Nucleoporins' exclusive amino acid sequence features regulate their transient interaction with and selectivity of cargo complexes in the nuclear pore

Mohaddeseh Peyro[a], Andrew M. Dickson[a], and Mohammad R. K. Mofrad[a,b,*]
[a]Molecular Cell Biomechanics Laboratory, Departments of Bioengineering and Mechanical Engineering, University of California, Berkeley, CA 94720; [b]Molecular Biophysics and Integrative Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

ABSTRACT Nucleocytoplasmic traffic of nucleic acids and proteins across the nuclear envelop via the nuclear pore complexes (NPCs) is vital for eukaryotic cells. NPCs screen transported macromolecules based on their morphology and surface chemistry. This selective nature of the NPC-mediated traffic is essential for regulating the fundamental functions of the nucleus, such as gene regulation, protein synthesis, and mechanotransduction. Despite the fundamental role of the NPC in cell and nuclear biology, the detailed mechanisms underlying how the NPC works have remained largely unknown. The critical components of NPCs enabling their selective barrier function are the natively unfolded phenylalanine- and glycine-rich proteins called "FG-nucleoporins" (FG Nups). These intrinsically disordered proteins are tethered to the inner wall of the NPC, and together form a highly dynamic polymeric meshwork whose physicochemical conformation has been the subject of intense debate. We observed that specific sequence features (called largest positive like-charge regions, or lpLCRs), characterized by extended subsequences that only possess positively charged amino acids, significantly affect the conformation of FG Nups inside the NPC. Here we investigate how the presence of lpLCRs affects the interactions between FG Nups and their interactions with the cargo complex. We combine coarse-grained molecular dynamics simulations with time-resolved force distribution analysis to disordered proteins to explore the behavior of the system. Our results suggest that the number of charged residues in the lpLCR domain directly governs the average distance between Phe residues and the intensity of interaction between them. As a result, the number of charged residues within lpLCR determines the balance between the hydrophobic interaction and the electrostatic repulsion and governs how dense and disordered the hydrophobic network formed by FG Nups is. Moreover, changing the number of charged residues in an lpLCR domain can interfere with ultrafast and transient interactions between FG Nups and the cargo complex.

INTRODUCTION

The nuclear pore complex (NPC), as the largest protein complex in the cell, has been the subject of research for the past few decades. Despite extensive research, there is still not enough information about how this protein complex performs its intricate function. NPC is responsible for bidirectional transport of cargo through the nuclear envelope. The primary, and interesting, open question about NPC's function is how it enables transport to happen in a fast yet selective manner (Aramburu and Lemke, 2017; Mulder, 2018). This feature, known as the transport paradox, is an active area of research.

NPC only enables transport of cargos that have specific properties and facilitates transport process in the order of milliseconds (Wolf and Mofrad, 2009; Azimi *et al.*, 2014; Jovanovic-Talisman and Zilman, 2017; Tan *et al.*, 2018; Matsuda and Mofrad, 2021).

FG nucleoporins, i.e., primarily disordered proteins rich in phenylalanine-glycine (FG) repeats, are known to be the major role players in nucleocytoplasmic transport (Soheilypour and Mofrad, 2016; Jovanovic-Talisman and Zilman, 2017). FG nucleoporins form a hydrophobic barrier that blocks passive transport of cargos larger than ~40 kDa, while allowing smaller molecules to freely diffuse (Wolf and Mofrad, 2008; Jamali *et al.*, 2011; Azimi and Mofrad, 2013; Timney *et al.*, 2016; Jovanovic-Talisman and Zilman, 2017). Large molecules, however, can actively pass through the nuclear pore by binding to transport receptors. Transport receptors, in turn, make weak and transient interactions with FG Nups, enabling them to transport through the NPC (Hough *et al.*, 2015; Raveh *et al.*, 2016). From a biophysical perspective, the transient hydrophobic interactions enable the cargo complex to overcome the hydrophobic barrier made by FG Nups.

Different experimental and computational methods have been utilized to study nucleocytoplasmic transport. Due to the complex nature and intricate structure of the NPC, the detailed underlying mechanism of transport through the NPC is usually not tractable via experimental methods. Computational methods, on the other hand, can easily explore the transport mechanism with a higher temporal and spatial resolution. Nevertheless, due to high computational expense, these methods are usually unable to reach high enough timescales that can fully describe the transport process. The spatiotemporal scale that is necessary for unveiling the transport mechanism is currently unattainable by either experimental or computational techniques (Moussavi-Baygi and Mofrad, 2016; Sakiyama *et al.*, 2016; Soheilypour and Mofrad, 2016; Soheilypour and Mofrad, 2018). However, integration of experimental and computational techniques can lead to invaluable insights about the function of the NPC.

FG Nups, the major role players in the transport process, are intrinsically disordered, i.e., they lack a fixed secondary structure. Therefore, the usual structure–function relationship cannot be applied to them (Wright and Dyson, 1999; Uversky, 2013) primarily due to high mutation rates (Brown *et al.*, 2002; DeGrasse *et al.*, 2009; Aramburu and Lemke, 2017; Jovanovic-Talisman and Zilman, 2017). As a result, efforts have been made to identify correlations between sequences of FG Nups and their conformation and possible role in nucleocytoplasmic transport instead. Yamada *et al.* showed that the ratio of charged residues to hydrophobic residues determines whether FG Nups form a collapsed coil, a relaxed coil, or an extended coil conformation (Yamada *et al.*, 2010). Tagliazucchi *et al.* studied the effect of amino acid sequences of FG Nups on transport of cargo complex through the NPC (Tagliazucchi *et al.*, 2013). Ando *et al.* took a more generic approach and analyzed the location of clusters of charged residues and FG motifs in the sequences of more than 1100 FG Nups from more than 250 species. They found that clusters of FG motifs are located more toward the center of the NPC and clusters of charged residues are located toward the scaffold of the NPC (Ando *et al.*, 2013). Ghavami *et al.* analyzed the change in FG Nups conformation inside yeast NPC on change in sequences of FG Nups (Ghavami *et al.*, 2014).

Our previous work on the distribution of different types of amino acids in sequences of FG Nups showed that, besides previously found patterns such as FG Nups having charge-rich and charge-poor regions (Huang *et al.*, 2019; Yamada *et al.*, 2010), more specific charge distribution patterns may be exhibited in the FG Nups. By analyzing a large dataset of FG Nups, we found an evolutionarily conserved sequence feature in the majority of FG Nups: long sequences of uninterrupted positively charged residues with a low charge density located at the N terminus of FG Nups toward the center of the NPC. We named these patterns largest positive like charge regions (lpLCRs) (Peyro *et al.*, 2015b, 2021). Figure 1 shows a schematic of charge distribution and the conformation of the disordered domain of yeast FG Nup with its lpLCR boxed. As the next step of the study, we used a one-bead-per-amino-acid-coarse-grained model to study the biophysical effect of lpLCRs on the conformation of FG Nups inside the NPC. Our simulations suggested that lpLCRs increase the movement of FG Nups inside the NPC and lead to a more even distribution of FG motifs in the NPC as well as more cross-interactions between FG Nups (Peyro *et al.*, 2015a).

In the present study, we combine a coarse-grained molecular dynamics model with time-resolved force distribution analysis (TRFDA) to further analyze the role of lpLCRs on nucleocytoplasmic transport. This study is the first to apply TRFDA to disordered proteins. We analyze the role of lpLCR on nucleocytoplasmic transport by exploring how lpLCRs affect the conformational ensemble of FG Nups, movement of the cargo complex, and the interaction of the cargo complex with FG Nups. For this purpose, first, the NPC nanochannel is subdivided into 17 ring cross-sections that each maintain the eightfold symmetry of the NPC but contain only one type of FG Nup. A simplified representation of cargo complex is included in the system, considering the cargo as a sphere and the transport protein as seven hydrophobic elements (phenylalanine) on the sphere. Second, the whole-NPC model including all the FG Nups and a cargo complex initially located at two-thirds the height of NPC is analyzed. The results of the simulations are as follows.

RESULTS

It is important to note that throughout this study, the term "lpLCR" refers to the lpLCR existing in the sequence of FG Nups (Peyro *et al.*, 2021). Ring cross-sections of the NPC containing only one type of FG Nup were studied by dividing the axial length of the entire NPC into separate rings that each contain only one type of FG Nup, copied in an eightfold symmetric manner. The ring complexes were simulated with coarse-grained models of the section of the NPC scaffold, the anchoring nup copies, and a spherical cargo bound to seven phenylalanines that represent transport proteins (see *Materials and Methods* for details of why this geometry is used for the cargo complex). This geometry accounts for FG Nup–FG Nup interactions and FG Nup–cargo complex interactions in each cross-section of the NPC. Studying the behavior of Nups and cargo in these isolated rings can provide detailed information about how each type of FG Nup may contribute to the transport process. For the second stage, a whole-NPC model is built from all FG Nups and a cargo complex located at two-thirds the height of the NPC. To distinguish and quantify the behavior of FG Nups and the cargo complex in each section, density maps of these entities were studied. Density maps are extracted in wild type and two mutant (negative and positive control) cases. In the negative control, the charged residues in the lpLCR were mutated to alanine, while in the positive control, the same number of charged residues in the lpLCR was doubled by mutating alanines to positively charged residues.

