Coarse-grained molecular dynamics simulation of transport through the nuclear pore complex

Ghavami, Ali

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2014

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 18-11-2018
Chapter 3

Probing the disordered domain of the nuclear pore complex

Abstract

The distribution of disordered proteins (FG-nups) that line the transport channel of the nuclear pore complex (NPC) is investigated by means of coarse grained molecular dynamics simulations. A one-bead-per-amino-acid model is presented which accounts for the hydrophobic/hydrophilic and electrostatic interactions between different amino acids, polarity of the solvent and screening of free ions. The results indicate that the interaction of the FG-nups forms a high density doughnut-like distribution inside the NPC which is rich in FG-repeats. We show that the obtained distribution is encoded in the amino acid sequence of the FG-nups and is driven by both electrostatic and hydrophobic interactions. To explore the relation between structure and function, we have systematically removed different combinations of FG-nups from the pore to simulate “inviable” and “viable” NPCs that were previously studied experimentally. The obtained density distributions show that the maximum density of the FG-nups inside the pore does not exceed 185 mg/ml in the inviable NPCs, while for the wildtype and viable NPCs this value increases to 300 mg/ml. Interestingly, this maximum density is not correlated to the total mass of the FG-nups, but depends sensitively on the specific combination of essential Nups located in the central plane of the NPC.

3.1 Introduction

Fast and selective transportation of macromolecules between the cytoplasm and the nucleoplasm is essential for the proper functioning of eukaryotic cells. This is accomplished by the nuclear pore complex (NPC) which is embodied in the nuclear envelope membranes and moderates the transport of molecules in a size-selective and directional manner. The NPC is a large molecular assembly with an estimated mass of 44-70 MDa for yeast (Cronshaw et al. 2002, Rout et al. 2000) that provides bidirectional pathways for passive transport of small molecules and facilitated transport of larger proteins (Peters 1984, Feldherr and Akin 1997, Peters 2009, Terry
The active transport mechanism of large macromolecules is driven by soluble nuclear transport factors (NTF) which mostly belong to the Karyopherin family (Kap). During import or export the appropriate NTF binds to cargo with nuclear import or export signals, upon which the NTF-cargo complex is translocated through the NPC.

The NPC has an eight-fold symmetrical structure and is composed of approximately 30 different proteins called nucleoporins (nups) (Yang et al. 1998, Rout et al. 2000). The Nups fall into different subgroups based on their function: transmembrane Nups that attach the NPC to the membrane, structured Nups that form the core scaffold of the NPC and maintain its shape and, finally, FG-nups. The FG-nups comprise about 30% of all Nups and are found to be intrinsically disordered and have many phenylalanine-glycine (FG) repeats in their amino acid sequence (Denning et al. 2003). They line the central channel of the NPC and are anchored to the scaffold through their structural domain. These FG-nups have been shown to be essential for the viability of yeast and presumably all eukaryotes (Strawn et al. 2004). However, how the biophysical properties of the FG-nups determine their function in passive and active transport is subject to intense debate.

Different models have been proposed to explain the role of the FG-nups during nuclear transport. The selective phase model presumes that the weak FG-FG interactions form a homogeneous cross-linked network (a hydrogel) inside the NPC. The Kaps can locally break the cross-links in the network and melt through the gel due to their higher affinity to the FG-repeats (compared to FG-FG affinities); the space between the cross-links serve as a sieve and allows for free diffusion of smaller molecules (Frey et al. 2006). The virtual gate model suggests that the brush-like structure formed by the disordered FG-nups, repels non-specific cargoes, but allows Kap-associated cargoes to overcome this entropic barrier because of the low affinity interactions between the Kaps and the FG-repeats (Rout et al. 2003). The reversible collapse model can be considered as an extension of the virtual gate model in which the active transport is facilitated by a conformational change of the FG-nups due to the presence of the Kaps (Lim et al. 2007). The interaction between Kaps and FG-nups results in a local collapse of the Nups towards their anchor point providing enough space for translocation of the Kap-cargo complex. The reduction of dimensionality model suggests that the wall of the transport channel is covered with an FG-NTF bilayer leaving a 5-10 nm tube for passive diffusion of small molecules at the center of the pore. The active transport is then facilitated by a 2D random walk of the Kaps over the NTF surface (Peters 2005, Peters 2009). Finally, the forest model is based on the Stokes radius and dimension of the individual FG-nup domains. Those FG-nups that have a fully collapsed conformation form ‘shrubs’ near the scaffold and those that consist of an extended domain next to a collapsed domain form ‘trees’, resembling an FG-nup forest landscape. It has been proposed that this configuration forms two distinct transport pathways, one at the center and the other near the scaffold, which are used for active and passive transport, respectively (Yamada et al. 2010). However, no consensus have been reached so far on a prevailing model.

