (green squares), complexes (blue triangles), and cargoes (red circles). The correlation coefficient is $r^2 = -0.93$. Found at: doi:10.1371/journal.pcbi.1000747.s001 (0.10 MB TIF)

Figure S2 Heat map of the physical properties of transport receptors, cargo proteins, and transport receptor-cargo complexes in *Homo sapiens*. Each row corresponds to a different protein or complex (Table S1), and each column to a different property (Table S2). The value of each property (with the exception of the isoelectric point) was obtained for every protein by summing the contribution from every amino acid in its sequence, and normalizing by its sequence length. Each property is normalized to have mean zero over the entire set of proteins. Bright red and green correspond to 3 standard deviations above and below this mean, respectively. Hierarchical clustering, based on a Euclidean distance metric, groups similar proteins together. The profiles of individual transport receptors resemble each other, but are visibly different to the profiles of individual cargo proteins. Found at: doi:10.1371/journal.pcbi.1000747.s002 (1.27 MB TIF)

Figure S3 A. The physical properties of H.s. nucleoporins displayed in a heat map. Each column represents a different property (Table S2), and each row represents a different nucleoporin (Table S3). The value of each property was obtained as in Fig. 1. Each property is normalized to have mean zero over the entire set of proteins; bright red and green correspond to 3 standard deviations above and below this mean, respectively. Clustering the proteins using a Euclidean metric separates them into two major groups. This plot reveals that FG-nucleoporins are

biophysically distinct from other, possibly structural, nucleoporins. Gle1 (N23, marked with an asterisk), does not contain FG-repeats but appears to have similar properties to FG-nucleoporins. B. 2D property space representation of H.s. nucleoporins in context of the human proteome. The hydrophobicity index (the first principal component of the 27 hydrophobicity scales) is plotted against the net charge at intracellular pH. A subset of nucleoporins is characterized by low hydrophobicity and net positive charge, while the other group is relatively hydrophobic. C. 2D property space representation of H.s. nucleoporins. As is the case for the yeast NPC, many human FG-nucleoporins carry net positive charge and are relatively hydrophilic, while most other nucleoporins are collectively net negatively charged and more hydrophobic. The localization of nucleoporins on this 2D plot may allow predictions concerning their localization within the NPC.

Found at: doi:10.1371/journal.pcbi.1000747.s003 (1.15 MB TIF)

Figure S4 Crystal structure of importin β [40,45], with negative residues (aspartic and glutamic acid) highlighted in blue. This structure reveals that the negative charge is distributed over the surface of the protein. High sequence homology of importin β -like transport receptors, both within species and between species, suggests that net negative surface charge is a conserved property of this protein family.

Found at: doi:10.1371/journal.pcbi.1000747.s004 (0.70 MB TIF)

Table S1 Compilation of nuclear transport receptors, cargos, and cognate transport receptor-cargo complexes from *Saccharomyces cerevisiae* (A, B) and *Homo sapiens* (C, D) as analyzed in this manuscript.

Found at: doi:10.1371/journal.pcbi.1000747.s005 (0.28 MB DOC)

Table S2 Compilation of biophysical properties analyzed in this manuscript

Found at: doi:10.1371/journal.pcbi.1000747.s006 (0.12 MB DOC)

Table S3 Compilation of yeast and human nucleoporins analyzed in this manuscript.

Found at: doi:10.1371/journal.pcbi.1000747.s007 (0.20 MB DOC)

Table S4 Signaling proteins from *Homo sapiens* analyzed in this paper.

Found at: doi:10.1371/journal.pcbi.1000747.s008 (0.06 MB DOC)

Acknowledgments

We thank Jacob Israelachvili, Andrew Murray, Carey Nadell, Erin O'Shea, Michael Rexach, Thomas Schwartz and Boris Shraiman, for important conversations.

Author Contributions

Conceived and designed the experiments: LJC MPB KR. Analyzed the data: LJC MPB KR. Contributed reagents/materials/analysis tools: LJC MPB KR. Wrote the paper: LJC MPB KR.