The results of our single FG Nup simulations (please see *Materials and Methods* for details of the simulation setup for single Nups), combined with differences in charge distribution, suggest that yeast FG Nups can be categorized into four groups (Figure 1):

1. Nsp1, Nup100, and Nup116, which are long FG Nups and contain both charge-rich domains and lpLCRs. Nsp1 has 617 residues in its disordered domain. Its lpLCR domain is 176 residues
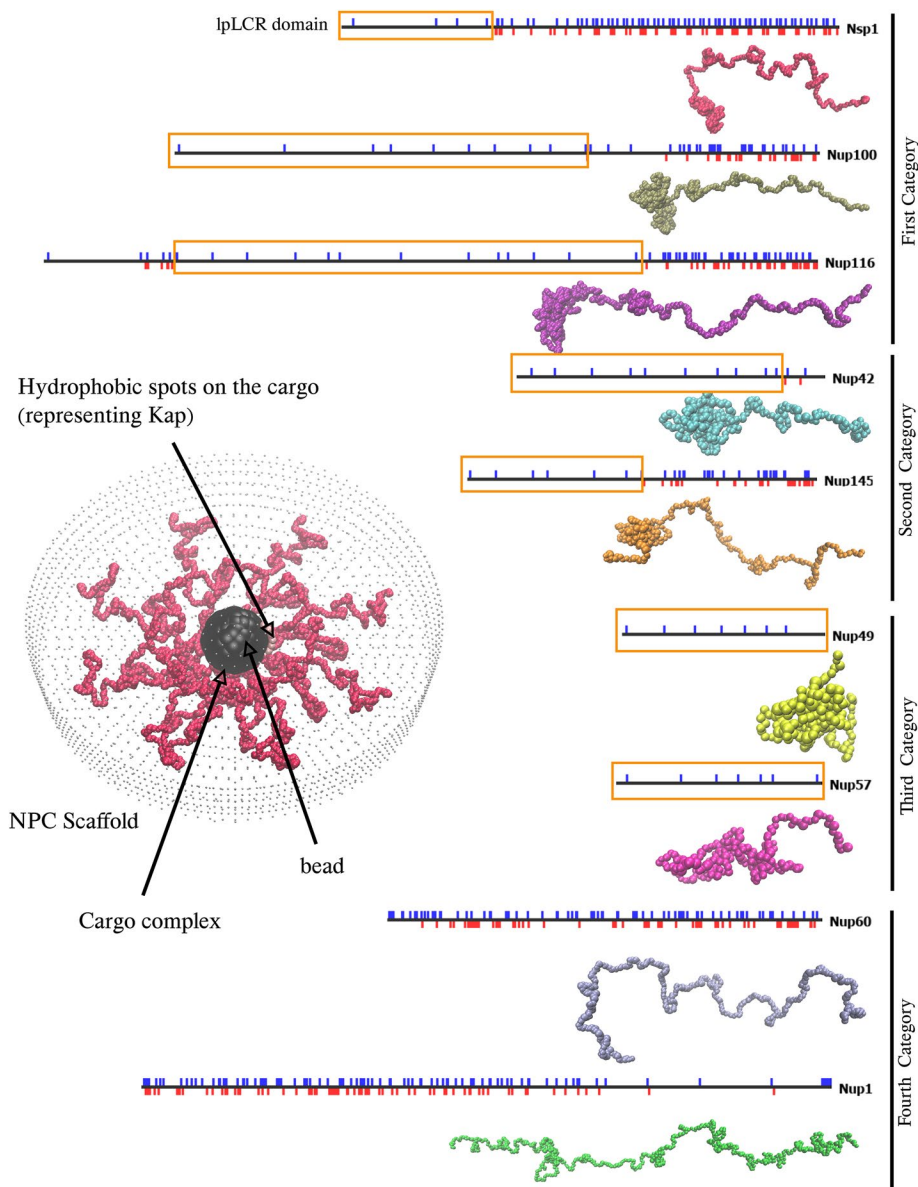
FIGURE 1: Initial conformation of a ring complex. Eight copies of a Nup (in this figure, Nsp1) are tethered to the scaffold of the NPC. The scaffold is represented by small noninteracting beads. The cargo complex is made up of a spherical cargo with seven Phe residues attached to represent Kap (Ghavami *et al.*, 2014). The cargo complex is initially located at the center of the ring. Right: charge distribution of yeast FG Nups and a snapshot of their conformation in single Nup simulations. In charge distribution schematics, the FG Nup sequence is schematically drawn as a black line with positively charged residues shown with small blue vertical lines pointing upward and negatively charged residues shown with small red vertical lines pointing downward. The lpLCR of each Nup is boxed. Only the disordered domain of the Nups is drawn and modeled. First category: Nsp1, Nup100, and Nup116. These are long FG Nups and contain both charge-rich domains and lpLCRs. These proteins form a globular head (lpLCR domain) and a relaxed coil tail (non-lpLCR, charge-rich domain). Second category: Nup42 and Nup145. These are medium-length FG Nups and contain both charge-rich domains and lpLCR. These proteins form a globular head (lpLCR domain) and a relaxed coil tail (non-lpLCR, charge-rich domain). Third category: Nup49 and Nup57. These proteins are short and do not have a charge-rich domain. Their length is made up of an lpLCR and their conformation only includes a globular head. Fourth category: Nup1, Nup60. These Nups do not feature lpLCRs .Due to the lack of an lpLCR domain, the entire length of these proteins forms a relaxed coil.

and contains four positive charges. This protein forms a collapsed coil head (lpLCR domain) and a relaxed coil tail (non-lpLCR, charge-rich domain). Nup100 has 800 residues in its disordered

domain with an lpLCR domain of 511 residues that contain 10 positive charges. Nup116 has 960 residues in its disordered domain with an lpLCR domain of 572 residues that contain 14 positive charges. The conformations of Nup100 and Nup116 look similar to that of Nsp1. In the next section, it will be shown that FG Nups in category 1 form a high-density conformation at the center of the ring cross-sections of NPC.

2. Nup42 and Nup145, which are medium-length FG Nups and contain both charge-rich domains and lpLCR. Nup42 has 382 residues in its disordered domain with an lpLCR domain of 317 residues that contain seven positive charges. Nup145 has 433 residues in its disordered domain with an lpLCR domain of 217 residues that contain seven positive charges. These proteins form a collapsed coil head (lpLCR domain) and a relaxed coil tail (non-lpLCR, charge-rich domain). In the next section of *Results*, it will be shown that FG Nups in category 2 form a donut-shaped high-density conformation near the scaffold of the ring. In the donut-shaped conformations, Nup molecules cross-interact with each other.

3. Nup49 and Nup57, which are short FG Nups, have a very low charge density and contain only positively charged residues; therefore, the entire sequence is an lpLCR. Nup49 has 251 residues in its disordered domain with an lpLCR domain of 251 residues that contain seven positive charges. Nup57 has 255 residues in its disordered domain with an lpLCR domain of 255 residues that contain seven positive charges. These proteins do not have a charge-rich domain and their conformation only includes a collapsed coil head.

4. Nup1 and Nup60 and Nup159, which do not feature lpLCRs. Nup1 has 856 residues in its disordered domain. Nup60 has 539 residues in its disordered domain. None of these FG Nups feature an lpLCR domain. Due to lack of an lpLCR domain, the entire length of these proteins forms a relaxed coil conformation. All yeast FG Nups (except for Nup159 that features a structured domain in its N-terminus which cannot be captured by the model used in this study) and their conformation, as well as a representation of the initial setup of ring simulations, can be seen in Figure 1. The ring shown in Figure 1 is an Nsp1-containing ring cross-section. The lpLCR domain of each Nup is boxed in this figure.

The outline of the *Results* section is as follows: first, the density maps of important elements of the system, i.e., Nup, cargo, FG motifs, and cargo hydrophobic spots are presented for representatives of each category of Nups mentioned earlier. Second, the force distribution analysis (FDA) of the interaction within and between the Nups in each ring is presented. Third, the density map and the FDA of FG Nups interactions with the cargo complex within each ring cross-section will be presented. At last, the whole-NPC model is analyzed by investigating the FG network and movement of the cargo complex in the system.

Analysis of trajectory of FG Nups and the cargo complex in the ring cross-section of FG Nups

Based on the categories mentioned earlier, we present the results of one of the FG Nups as a representative of each category. The ring-containing cross-section of FG Nups was analyzed by running the coarse-grained model of the system for 1 µs. In these simulations, eight copies of each type of Nup were tethered to the scaffold and the cargo complex was initially located at the center of the ring (please see Figure 1). The system was equilibrated for 50 ns and the simulation was run for 1 µs. This was long enough to capture the behavior of the system. It is important to note that, due to the disordered nature of the FG Nups, the system does not converge to a certain state; however, the conformational ensemble of the Nups shows an average behavior that can be considered equivalent to "convergence" in the case of structured protein simulations. For each Nup, the negative control and the positive control simulations were run with the same initial conditions and the results are presented next to the wild-type system. Density maps of key elements of these simulations are drawn for each case, i.e., the Nup, cargo, FG motifs, and hydrophobic spots on cargo that represent Kap. Density maps are the best method to analyze the way each of the elements is distributed in the ring. To draw the density maps, the ring is divided into a granular grid, and the number of times an element is spotted in each cell is counted. Cells are then colored according to their counts (color range is presented next to density maps).

Nsp1

Density maps of the Nsp1-containing ring are presented in Figure 2 as a representative of the first category. Density maps show that in wild-type and mutant simulations, Nsp1 occupies most of the space inside the ring and leaves a small free space for the cargo (Figure 2A). This behavior is expected, since Nsp1 is a long FG Nup in the yeast NPC (Nsp1 contains 617 residues in its disordered domain) and eight copies of this molecule occupy most of the space inside the ring. As shown in Figure 1, the Nsp1 sequence is made up of two distinct sections: a highly charged section with a relaxed coil conformation that is located toward the NPC scaffold and, a second section that is located toward the center of NPC and has a lower charge density and a collapsed coil conformation (Yamada *et al.*, 2010). The second section is the Nsp1 lpLCR domain, which has four positively charged residues. The collapsed coil lpLCR domains meet at the center of the Nsp1 ring and form a high-density region. Comparing the behavior of the system in the case of wild type and mutants shown in Figure 2 demonstrates that mutating the charged residues in the lpLCR to alanine (negative control) increases the density of Nsp1 molecules at the center of the ring by 25% (Figure 2A). Also, this mutation slightly decreases the movement of the cargo complex inside the ring (Figure 2B). On the other hand, in the positive mutation where four alanines are mutated to charged residues, the Nups form a more even distribution inside the ring, implying that they have more cross-interactions rather than being trapped in

aggregated configurations. This change in Nup behavior led to a 25% decrease in the density of Nups at the center of the ring (Figure 2A). Also, the cargo has slightly more freedom of movement compared with the wild-type system (Figure 2B). FG motif density maps show the same trend as Nups density maps (Figure 2C) since FG motifs are abundantly present along the length of Nsp1 Nups. The density map of the cargo hydrophobic spots shows the same trend as the cargo itself, as they are attached to it (Figure 2D). The density maps of FG motifs and cargo hydrophobic spots are presented, as they bind to each other to facilitate the transport process.

Nup42

Nup42 is the representative of the second category. It is one of the medium-length FG Nups in yeast NPC. This Nup has a short charge-rich domain and a long lpLCR domain. Over the course of the simulation, the lpLCR domain of Nup42 forms a collapsed coil conformation, and the short charge-rich domain forms a small relaxed coil conformation that keeps the globular lpLCR domain shortly distant from the scaffold (see Figure 1, Supplemental Figure S1A, and Supplemental Video S1). The presence of the relaxed coil domains provides more freedom of movement for the lpLCR domains and allows cross-interaction between the Nup42 molecules (Supplemental Figure S1A). In all three simulations (wild type and two controls), the Nups form cross-interacting aggregates near the scaffold (Supplemental Figure S1A) and the cargo complex explores the remaining space in the ring complex. In the negative control, the lpLCR domains form highly dense conformations, and the cross-interactions between the Nups decrease as they get trapped into local interactions with their neighboring molecules (Supplemental Video S2). The maximum density of Nup molecules increases by a significant 100% value. The cargo's movement is a little less uniform compared with the wild-type system (Supplemental Figure S1B). In contrast, in the positive control the added charged residues cause the lpLCR domain to form a more even distribution compared with wild type (Supplemental Video S3). The maximum density of the Nup molecules is decreased by ~50%, which means that the Nups are more evenly distributed. The mutated Nup42 molecules form dynamic cross-interactions in which the Nups easily bind and unbind. The FG motif density maps follow the same trend as the Nup density maps (Supplemental Figure S1C). The density map of the hydrophobic spots on the cargo shows that the presence of these elements are uniform across the ring except for a high density in one spot (Supplemental Figure S1D). In the negative control, three high-density spots exist that are very close to the location of high density of FG motifs (Supplemental Figure S1C). This implies that the hydrophobic spots interact with FG motifs at these spots. In the positive control, the density map of cargo hydrophobic spots shows a high density at the center. Considering that Nup molecules are located toward the scaffold, a high density of cargo hydrophobic spots at the center means that the cargo and its hydrophobic spots spend most of the time traveling between Nup42 molecules rather than interacting with a few of the Nups for a long time.

Nup145 is another Nup in this category that is also of medium length. The Nup145 lpLCR domain is shorter than that of Nup42, but its non-lpLCR domain is longer. The density maps of Nup, cargo, FG motifs, and cargo hydrophobic spots presented in Supplemental Figure S2 show that the behavior of Nup145 is similar to that of Nup42.

Nup49

Nup49 is the shortest FG Nup in the yeast NPC. Nup49 has seven positively charged residues and no negatively charged residues, so
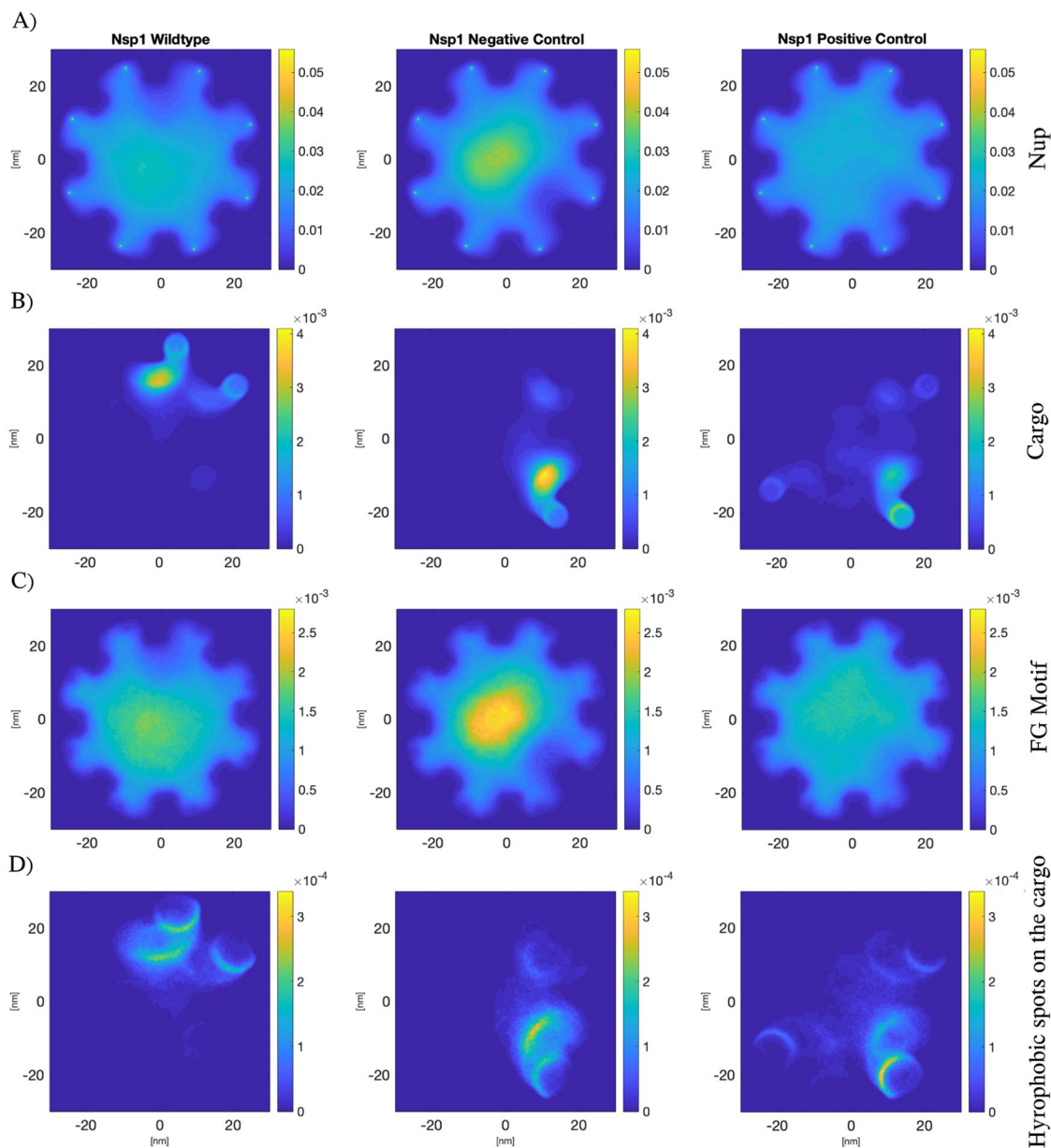
FIGURE 2: Density maps of (A) Nups and (B) cargo (C) FG motifs (D) hydrophobic spots on the cargo complex, inside ring of Nsp1. The plots show wild type, negative control, and positive control Nsp1 simulation results from left to right. In the negative control charged residues in lpLCR are mutated to alanine and in the positive control charged residues are added to the system by mutating four alanines to lysine. Nsp1 has a charge-rich domain (residues 177–617) and an lpLCR domain (residues 1 to 176), which contains only four positively charged residues (see Figure 1). The lpLCR domain forms a collapsed coil conformation and the non-lpLCR domain forms a relaxed coil conformation. The collapsed lpLCR domains meet at the center of the ring and form a highly aggregated conformation at the center (see A). In the negative control, density of Nsp1 molecules at the center increases by 25% while in the positive control, this density decreases by 25%. Also, movement of cargo becomes slightly more limited in the negative control, while the cargo gains slightly more freedom of movement in the positive control (see B). Density map of FG motifs (C) shows a similar pattern to Nups because Nsp1 are abundant in FG motifs their entire amino acid sequence. Density map of cargo hydrophobic spots (D) shows a similar pattern to that of cargo.

its entire sequence can be considered an lpLCR (see Figure 1). The density maps of Nup, cargo, FG motifs, and cargo hydrophobic spots are presented in Supplemental Figure S3. Normal Nup 49 molecules form highly collapsed coil conformations toward the scaffold of the ring (Supplemental Figure S3A). Their very collapsed conformation and short length inhibits their cross-interaction with each other. The reason behind this behavior is the low charge den-

sity of this molecule combined with high hydrophobicity of all FG Nups. In the negative control, Nup molecules form even more collapsed conformations compared with wild-type molecules. Mutating the charged residues to alanine causes the density of the isolated Nup49 aggregates to increase by ~10%. In the positive control simulation, additional charged residues allow it to have a larger end-to-end distance and a less compact conformation, decreasing the

A)

Wildtype   Negative Control   Positive Control   B)

Time

Heatmap of punctual stress in Nsp1 (KJ/mol nm)

C)   Heatmap of punctual stress in LCR domain of Nsp1 (KJ/mol nm)

D)   Wildtype   Negative Control   Positive Control   E)

Time

Heatmap of punctual stress in Nup42 (KJ/mol nm)

maximum density of the positive control molecules by almost 20%. Density maps of cargo show that in all three cases (Supplemental Figure S3B), the cargo has a large space to explore inside the ring, but the area is larger and less dense in the negative control, wild type, and the positive control, respectively. The area negatively correlates with the Nup 49 end-to-end distance. In the positive control, Nups are less aggregated, so the cargo has less space to cover inside the ring, while in the wild type Nups are more compact, so the cargo has a relatively larger area to cover. In the negative control where Nups are most collapsed, the cargo has the largest space to explore, but spends a long time close to some of the Nups, which are shown as high-density points in the density maps. Density maps of FG motifs (Supplemental Figure S3C) show the same pattern as Nups and the density map of hydrophobic spots (Supplemental Figure S3D) show the same trend as the cargo itself.

Nup57 simulations show that the ring of Nup57 including the cargo shows similar behavior to the Nup49 ring (Supplemental Figure S4). Nup57 is the second shortest Nup in yeast NPC with no negatively charged residues and a few positively charged residues in its sequence. In terms of length, charged residue content, density, and distribution, Nup49 and Nup57 are very similar and as with Nup49, charged residues appear to decrease the strength of bonds between FG Nup and cargo while also expanding the space occupied by FG Nups (Supplemental Figure S4, A and B).

Nup1
Nup1 is in the fourth category with Nup60, and Nup159. These Nups do not contain an lpLCR and are analyzed here as a natural control for the other Nups that do contain lpLCRs. Notably, there is no lpLCR to mutate, so only the wild-type system is analyzed. All of these Nups are long and have long charge-rich domains. In both the Nup1 and the Nup 60 density maps (Supplemental Figures S5A and S6A), we see that the entire ring is covered, but neither forms high-density conformations at the center. The cargo spends some time at the center and some time around the edges (Supplemental Figures S5B and S6B). The reason for this is that FG motifs form high-density areas near the tethering point of the Nups (Supplemental Figures S5C and S6C) and hydrophobic spots on the cargo complex are also mainly present around the edges where a high density of FG motifs exists (Supplemental Figures S5D and S6D).

It is noteworthy that the FG Nups in the fourth category (Nup1, Nup60, and Nup59) are located at the extremities of the NPC. These Nups are long, but do not feature an lpLCR. Hence, they do not form high densities and the two ends of the NPC on the top and bottom are not blocked by a dense network of hydrophobic residues. However, based on the ring cross-section analysis of the other Nups, the cargo will face a dense network formed by Nup42, Nsp1, Nup116, and Nup100 on the top and Nsp1 at the bottom.