References

- Kose S, Imamoto N, Tachibana T, Shimamoto T, Yoneda Y (1997) Ran-assisted nuclear migration of a 97-kD component of nuclear pore-targeting complex. *J Cell Biol* 139: 841–849.
- Ribbeck K, Lipowsky G, Kent HM, Stewart M, Görlich D (1998) NTF2 mediates nuclear import of Ran. *EMBO J* 17: 6587–6598.
- Schwoebel ED, Talcott B, Cushman I, Moore MS (1998) Ran-dependent signal-mediated nuclear import does not require GTP hydrolysis by Ran. *J Biol Chem* 273: 35170–35175.
- Englmeier L, Olivo JC, Mattaj IW (1999) Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis. *Curr Biol* 9: 30–41.
- Ribbeck K, Kutay U, Paraskeva E, Görlich D (1999) The translocation of transport-cargo complexes through nuclear pores is independent of both Ran and energy. *Curr Biol* 9: 47–50.
- Ribbeck K, Görlich D (2001) Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* 20: 1320–1330.

7. Yang W, Gelles J, Musser SM (2004) Imaging of single-molecule translocation through nuclear pore complexes. *Proc Natl Acad Sci U S A* 101: 12887–12892.
8. Gorlich D, Kutay U (1999) Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* 15: 607–660.
9. Conti E, Muller CW, Stewart M (2006) Karyopherin flexibility in nucleocytoplasmic transport. *Curr Opin Struct Biol* 16: 237–244.
10. Stewart M, Baker RP, Bayliss R, Clayton L, Grant RP, et al. (2001) Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import. *FEBS Lett* 498: 145–149.
11. Stewart M (2007) Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol* 8: 195–208.
12. Rout MP, Aitchison JD, Suprpto A, Hjertaas K, Zhao Y, et al. (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol* 148: 635–651.
13. Vasu SK, Forbes DJ (2001) Nuclear pores and nuclear assembly. *Curr Opin Cell Biol* 13: 363–375.
14. Cronshaw JM, Krutchinsky AN, Zhang W, Chait BT, Matunis MJ (2002) Proteomic analysis of the mammalian nuclear pore complex. *J Cell Biol* 158: 915–927.
15. Tran EJ, Wentse SR (2006) Dynamic nuclear pore complexes: life on the edge. *Cell* 125: 1041–1053.
16. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, et al. (2007) The molecular architecture of the nuclear pore complex. *Nature* 450: 695–701.
17. Rout MP, Wentse SR (1994) Pores for thought: nuclear pore complex proteins. *Trends in Cell Biology* 4: 357–364.
18. Denning DP, Uversky V, Patel SS, Fink AL, Rexach M (2002) The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. *J Biol Chem* 277: 33447–33455.
19. Denning DP, Patel SS, Uversky V, Fink AL, Rexach M (2003) Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc Natl Acad Sci U S A* 100: 2450–2455.
20. Akey CW, Goldfarb DS (1989) Protein import through the nuclear pore complex is a multistep process. *J Cell Biol* 109: 971–982.
21. Grote M, Kubitschek U, Reichelt R, Peters R (1995) Mapping of nucleoporins to the center of the nuclear pore complex by post-embedding immunogold electron microscopy. *J Cell Sci* 108(Pt 9): 2963–2972.
22. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, et al. (2007) Determining the architectures of macromolecular assemblies. *Nature* 450: 683–694.
23. Rexach M, Blobel G (1995) Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell* 83: 683–692.
24. Bayliss R, Ribbeck K, Akin D, Kent HM, Feldherr CM, et al. (1999) Interaction between NTF2 and xFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J Mol Biol* 293: 579–593.
25. Bayliss R, Littlewood T, Stewart M (2000) Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* 102: 99–108.
26. Fribourg S, Braun IC, Izaurralde E, Conti E (2001) Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. *Mol Cell* 8: 645–656.
27. Strawn LA, Shen T, Wentse SR (2001) The GLFG regions of Nup116p and Nup100p serve as binding sites for both Kap95p and Mex67p at the nuclear pore complex. *J Biol Chem* 276: 6445–6452.
28. Bayliss R, Leung SW, Baker RP, Quimby BB, Corbett AH, et al. (2002) Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J* 21: 2843–2853.
29. Bayliss R, Littlewood T, Strawn LA, Wentse SR, Stewart M (2002) GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. *J Biol Chem* 277: 50597–50606.