Force distribution analysis of the interaction between FG Nups
TRFDA was performed on the Nups in the ring cross-sections of the NPC with the aim to gain a deeper understanding of the effect of the presence of lpLCRs in FG Nups and how mutating lpLCRs can influence the force distribution within the Nups.

Nsp1
The results of TRFDA on Nsp1 molecules within Nsp1 containing ring cross-sections are presented in Figure 3. Heatmaps show the punctual stress of each residue along the length of the simulation (Figure 3A). Punctual stress is representative of the sum of all the interaction forces acting on each residue, which allows for identifying "hotspots" in the protein (see *Materials and Methods* for a full explanation of the term punctual stress). Time and residue numbers are represented in the x and y axes, respectively. Heatmaps of punctual stress are drawn for the wild type and the negative and positive controls, and a schematic of charge distribution in each case is drawn next to it. In the charge distribution schematic, the little blue

FIGURE 3: TRFDA of Nsp1 and Nup42 molecules. (A) Heatmaps of punctual stress in Nsp1 molecules caused by the interaction of Nsp1 molecules with themselves and with each other. The y-axis represents residue number (1 to 617 from top to bottom) and The x-axis represents time. A schematic of charge distribution of each case is drawn next to their heatmap (small blue lines are positively charged and small red lines are negatively charged residues). The residues in the non-lpLCR domain (177–617) experience on average less punctual stress, while the residues in the lpLCR (1–176) experience on average higher punctual stress. In the wild type and the two controls, the same residues experience a high punctual stress, but the magnitude of punctual stress is higher to lower in the negative control, wild type, and positive control, respectively. (B) Time-averaged punctual stress for the residues of wild type, negative control, and positive control compared with each other. All three cases show similar punctual stress values in the non-lpLCR domain, while the values differ in the lpLCR domain. The negative control has the highest punctual stress and the positive control has the lowest punctual stress, and the wild type is in the middle of the two. The difference in punctual stress of the three cases is presumably due to the difference in their conformational ensemble. The negative control has the most aggregated/collapsed conformation and the residues are experiencing the highest punctual stress, while the positive control has the least aggregated/collapsed conformation and the residues are experiencing the least punctual stress, and the wild type is between the two. As can be seen both in the heatmaps and in the time-averaged graphs, the spikes of punctual stress happen at the same residues for the wild type and the mutants. This implies that mutating the number of charged residues in the lpLCR domain does not change the network of interacting residues; it only affects the intensity of the interaction. (C) Zoomed-in view of heatmap of punctual stress for the lpLCR domain of Nsp1 (residues 1–176). The charge distribution of each case is shown next to them (blue lines are positively charged residues). The residues/group of residues experiencing high punctual stresses are shown next to the heatmaps. FG motifs, phenylalanines, and only a few leucines and two glutamines are experiencing a high punctual stress. Therefore, interacting residues are mostly Phe that form a hydrophobic network. In the negative control, there are no charged residues to disturb the highly dense hydrophobic network. In wild type, the presence of four charged residues in the lpLCR domain partially interfered with the hydrophobic interactions and decreased the intensity of interaction. This effect is enhanced in the positive control, and while the same residues are interacting, their punctual stress is decreased as there are more charges to disturb the strong hydrophobic interactions. (D, E) TRFDA of Nup42 molecules. The y-axis represents residue number (1 to 382 from top to bottom) and the x-axis represents time. The lpLCR domain is residue 1–317 and the non-lpLCR domain is residue 318–382. The format of the figure and the results are similar to that of Nsp1.

	Nsp1(lpLCR domain)	Nup42(lpLCR domain)	Nup49(lpLCR domain)
Wild type	4.08 + –0.12 nm	3.95 + –0.24	3.23 + –0.11
Negative control	3.25 + –0.11 nm	3.10 + –0.31	2.38 + –0.08
Positive control	4.55 + –0.18 nm	5.62 + –0.29	4.43 + –0.33

TABLE 1: Average distance between Phe residues in the lpLCR domain of Nsp1, Nup42 and Nup49 molecules.

lines represent positively charged residues and the little red lines represent negatively charged residues. The time-averaged value of the punctual stress for each residue is drawn in Figure 3B. The wild-type, negative control, and positive control graphs are drawn in the same plot. As both the heatmaps and the time-averaged plot show, punctual stresses are higher in the lpLCR domain (residues 1–176), indicating that residues in the lpLCR domain generally interact more intensely with each other. lpLCR domains have low charge and form the collapsed coil domain in Nsp1 (see Figure 1). The TRFDA results demonstrate that the denser formations allowed by a low charge density make for stronger, more frequent interactions between residues. On the other hand, the charge-rich domains form relaxed coil conformations in which residues have less interaction within and between the Nups.

As the time-averaged plot depicts, wild type and the mutants have the same average punctual stress in the non-lpLCR domain (residues 176–617) but the values differ in the lpLCR domain. The average punctual stress values are ordered from higher to lower in the negative control, wild type, and positive control. The heatmaps show a similar trend, with the negative control having the highest punctual stress, the positive control the least, and the wild type in the middle. This observation is in line with the behavior observed in Figure 2A. Negative control molecules have the most collapsed conformation and experience the highest punctual stress, while the positive control molecules have the least aggregated conformation and the Nups experience the least amount of punctual stress.

Interestingly, our results show that the same residues experience a high punctual stress for each of the wild-type and mutant experiments, but the magnitude of punctual stress is different in the three systems (Figure 3C). The highly stressed residues are mostly FG motifs and phenylalanines along with a few leucines and one case of a pair of glutamines located next to each other. This is the case for all the eight copies of the Nsp1 molecules in the ring and the wild type and the two mutants. This is a significant observation in the mechanism of interaction of Nsp1 molecules with each other and implies that changing the number of charged residues affects the intensity of interaction of residues in the lpLCR domain, but does not change the network of residues interacting with each other.

Nup42
The results of TRFDA on Nup42 molecules within Nup42 containing ring cross-sections are presented in Figure 3, D and E. The charge distribution of each case is schematically shown next to it. Nup42 has a short non-lpLCR domain (residues 318–382) and a long lpLCR domain (residues 1–317). Similar to the trend seen in Nsp1 (Figure 3, A–C), lpLCR domains show a larger punctual stress than non-lpLCR domain in Nup42 and its two mutants (Figure 3, D and E). Also, similar to Nsp1, the negative control shows the highest punctual stress and the positive control shows the lowest punctual stress. This is in line with the conformation of Nup42 molecules shown in Supplemental Figure S1A. The negative control has the most collapsed conformation and allows for stronger interactions to form within and between molecules. The positive control has the least compact conformation and the least punctual stress.
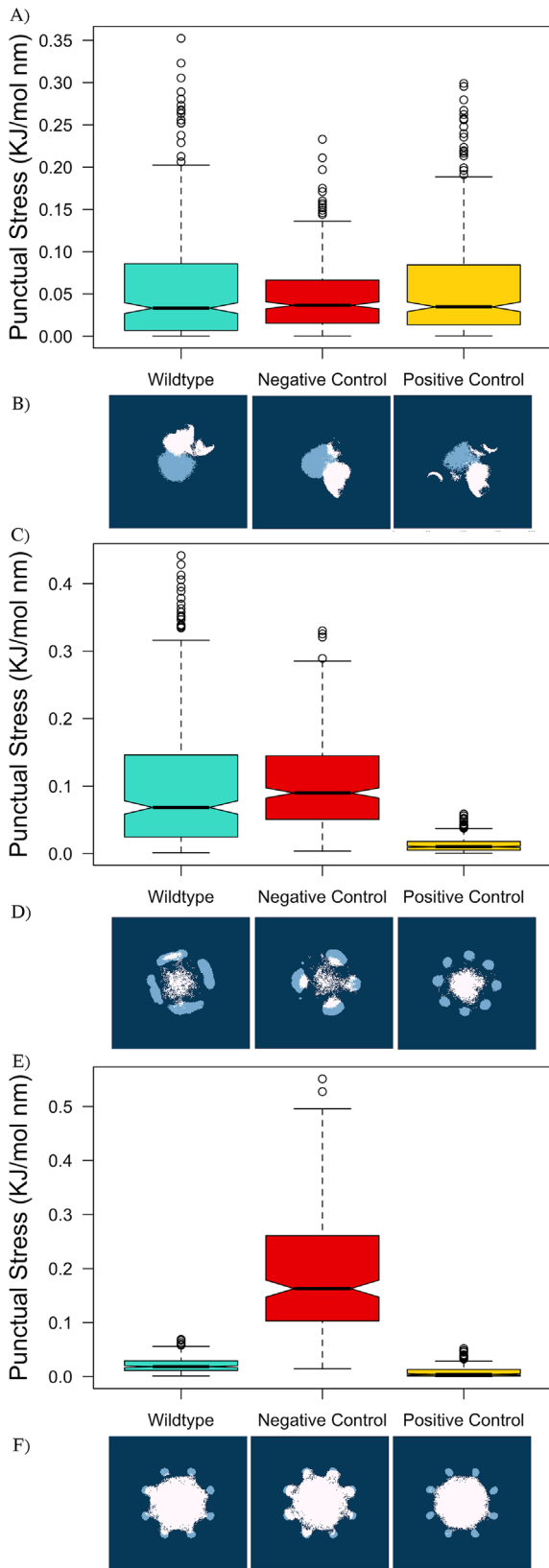
The interesting point is that in the case of Nup42, the wild type and negative control case show similar behavior while the positive control is significantly different. The residues experiencing a high punctual stress are similar in wild type and negative control, with the magnitude of punctual stress higher in the negative control. However, some of those residues experience low punctual stress values. The reason might be due to the fact that with adding the charged residues to the lpLCR domain, the charge distribution in the positive control is very similar to the non-lpLCR domain. Therefore, the positive control depicts a significantly different behavior compared with wild type and negative control. However, the interacting residues (with high punctual stress) are still the same in all three cases and are mainly phenylalanines with few other hydrophobic residues. Therefore, similar to Nsp1, the network of interacting residues has not changed, but the magnitude of punctual stress has.

Nup49
The results of TRFDA on Nup49 containing ring cross-sections are presented in Supplemental Figure S7. Nup49 does not have a charge-rich domain, so its entire length is an lpLCR (see Figure 1). The trends are similar to those seen in Nsp1 and Nup42, with the negative control having the highest punctual stress due to higher aggregation and the positive control having the lowest punctual stress due to lower aggregation. Similar to Nsp1 and Nup42, certain residues/ (FG motifs, phenylalanines, and a few other hydrophobic residues) experience a high punctual stress with the punctual stresses being the highest in the negative control and lowest in the positive control. Therefore, in the case of Nup49 as well, number of charged residues has not changed the network of interacting residues, but has affected the magnitude of punctual stress.

Considering that in all the Nups studied here, Phe residues were the main interacting residues, and the average distance of them was compared in wild type and mutant in each ring cross-section. The results are presented in Table 1 and show that the average distance of Phe residues decreases in the negative control and increases in the positive control. This implies that the number of charged residues in the lpLCR regulates the spacing between the Phe residues and influence the force distribution of Nups in this way. Additionally, the radius of gyration for wild type and control cases was calculated in the lpLCR domain of Nups. Mutating the charged residue in the lpLCR domain to alanine reduces the radius of gyration of the lpLCR domain and adding charged residues increases its radius of gyration (Supplemental Table S1). Further analysis was performed to investigate whether the distance between Phe residues is correlated with the forces that the residues experience. At each timestep of our simulations, we calculated both the average distance between Phe residues and the average force experienced by Phe residues. Calculated over all timesteps, the Pearson product–moment correlation coefficient gives the correlation between the two values. The results for Nsp1 are presented in Supplemental Table S2. The negative correlation values indicate that for all simulations, force tends to increase as intradistances decrease. The greater magnitude of the correlation for negative control indicates that in its dynamics Phe residues are forced apart more directly when they draw close. The

A)

Punctual Stress (KJ/mol nm) — Wildtype, Negative Control, Positive Control

B)

C)

Punctual Stress (KJ/mol nm) — Wildtype, Negative Control, Positive Control

D)

E)

Punctual Stress (KJ/mol nm) — Wildtype, Negative Control, Positive Control

F)

above analyses indicate that the number of charged residues in the lpLCR regulates the spacing between the Phe residues and as a result, govern the force distribution and dynamics of FG Nups.

Further statistical analysis has been done on the TRFDA data (Figure S8 and Figure S9). Please refer to the Supplemental Material for an explanation of the methods and results.

How does mutating the lpLCR domain affect the interaction between FG Nups and the cargo complex?

The results presented so far are on the interaction of Nups with each other. In this section, the effect of mutation of lpLCR on the interaction of FG Nups with the cargo complex is discussed.

Using TRFDA, punctual stress of Nups caused by the interaction of Nup molecules with the cargo complex is calculated. It is worth noting that these punctual stresses are much smaller than the punctual stresses caused by the interaction of Nups with each other. The reason is that the number of interactions happening between the residues of the Nups is much larger than the number of interactions between the Nups and the cargo complex. The boxplot of the punctual stress values for wild type and the two mutants of Nsp1 are shown in Figure 4A. The boxplot depicts similar punctual stress values in the three cases, meaning that lpLCR mutations have not significantly affected the interaction of Nsp1 molecules with the cargo complex. Combined density maps of FG motifs and hydrophobic spots on the cargo are shown in Figure 4B with blue and white dots, respectively. For the sake of clarity of the image, the points with low-density maps were set to zero density. Colocalization of white and blue dots can imply long interactions between FG motifs and cargo hydrophobic spots. Density maps in Figure 4B are not significantly different as well.

The same analysis is done for the interaction of Nup42 molecules with the cargo complex (Figure 4, C and D). The results show that the interaction of cargo complexes with wild-type Nup42 molecules becomes slightly stronger in the case of the negative control, but these interactions become significantly weaker in the case of the positive control. The density maps (Figure 4D) are in line with the boxplots. Mutating the lpLCR domains by adding charged residues significantly affects the interaction of Nup42 molecules with each

significantly different. (B) Density maps of FG motifs (shown with blue dots) and cargo hydrophobic spots (shown with white dots) are drawn together. Colocalization of white and blue dots can imply interaction between FG motifs and cargo hydrophobic spots. Therefore, interaction of Nsp1 molecules and hydrophobic spots on the cargo is not significantly affected by mutating the lpLCR domains.
(C, D) Interaction of Nup42 molecules with the cargo complex. The format of the plots is similar to plots for Nsp1. (C) The boxplots of the negative control are slightly higher than wild type, but the boxplot of positive control is significantly lower than wild type. (D) Blue and white dots depict colocalization in wild type and negative control, but do not show any colocalization in the positive control. This implies that interactions between Nups and cargo are much weaker and more transient in the positive control compared with the other two cases.
(E, F) Interaction of Nup49 molecules with the cargo complex. (E) The average punctual stresses in the negative control are 10 times higher than wild type, while they are three times lower in the positive control compared with wild type. (F) Blue and white dots depict no colocalization in the positive control, some colocalization in wild type, and a lot of colocalization in the negative control. These are in line with boxplot results. Therefore, interactions between Nups and cargo are the most weak and transient in the positive control and stronger in wild type. Very strong interactions in the negative control might interfere with the transient interactions necessary for nucleocytoplasmic transport.

FIGURE 4: Interaction of Nsp1, Nup42 and Nup49 molecules with the cargo complex. (A) TRFDA was performed to find the punctual stress experienced by Nsp1 molecules caused by the interaction of Nsp1 with the cargo complex. Boxplot of time-averaged values of punctual stress in all 617 residues of all copies of Nsp1 is drawn for wild type and two controls. The boxplots of the three cases are not

other and the interaction of Nup42 molecules with the cargo complex. The interactions become more transient and less strong.

The Nup49 ring complex was also similarly analyzed. Results in Figure 4, E and F show that interaction of Nup49 molecules with the cargo complex is largely affected by mutations to the lpLCR domain. The average punctual stress in wild type is three times larger than the positive control and the average punctual stress in the negative control is 10 times larger than the positive control. The combined density maps show how this occurs (Figure 4F). The hydrophobic spots on the cargo are colocalizing with FG motifs to a large extent in the negative control, but showing no colocalization in the positive control. The results imply that mutation of lpLCR domains in Nup49 affects the conformation of molecules (Supplemental Figure S3A) and has a major impact on the interaction of the cargo complex and Nup49 molecules.

Our results suggested a nonuniversal effect of changing the lpLCR charge density on the interaction of FG Nups with the cargo complex; whereas in some cases the effect was significant, in other cases changing the lpLCR charge density didn't appreciably impact the interaction of cargo with the FG Nups. While a general rule cannot be made with the few datapoints presented here, based on the current data, one can postulate that the effect of mutations on different Nups may vary based on the ratio of the length of the lpLCR domain versus the length of the disordered domain of each Nup. In the case of Nsp1, the lpLCR domain is almost 25% of the length of the disordered domain of the Nsp1, and mutations in the lpLCR domain do not have a significant influence on the interactions between Nsp1 molecules and the cargo complex. In the case of Nup 42, on the other hand, the lpLCR domain covers some 80% of the disordered domain of this Nup. Mutating the charged residues to alanine in the lpLCR domain slightly increases the interactions between the Nups and the cargo complex and adding charged residues to the lpLCR domain decreases the interactions. We can speculate that since the lpLCR domain covers a large portion of the sequence, both removing and adding charged residues affect the interaction of Nup42 molecules with the cargo complex. The same is true for Nup 49. Since the entire length of Nup 49 is an lpLCR, both removing and adding charged residues impact the interaction of Nup 49 molecules with the cargo complex.

A common observation between all the cases of all the rings analyzed is that the conformational ensemble of Nups does not significantly change with the presence of a cargo complex in the system. This can be concluded from comparison of our previous study (Peyro et al., 2015a) with the density maps of Nups in Figure 2 and Supplemental Figures S1 and S3. The density maps look very similar in the presence and absence of a cargo complex. This behavior was expected based on comparison of the punctual stress values caused by the interaction of Nups with each other and the values caused by the interaction of FG Nups with the cargo complex, which shows that the Nup–Nup interactions are significantly dominant.

It is worth noting that in all of the simulations in this study and our previous study (Peyro et al., 2015a), the general behavior of FG Nups is in line with the behavior observed in Sakiyama et al. (2016) via high-speed AFM. Although the timescale of the observations are different, the average behavior is similar. In line with their observation, our simulations depict a highly dynamic and flexible behavior by FG Nups, which go through transient entanglements yet do not form a cohesive network.

Whole-NPC results
Distribution of FG motifs in the NPC.
A model of the whole NPC was created including all of the comprising FG Nups, the scaffold of
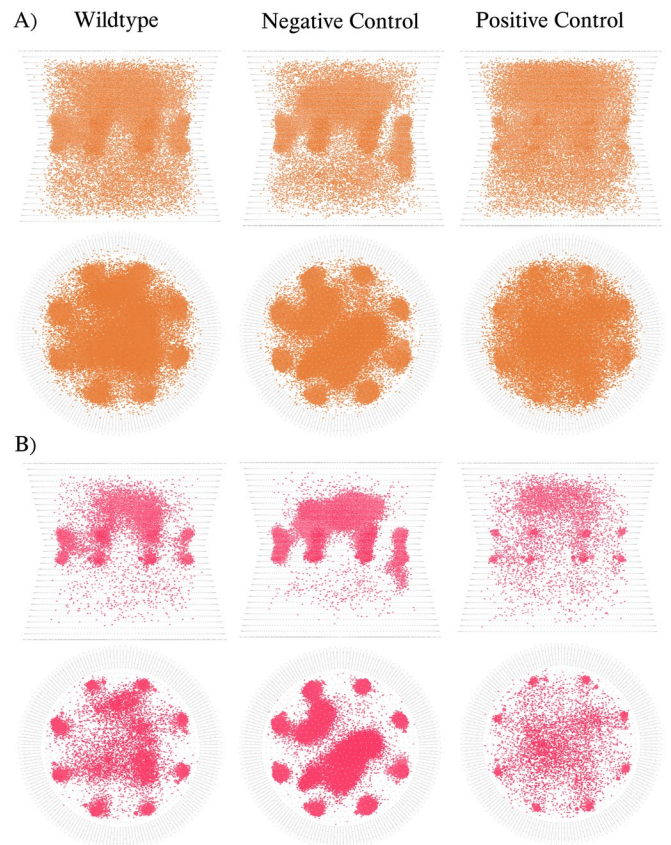
FIGURE 5: FG density maps of the whole-NPC model for wild type and two controls for density threshold of 1.5% of maximum density (A) and 2% of maximum density (B). The side view of the NPC is presented at the top row and the top view is presented at the bottom row. Similar to the behavior observed in ring cross-sections, the negative control makes the FG motifs more clustered, and the positive control makes them more evenly distributed within the space inside the NPC. Maximum density of FG motifs increases by 26% in the negative control and decreases by 15% in the positive control. Both the wild-type system and the negative control depict high-density FG clusters in the system, but the 15% decrease in density affects the shape of clusters significantly and scatters them remarkably. Comparing the results with the system excluding the cargo complex (Peyro et al., 2015a) shows that the FG network of the wild-type system behaves almost similarly in the absence or presence of cargo, but this behavior is different in the case of the negative control. In absence of cargo complex, FG motifs in the negative control form a very high-density cluster at the center. In the presence of cargo, FG motifs cluster density increases, but the presence of cargo at the center of the system prevents FG networks from forming a very dense cluster at the center. Rather, they form high-density clusters with an opening at the center.