30. Bednenko J, Cingolani G, Gerace L (2003) Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. *J Cell Biol* 162: 391–401.
31. Cushman I, Palzkill T, Moore MS (2006) Using peptide arrays to define nuclear carrier binding sites on nucleoporins. *Methods* 39: 329–341.
32. Ribbeck K, Gorlich D (2002) The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J* 21: 2664–2671.
33. Grant RP, Neuhaus D, Stewart M (2003) Structural basis for the interaction between the Tap/NXF1 UBA domain and FG nucleoporins at 1A resolution. *J Mol Biol* 326: 849–858.
34. Strawn LA, Shen T, Shulga N, Goldfarb DS, Wentse SR (2004) Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat Cell Biol* 6: 197–206.
35. Frey S, Richter RP, Goerlich D (2006) FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* 314: 815–817.
36. Patel SS, Belmont BJ, Sante JM, Rexach MF (2007) Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* 129: 83–96.
37. Jovanovic-Taliman T, Tetenbaum-Novatt J, McKenney AS, Zilman A, Peters R, et al. (2009) Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. *Nature* 457: 1023–1027.
38. Quimby BB, Leung SW, Bayliss R, Harreman MT, Thirumala G, et al. (2001) Functional analysis of the hydrophobic patch on nuclear transport factor 2 involved in interactions with the nuclear pore in vivo. *J Biol Chem* 276: 38820–38829.
39. Frey S, Gorlich D (2007) A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* 130: 512–523.
40. Liu SM, Stewart M (2005) Structural basis for the high-affinity binding of nucleoporin Nup1p to the *Saccharomyces cerevisiae* importin-beta homologue, Kap95p. *J Mol Biol* 349: 515–525.
41. Vetter IR, Arndt A, Kutay U, Gorlich D, Wittinghofer A (1999) Structural view of the Ran-Importin beta interaction at 2.3 A resolution. *Cell* 97: 635–646.
42. Chook YM, Blobel G (1999) Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. *Nature* 399: 230–237.
43. Bullock TL, Clarkson WD, Kent HM, Stewart M (1996) The 1.6 angstroms resolution crystal structure of nuclear transport factor 2 (NTF2). *J Mol Biol* 260: 422–431.
44. Stewart M, Kent HM, McCoy AJ (1998) Structural basis for molecular recognition between nuclear transport factor 2 (NTF2) and the GDP-bound form of the Ras-family GTPase Ran. *J Mol Biol* 277: 635–646.
45. Liker E, Fernandez E, Izaurralde E, Conti E (2000) The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain. *EMBO J* 19: 5587–5598.
46. Rice P, Longden I, Bleasby A (2000) EMBOS: The European Molecular Biology Open Software Suite. *Trends in Genetics* 16: 276–277.
47. Dingwall C, Laskey RA (1991) Nuclear targeting sequences—a consensus? *Trends Biochem Sci* 16: 478–481.
48. Rout MP, Aitchison JD, Magnasco MO, Chait BT (2003) Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol* 13: 622–628.
49. Zilman A, Di Talia S, Chait BT, Rout MP, Magnasco MO (2007) Efficiency, selectivity, and robustness of nucleocytoplasmic transport. *PLoS Comput Biol* 3: e125.
50. Nelson P (2008) *Biological Physics*; O'Neill J, editor. New York: Clancy Marshall.
51. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. *Science* 268: 1144–1149.
52. Schreiber G, Haran G, Zhou HX (2009) *Fundamental Aspects of Protein-Protein Association Kinetics*. *Chem Rev*.
53. Zhou HX, Rivas G, Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 37: 375–397.
54. Ben-Efraim I, Gerace L (2001) Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import. *J Cell Biol* 152: 411–417.
55. Pyhtila B, Rexach M (2003) A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex. *J Biol Chem* 278: 42699–42709.
56. Denning DP, Rexach MF (2007) Rapid evolution exposes the boundaries of domain structure and function in natively unfolded FG nucleoporins. *Mol Cell Proteomics* 6: 272–282.
57. Jakel S, Albig W, Kutay U, Bischoff FR, Schwamborn K, et al. (1999) The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *Embo J* 18: 2411–2423.
58. Fagotto F, Gluck U, Gumbiner BM (1998) Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. *Curr Biol* 8: 181–190.