the NPC, and a 10-nm spherical cargo accompanied by seven hydrophobic spots that represent Kap. Initially, the cargo complex was located at the cytoplasmic side of the NPC, but it was not able to pass through the pore over the course of the 100-ns simulation. The cargo complex would penetrate the hydrophobic FG network to a small extent but was repelled back and could not reach the nucleoplasmic side. In reality, the transport process happens in the scale of milliseconds; however, due to the very high computational expense of simulating such a large system of molecules, we are not able to reach those timescales. Therefore, we tried to accelerate the transport process by helping the cargo complex overcome the

energy barrier of the high-density network of FG Nups that exists toward the cytoplasmic side of the NPC. For this purpose, we initially positioned the cargo at two-thirds height of the NPC and ran the simulation after minimizing the energy of the system. With this initial condition, the cargo was able to pass through the pore. This intervention allowed us to see the transport over the course of the simulation. This initial condition was kept the same for all of the whole-NPC simulations, either wild type or mutant. We repeated the simulation for negative and positive controls by replacing the wild-type Nups with negative and positive control Nups, respectively, and in both of these cases, the cargo complex was able to reach the nucleoplasmic side. FG density maps of wild type and two control systems are presented in Figure 5, A and B for two density thresholds of 1.5 and 2% of maximum density, respectively.

Based on observations in Figure 5, the whole NPC shows a similar behavior to ring cross-sections when lpLCR domains are mutated. The FG domains become more concentrated in the negative control and less concentrated/more evenly distributed in the positive control. The maximum density of FG motifs in the model increases by 26% in the negative control and decreases by 15% in the positive control. The interesting observation is that in both wild type and negative control, a high-density cluster of FG motifs is formed at the top-central region of the NPC, but the 15% decrease in density of FG motifs significantly affected the high-density FG network formed toward the center of the pore and scattered them in the rest of the space inside the NPC.

Results of Figure 5B can be compared with Figure 8 in Peyro *et al.* (2015a), where the whole NPC is modeled in the absence of any cargo complex. In the absence of a cargo complex, the wild-type system showed a donut-shaped cluster of FG motifs toward the upper side of NPC, while the negative control showed a high-density network at the center of the NPC that blocked the pore. The wild-type system including the cargo complex shows a similar behavior (Figure 5B, left figure). However, the negative control system with a cargo complex (Figure 5B, middle figure) behaves differently from the system in the absence of cargo. The system including the cargo does form high-density FG clusters, but they have an opening at the center caused by the cargo complex. As will be shown in Figure 6, in the wild-type and negative control systems, the cargo spends most of its time at the center of the NPC. The presence of several FG Nups filling up the space inside the NPC makes the cargo keep toward the central area, and this prevents FG motifs from forming a cluster that blocks the center of the NPC. This observation implies that while the FG motif network is more clustered in the negative control, the presence of even one cargo complex would prevent FG Nups from collapsing into high-density conformations at the center. The difference in the positive control seems to be more fundamental though, as FG clusters become very scattered and significantly decrease in density.

Movement of the cargo complex inside the NPC. Movement of the cargo complex in the whole-NPC model was analyzed and compared between wild type and the two controls. The density map of the center of mass of the cargo in wild type and the two control cases was drawn in Figure 6. Green dots represent the center of mass of the cargo in the wild-type system, red dots represent the center of mass of the cargo in the negative control system, and blue dots represent the center of mass of the cargo in the positive control system. It is clearly visible in Figure 6 that the cargo has more freedom of movement in the positive control (shown with blue) and less freedom of movement in the negative control (shown with red). These results are in line with our observations in the ring cross-sections of the NPC.
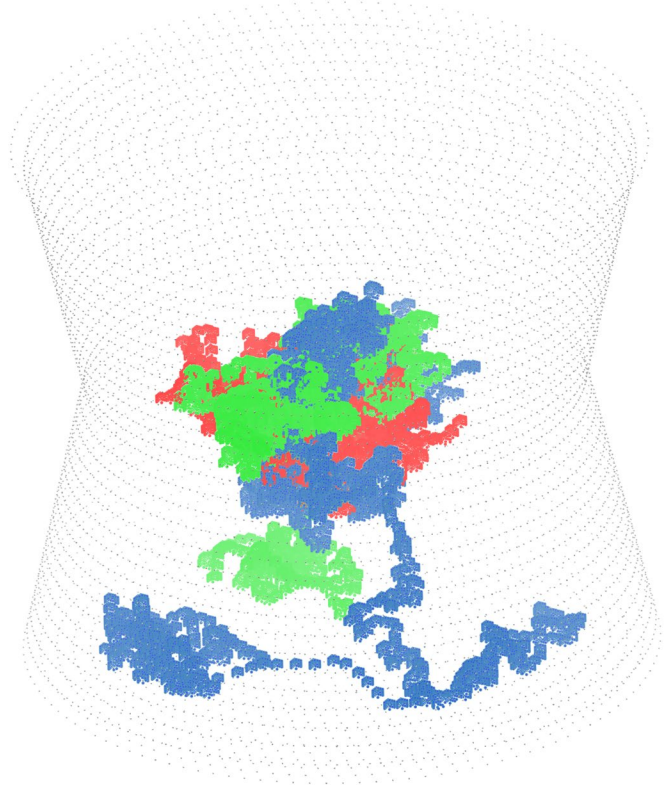

FIGURE 6: Density map of the center of mass of cargo. Green dots represent cargo in the wild-type system, red dots represent cargo in the negative control system, and blue dots represent cargo in the positive control system. A significant difference is observed in the movement of cargo in the three different systems, with the cargo in the negative control system sweeping a smaller space in the system and the cargo in the positive control system sweeping a larger space. This observation correlates with results in Figure 5, where in the negative control, FG motifs are more clustered and in the positive control, they are more scattered.

Conformational ensemble of each FG Nup layer inside the NPC

Here the density map of each layer of FG Nups is visualized in the whole-NPC model. The side view and the top view of the density map of each layer of FG Nup is presented in Figure 7. The density threshold is set to 1.5% of maximum density. The Nups are ordered in the way they appear in the NPC from the cytoplasmic side to the nucleoplasmic side (Alber *et al.*, 2007) (ordered from left to right and top to bottom in Figure 7). Some of the FG Nups like Nsp1 appear at multiple layers in the NPC, so each of the different layers are drawn separately. The side-view results show that Nsp1 copies and Nup116 travel far from their tethering point within the NPC. The other FG Nups stay within their tethering point region. However, the density maps of almost all of FG Nups overlap in terms of their location within the NPC.

Density maps of different layers of FG Nups are drawn in Supplemental Figure S10 with a 3% density threshold to compare with results from isolated ring cross-sections. Slight differences exist in density maps of Nsp1 copies and Nup116 rings in the whole-NPC model and ring cross-sections. The reason might be because the whole NPC is a crowded system and the movement of the cargo complex happens slowly (compared with the simulation time). Therefore, the cargo spends a long time at each layer and causes Nsp1 and Nup116 molecules not to fully cluster at the center. It is
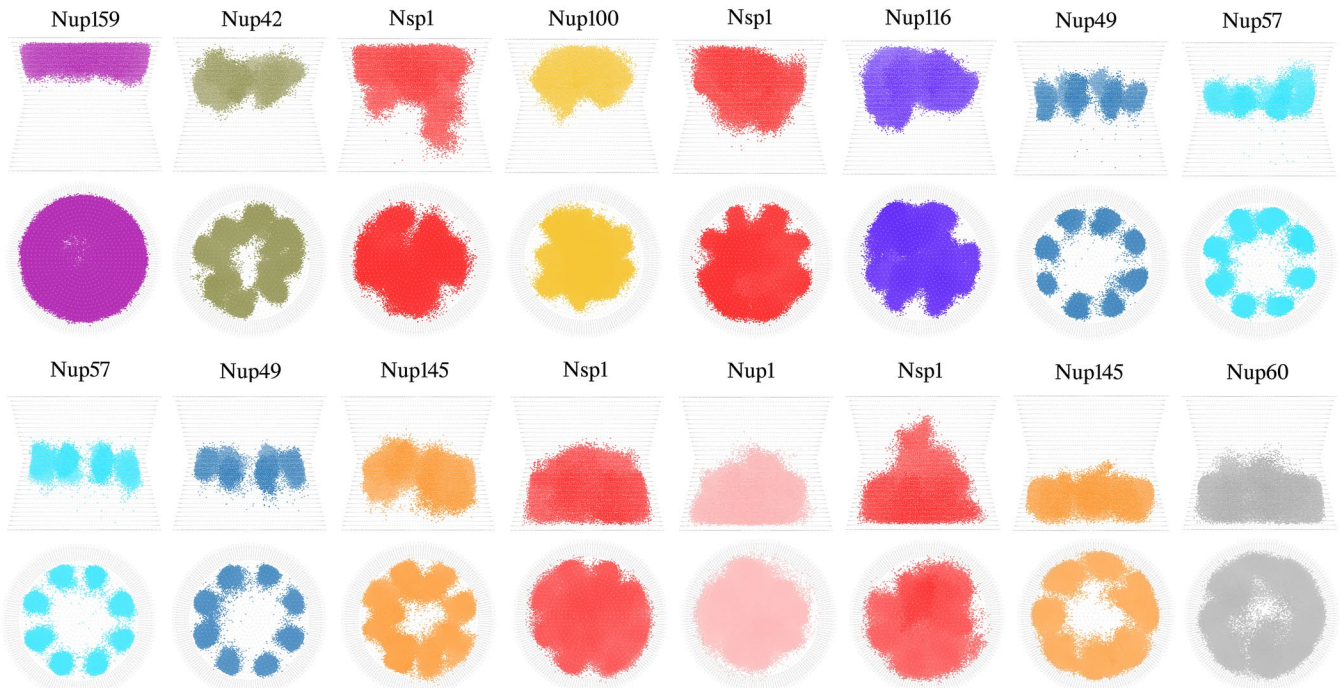
FIGURE 7: The density map of each layer of FG Nups in the whole-NPC model. Density threshold is set to 1.5% of maximum density for all the cases. The density maps are ordered in the same manner they appear in NPC (Alber *et al.*, 2007). The density maps from left to right and from top to bottom correspond to layers of FG Nups in yeast NPC from the cytoplasmic side to the nucleoplasmic side. Nsp1 copies and Nup116 travel far from their tethering point and explore a large space inside the NPC. The other Nups stay within a short distance from their tethering point, but the spaces of the FG Nups sweep overlap with each other.

also interesting to note that while the other FG Nups are located close to their neighboring FG Nups and share the space with them, their conformational ensemble is not influenced by the presence of neighboring molecules and looks similar to their isolated ring behavior.

DISCUSSION
Overview
The role of FG Nup sequences in NPC function is still not fully understood, but there is compelling evidence that their function in nucleo-cytoplasmic transport depends on lpLCRs. In this study, we employed a biophysics approach to analyze the effect of this specific feature. To make the analysis tractable, the NPC was analyzed at the levels of individual FG Nups, ring cross-sections, and the whole NPC, with all models derived from the yeast NPC. For all FG Nups, in addition to modeling their baseline properties, we also evaluated their interaction with potential cargo. Our results suggest that lpLCRs are differentiated from the remainder of the FG Nup by both their conformation and the forces their residues experience, and that these differences are necessary for the selective transport of cargo.

FG Nups have been shown to be different from other disordered proteins in terms of sequence composition (Peyro *et al.*, 2021). Comparing FG Nups with other disordered proteins (Peyro *et al.*, 2021), FG Nups have almost the same percentage of hydrophobic amino acids, which are considered to be order-promoting (Campen *et al.*, 2008), while they have considerably fewer charged residues, which are considered to be disorder-promoting (Campen *et al.*, 2008). Here ordered and disordered proteins refer to the presence or lack of secondary structures in the protein's three-dimensional conformation. Our recent study shows that lpLCRs are exclusive features of sequences of FG Nups that differentiate FG

Nups from DisProt proteins (Peyro *et al.*, 2021). Removing the lpLCR domains from their sequences, FG Nups would feature charge distributions that appear identical to those in other DisProt proteins.

Based on our previous study (Peyro *et al.*, 2015a) and our current simulations, the non-lpLCR domains of FG Nups form relaxed coil conformations while the lpLCR domains form a collapsed conformation. While all domains lack structure throughout, the presence of extra order-promoting and fewer disorder-promoting residues in the lpLCR domains causes them to form a collapsed conformation while the non-lpLCR domains form less dense relaxed coil conformation. Our new results further prove the link between lpLCR charge distributions and their ensemble conformation and go on to demonstrate how the length and charge of lpLCRs influence interactions with cargo in the NPC. Based on the difference in sequence composition between lpLCR versus non-lpLCR domains and their correlation with conformation of these domains, here in the *Discussion* we call the collapsed and aggregated conformation less "disordered" and the relaxed open conformation more "disordered."

lpLCR ensemble conformation
Overall, our results from the wild-type and control simulations show the differences charge distributions in the lpLCR make toward Nup–Nup interactions, Nup–cargo interaction, movement of cargo complexes inside the NPC, and the characteristics of the FG motif network.

In the case of Nup–Nup interactions, despite the variety of lengths and sequence structures within our FG Nups we observed similar patterns in the difference between the wild-type and the mutated rings for all of the yeast FG Nups. Mutating the charged residues in

the lpLCR domain to alanine in our negative control experiment consistently collapsed the lpLCR, allowing Phe residues to interact more closely and form a stronger hydrophobic network. These stronger individual interactions are demonstrated by the increase in forces on these residues, as well as by the increase in the variance of these forces. Stronger interactions are also harder to break, and in all negative control experiments, the Nups get trapped into highly aggregated conformations and can hardly overcome the energy barrier to break the bonds and interact with new molecules. Therefore, the system transforms into a less "disordered," frozen state.

For our positive control, mutating some of the alanines in the lpLCR domain to positively charged residues also consistently produced similar results. The lpLCR domain becomes less collapsed/aggregate, while Phe residues become separate and form weaker bonds with each other. Therefore, the punctual stress in Phe residues decreases. Weaker interactions in the positive control are easier to break; hence, the Nups easily bind and unbind and can more easily cross-interact with new neighboring molecules. Therefore, the system transforms to a more "disordered" dynamic network of residues, which we can see from the random and widely spread distribution of the FG Nup residues for positive control. In general, the number of charged residues in the lpLCR regulates the "disorderedness" of the network of FG Nups and by extension the degree of interaction inside the network of FG Nup Phe residues. In wild-type and negative control lpLCR domains, we see that hydrophobic interactions dominate, indicating that the residues form a dense hydrophobic network, while in non-lpLCR domains, or positive control lpLCR, electrostatic repulsion dominates against hydrophobic interactions and the Nups do not collapse into aggregated conformations.

Mutating the lpLCR domains had a similar effect on the whole-NPC model as well, confirming that the results observed in ring cross-sections are preserved in the higher order system. While wild-type and negative control systems form high-density FG networks toward the central area of the NPC, the density of FG motifs in the same area is decreased for the positive control to an extent that the network of FG motifs is significantly influenced. Additionally, mutating the lpLCR domains has a remarkable effect on the movement of cargo complexes inside the NPC. As with the ring cross-section results, cargos have limited movement in the negative control and much more freedom in the positive control.

Cargo interactions
After confirming that lpLCR dynamics are regulated by charge distribution, we went on to verify that charge distribution also regulates cargo transport through these dynamics. However, differences in these cases were more qualified and depended on the circumstances of the FG Nup being tested.

Our results showed that the density of charged residues may or may not affect the interaction of the cargo complex with FG Nups, likely depending on the relative length of the lpLCR versus the full length of the nucleoporin sequence. That is, if the lpLCR domain covers a larger portion of the length of the disordered nucleoporin sequence, it is more likely that mutating the charged residues in the lpLCR domain will affect the interaction of FG Nups with the cargo complex.

Interpretation
Previous studies have shown that selectivity of NPC relies on both the properties of the FG network and the interaction between FG Nups and the cargo complex (Frey and Görlich, 2007; Zilman et al., 2007, 2010; Beck and Hurt, 2017; Hayama et al., 2018). Frey and Görlich showed that a saturated network of FG motifs is required for

NPC to maintain its selective properties (Frey and Görlich, 2007). Also, the quality of the interaction between FG Nups and cargo complex plays a crucial role in selectivity of the NPC. Large cargo complexes need to bind to FG Nups to be efficiently translocated through the NPC. However, if the binding is too strong, the cargo complex may get stuck to the FG Nups and therefore block the NPC. As a result, only the binding interaction within a range of certain affinities can lead to effective transport through the NPC. Therefore, the property of interaction between FG Nups and the cargo complex is directly correlated with selectivity of the NPC.

According to our results, the existence of lpLCR patterns in the FG Nup sequences seems to be a contributing factor enabling these transient, but frequent, interactions between transport factors and FG Nups. Our TRFDA results suggest that interactions between the cargo complex and the FG Nups are weakened as the number of charges in the lpLCR domains increases, which may in turn translate into shorter interactions because weaker bonds are easier to break. Videos of simulation trajectories for Nup42 (Supplemental Videos S1, S2, and S3) qualitatively imply that in the negative control, the cargo complex interacts with the Nups for longer periods of time as compared with wild type, whereas in the positive control, the interaction times decrease and interactions are more transient as compared with wild type. These results imply that the presence of charged residues in the lpLCR domains prevent the cargo complex from "sticking" to FG Nups. At the same time, it is reported that transient interactions rely on the plasticity of FG Nups (Milles et al., 2015). Our results suggest that removing charged residues in lpLCR causes the FG Nups to form highly aggregated conformations in which FG Nups are less likely to unbind and form new interactions, effectively reducing the plasticity (viscoelasticity) of the ultrafast interactions between the cargo complex and the FG Nups. The charge distribution of lpLCRs may therefore be necessary for sufficiently fluid behavior within the NPC.

It appears that the number of charged residues in the lpLCR is optimized to keep the network of Phe residues "intimate" enough to form a dense hydrophobic barrier inside the NPC while enabling fast and transient interactions to take place between FG Nups and the cargo complex. For the NPC to correctly filter cargo, its FG network needs to bind to them consistently, but also transiently, so that they do not become stuck.

CONCLUSION
The results from this study emphasize and add to the findings from our previous study (Peyro et al., 2015a) on the importance of lpLCRs to NPC function. While only a few charged residues are mutated in each sequence, a significant change can be observed in the conformational ensemble of FG Nups and interaction of the cargo complex. This is more remarkable considering the fact that NPC is a very robust macromolecular machine (Strawn et al., 2004), which is viable even when some of the FG Nups are removed. Although our results cannot directly predict how the transport process would be affected by lpLCR mutations, the effect of mutating lpLCRs on the governing factors of transport process is predicted. The results from this study call for experimental investigations to study the effect of lpLCRs on transport. Investigating this specific sequence feature can be useful in understanding the underlying mechanism of transport as well as designing artificial nanopores.

MATERIALS AND METHODS
Coarse-grained model
A one-bead-per-amino-acid-coarse-grained molecular dynamics model was used to analyze the behavior of FG Nups and the cargo

complex in the ring cross-sections of the NPC. This coarse-grained model is designed to study intrinsically disordered proteins. Amino acids are modeled as spherical beads with a mass of 120 Da and a distance of 0.38 nm. The force field developed for this model, originally developed by Ghavami *et al.*, accounts for different biophysical factors including bending and torsion potentials between neighboring beads, an implicit solvent, ion screening effect, and hydrophobic and electrostatic interactions (Ghavami *et al.*, 2013, 2014). Langevin dynamics simulations were performed using GROMACS molecular dynamics simulation software (Pronk *et al.*, 2013). A timestep of 0.02 ps was used in simulations. In all of the simulations, only disordered domains of FG Nups were used in the simulation, which are taken from(Yamada *et al.* (2010). This model has been previously used in modeling FG Nups (Ghavami *et al.*, 2014, 2016; Peyro *et al.*, 2015a).

Single Nup simulations were performed with an isolated single Nup fixed at one end. These simulations were run for 1 μs. In the initial configuration of the rings, the cargo complex was located at the center of the ring, and FG Nups had a random conformation as the starting point of the simulation.

Visualization
Density plots of the rings were generated by discretizing the ring into $0.5 \times 0.5 \times 0.5$ nm unit cells. The number of occurrences of each element was counted over the total simulation time. Gridcount package was used to generate the ring density plots (Beckstein and Sansom, 2003). Density plots of whole-NPC simulations were extracted using GROmaps (Briones *et al.*, 2019). All visualizations were performed via VMD 1.9.3 (Humphrey *et al.*, 1996).

Cargo complex
A simple representation of the cargo complex was used in the model. A general spherical shape was chosen for the cargo as a representative of the shape of different globular molecules passing through the pore. The cargo was modeled as a group of noninteracting beads with a diameter of 10 nm; the 10 nm diameter was chosen to ensure that the cargo is transported through active transport rather than free diffusion. Transport receptor (specifically Karyopherin β) was modeled as seven hydrophobic spots (phenylalanine) on the surface of the cargo. The shape the binding spots form on the cargo and their spacing was chosen based on the experimental study done by Isgro and Schulten (Isgro and Schulten, 2007) and the computational analysis was done by Ghavami *et al.* (Ghavami *et al.*, 2016).

TRFDA
TRFDA was performed on the ring complexes using TRFDA GROMACS tool (Costescu and Gräter, 2013). Punctual stress on each residue based on the nonbonded interaction of Nup–Nup or Nup–cargo complex was calculated through simulation time. Punctual stress was calculated based on pairwise interactions between residues. The pairwise interaction is represented by a pairwise force acting on the center of mass of the two residues. The pairwise forces are calculated based on the equation below:

$$F_{ri,rj} = \sum_{i \in ri, \; j \in rj} \vec{F_{ij}}$$

where *i* is an atom of residue *ri* and *j* is an atom of residue *rj*, where *ri* and *rj* are different residues. In the case of our one-bead-per-amino-acid-coarse-grained simulations, each residue was considered to be one atom in the TRFDA analysis. Punctual stress is then calculated in TRFDA as a sum of absolute values of forces acting on each residue. Therefore, punctual stress allows to identify points where pairwise forces accumulate and find "hot spots" in the protein (Costescu and Gräter, 2013).

PCA analysis
PCA analysis was done using the in-house Python scripts and the sklearn machine learning library. The codes used for this analysis are available at https://github.com/molecular-cell-biomechanics-lab.

REFERENCES
Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, *et al.* (2007). The molecular architecture of the nuclear pore complex. Nature 450, 695–701.

Ando D, Colvin M, Rexach M, Gopinathan A (2013). Physical motif clustering within intrinsically disordered nucleoporin sequences reveals universal functional features. PLoS One 8, e73831.

Aramburu IV, Lemke EA (2017). Floppy but not sloppy: Interaction mechanism of FG-nucleoporins and nuclear transport receptors. Semin Cell Dev Biol 68, 34–41.

Azimi M, Mofrad MRK (2013). Higher Nucleoporin-importinβ affinity at the nuclear basket increases nucleocytoplasmic import. PLoS One 8, e81741.

Azimi M, Bulat E, Weis K, Mofrad MRK (2014). An agent-based model for mRNA export through the nuclear pore complex. Mol Biol Cell: Special Issue on Quantitative Biology 25, 3643–3653.

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

Beckstein O, Sansom MSP (2003). Liquid-vapor oscillations of water in hydrophobic nanopores. Proc Natl Acad Sci USA 100, 7063–7068.

Briones R, Blau C, Kutzner C, de Groot BL, Aponte-Santamaría C (2019). GROmaps: A GROMACS-based toolset to analyze density maps derived from molecular dynamics simulations. Biophys J 116, 4–11.

Brown CJ, Takayama S, Campen AM, Vise P, Marshall TW, Oldfield CJ, Williams CJ, Keith Dunker A (2002). Evolutionary rate heterogeneity in proteins with long disordered regions. J Mol Evol 55, 104–110.

Campen A, Williams R, Brown C, Meng J, Uversky V, Dunker A (2008). TOP-IDP-Scale: a new amino acid scale measuring propensity for intrinsic disorder. Protein Pept Lett 15, 956–963.

Costescu BI, Gräter F (2013). Time-resolved force distribution analysis. BMC Biophys 6, 5.

DeGrasse JA, DuBois KN, Devos D, Siegel TN, Sali A, Field MC, Rout MP, Chait BT (2009). Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. Mol Cell Proteomics 8, 2119–2130.

Frey S, Görlich D (2007). A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. Cell 130, 512–523.

Ghavami A, van der Giessen E, Onck PR (2013). Coarse-grained potentials for local interactions in unfolded proteins. J Chem Theory Comput 9, 432–440.

Ghavami A, van der Giessen E, Onck PR (2016). Energetics of transport through the nuclear pore complex. PLoS One 11, e0148876.

Ghavami A, Veenhoff LM, van der Giessen E, Onck PR (2014). Probing the disordered domain of the nuclear pore complex through coarse-grained molecular dynamics simulations. Biophys J 107, 1393–1402.

Hayama R, Sparks S, Hecht LM, Dutta K, Karp JM, Cabana CM, Rout MP, Cowburn D (2018). Thermodynamic characterization of the multivalent interactions underlying rapid and selective translocation through the nuclear pore complex. J Biol Chem 293, 4555–4563.

Hough LE, Dutta K, Sparks S, Temel DB, Kamal A, Tetenbaum-Novatt J, Rout MP, Cowburn D (2015). The molecular mechanism of nuclear transport revealed by atomic-scale measurements. Elife 4, p.e10027.

Huang K, Tagliazucchi M, Park SH, Rabin Y, Szleifer I (2019). Molecular model of the nuclear pore complex reveals a thermoreversible FG-network with distinct territories occupied by different FG motifs. Preprint doi: 10.1101/568865.

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

Isgro TA, Schulten K (2007). Cse1p-binding dynamics reveal a binding pattern for FG-repeat nucleoporins on transport receptors. Structure 15, 977–991.

Jamali T, Jamali Y, Mehrbod M, Mofrad MRK (2011). Nuclear Pore Complex. Biochemistry and Biophysics of Nucleocytoplasmic Transport in Health and Disease, New York: Elsevier.

Jovanovic-Talisman T, Zilman A (2017). Protein transport by the nuclear pore complex: simple biophysics of a complex biomachine. Biophys J 113, 6–14.

Matsuda A, Mofrad MRK (2021). Free energy calculations shed light on the nuclear pore complex's selective barrier nature. Biophys J 2021 Jul 31;S0006-3495(21)00633-0, doi: 10.1016/j.bpj.2021.07.025.

Milles S, Mercadante D, Aramburu IV, Jensen MR, Banterle N, Koehler C, Tyagi S, Clarke J, Shammas SL, Blackledge M, et al. (2015). Plasticity of an Ultrafast Interaction between Nucleoporins and Nuclear Transport Receptors. Cell 163, 734–745.

Moussavi-Baygi R, Mofrad MRK 2016. Rapid Brownian motion primes ultrafast reconstruction of intrinsically disordered Phe-Gly-repeats inside the nuclear pore complex. Sci Rep, 6, Article number: 29991.

Mulder FAA (2018). Fuzzy and fast nuclear transport. J Biol Chem 293, 4564–4565.

Peyro M, Soheilypour M, Ghavami A, Mofrad MRK (2015a). Nucleoporin's like charge regions are major regulators of FG coverage and dynamics inside the nuclear pore complex. PLoS One 10, e0143745.

Peyro M, Soheilypour M, Lee BL, Mofrad MRK (2015b). Evolutionarily conserved sequence features regulate the formation of the FG network at the center of the nuclear pore complex. Sci Rep 5, 15795.

Peyro M, Soheilypour M, Nibber VS, Dickson AM, Mofrad MRK (2021). FG nucleoporins feature unique patterns that distinguish them from other IDPs. Biophys J 120, 3382–3391.

Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, et al. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29, 845–854.

Raveh B, Karp JM, Sparks S, Dutta K, Rout MP, Sali A, Cowburn D (2016). Slide-and-exchange mechanism for rapid and selective transport through the nuclear pore complex. Proc Natl Acad Sci USA 113, E2489–E2497.

Sakiyama Y, Mazur A, Kapinos LE, Lim RYH (2016). Spatiotemporal dynamics of the nuclear pore complex transport barrier resolved by high-speed atomic force microscopy. Nat Nanotechnol 11, 719–723.

Soheilypour M, Mofrad MRK (2016). Regulation of RNA-binding proteins affinity to export receptors enables the nuclear basket proteins to distinguish and retain aberrant mRNAs. Sci Rep 6, 35380.

Soheilypour M, Mofrad MRK (2018). Quality control of mRNAs at the entry of the nuclear pore: cooperation in a complex molecular system. Nucleus 9, 202–211.

Strawn LA, Shen T, Shulga N, Goldfarb DS, Wente SR (2004). Minimal nuclear pore complexes define FG repeat domains essential for transport. Nat Cell Biol 6, 197–206.

Tagliazucchi M, Peleg O, Kröger M, Rabin Y, Szleifer I (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.

Tan PS, Aramburu V, Mercadante D, Shammas SL, Lemke Correspondence EA (2018). Two differential binding mechanisms of FG-nucleoporins and nuclear transport receptors. Cell Rep 22, 3660–3671.

Timney BL, Raveh B, Mironska R, Trivedi JM, Kim SJ, Russel D, Wente SR, Sali A, Rout MP (2016). Simple rules for passive diffusion through the nuclear pore complex. J Cell Biol 215, 57–76.

Uversky VN (2013). Unusual biophysics of intrinsically disordered proteins. Biochim Biophys Acta 1834, 932–951.

Wolf CB, Mofrad MRK (2008). On the octagonal structure of the nuclear pore complex: insights from coarse-grained models. Biophys J 5, 2073–2085.

Wolf CB, Mofrad MRK (2009). Mechanotransduction: The role of nuclear pore mechanics and nucleocytoplasmic transport. In: Cellular Mechanotransduction: Diverse Perspectives from Molecules to Tissues, ed. Mofrad and Kamm, New York: Cambridge University Press, 417–437.

Wright PE, Dyson HJ (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293, 321–331.

Yamada J, Philips JL, Patel S, Goldfien G, Calestagne-Morelli A, Huang H, Reza R, Acheson J, Krishnan VV, Newsam S, 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.

Zilman A, Di Talia S, Chait BT, Rout MP, Magnasco MO (2007). Efficiency, selectivity, and robustness of nucleocytoplasmic transport. PLoS Comput Biol 3, e125.

Zilman A, Di Talia S, Jovanovic-Talisman T, Chait BT, Rout MP, Magnasco MO (2010). Enhancement of transport selectivity through nano-channels by non-specific competition. PLoS Comput Biol 6, e1000804.