One of the reasons that has hampered the understanding of nuclear transport is the absence of experimental techniques that can probe the structure and dynamics of the disordered proteins inside the transport channel and during transport. This has led to the development of computational approaches to gain insight on the conformation of the FG-nups inside the NPC. Due to the large size of the system, high-resolution (all-atom) molecular dynamics simulations are restricted to study only single FG-nups (Krishnan et al. 2008, Yamada et al. 2010) or a periodic
3.2 Methods

A coarse-grained molecular dynamics model is developed in which each amino acid is represented by one bead. The potential energy of the system is written as

\[ \phi = \phi_b + \phi_{hp} + \phi_{el}, \tag{3.1} \]

where \( \phi_b \) is the potential energy of the bonded interactions, which has three subcomponents for bond stretching \( \phi_{bond} \), bending \( \phi_{bend} \) and torsion \( \phi_{tor} \). An average mass of 120 Da is assigned to each bead and the distance between neighboring beads is fixed at \( b = 0.38 \text{ nm} \) using a stiff harmonic potential \( \phi_{bond} = K_b(r-b)^2 \) with \( K_b = 8038 \text{kJ} \cdot \text{nm}^{-2} \cdot \text{mol}^{-1} \). The bending and torsion potentials for the coarse grained model are extracted from the Ramachandran data of the coiled regions of protein structures as discussed in chapter 2 (Ghavami et al. 2012). It is widely accepted that FG-nups are intrinsically disordered and highly flexible, without evidence of secondary structure formation (Atkinson et al. 2013, Denning et al. 2003, Yamada et al. 2010, Lim, Huang, Koser, Deng, Lau, Schwarz-Herion, Fahrenkrog and Aebi 2006). This indicates that no stable, long-lasting hydrogen bonds are formed. In line with this, hydrogen bonding is not incorporated in our model. The model accounts for hydrophobic and hydrophilic interactions, the effect of the polarity of the solvent, the screening effect of free ions and the electrostatic interactions between charged amino acids (Ghavami et al. 2013). In order to take into account both attractive hydrophobic and repulsive hydrophilic interactions with one function, the following array of short FG-nup segments, end-grafted on a flat surface (Miao and Schulten 2010, Miao and Schulten 2009). On the other hand, several low-resolution approaches have been used to study the transport rate and accumulation of cargo molecules inside the NPC without considering the detailed interactions between the FG-nups (Mincer and Simon 2011, Moussavi-Baygi et al. 2011a). Due to the high level of coarse graining the full geometry of the NPC can be accounted in these approaches, but at the expense of loosing detail at the scale of individual amino acids (Moussavi-Baygi et al. 2011b). Recently, Tagliazucchi et al. (Tagliazucchi et al. 2013) developed a computational model that accounts for the amino acid sequence of the FG-nups by distinguishing 6 different families of amino acids. Then, FG-nups are distributed along the channel tethered to the surface of the NPC-scaffold and minimization of free energy is applied subject to axisymmetric boundary conditions. Using this model they studied the translocation of model particles through the NPC.

The goal of the current work is to probe the full three-dimensional disordered domain of the yeast nuclear pore complex by accounting for all FG-nups, each having a complete 20 amino acid resolution. Our one-bead-per-amino-acid molecular dynamics approach is based on experimentally-obtained hydrophobicity scales and calibrated against experimental Stokes radii of a wide range of FG-nup segments (Yamada et al. 2010). Using the model we will study the effect of electrostatic and hydrophobic interactions on the FG-nup conformation and explore the relation between FG-nup distribution and cell viability, as studied using genetically-modified yeast strains (Strawn et al. 2004).
potential is proposed:

$$\phi_{hp}(r) = \begin{cases} 
\varepsilon_{rep}\left(\frac{r}{\sigma}\right)^8 - \varepsilon_{ij}\left[\left(\frac{r}{\sigma}\right)^6 - \left(1 - \sigma/r\right)^6\right] & r \leq \sigma \\
(\varepsilon_{rep} - \varepsilon_{ij})\left(\frac{r}{\sigma}\right)^2 & \sigma \leq r,
\end{cases} \quad (3.2)$$

where $$\varepsilon_{ij} = \varepsilon_{hp}\sqrt{(\varepsilon_i\varepsilon_j)^{\alpha}}$$ is the strength of the interaction for each pair of amino acids $$(i,j)$$ and $$\sigma = 0.60$$ nm. The hydrophobicity scales obtained from partition energy measurements (Eisenberg 1984, Roseman 1988, Abraham and Leo 1987) are normalized between 0 and 1 and the average value of the experimental data is used to define the relative hydrophobic strength for each amino acid $$\varepsilon_i \in [0, 1]$$ (see Table 3.5 and 3.1 in the Appendix). The $$\varepsilon_{hp}$$ can be interpreted as the absolute hydrophobic strength between the most hydrophobic amino acids, while $$\varepsilon_{rep}$$ defines the intensity of the repulsive hydrophilic interactions. When $$\varepsilon_{ij} > \varepsilon_{rep}$$ the potential is attractive (hydrophobic interactions), when $$\varepsilon_{ij} = \varepsilon_{rep}$$ the potential is neutral, only accounting for excluded volume effects until $$r = \sigma$$ and when $$\varepsilon_{ij} < \varepsilon_{rep}$$ the potential is purely repulsive (hydrophilic interactions). The electrostatic interactions between charged amino acids are described by the modified Coulomb law

$$\phi_{el} = -\frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r(r)r} \exp(-\kappa r). \quad (3.3)$$

Since experiments suggest that the nuclear envelope is not a selective ion barrier (Oberleithner et al. 1993), a Debye screening coefficient similar to that of cytoplasm, $$\kappa = 1.0$$ nm$$^{-1}$$, is chosen (Colwell et al. n.d.). The distance-dependent dielectric constant of the solvent $$\varepsilon_r(r)$$ is included through the sigmoidal function $$\varepsilon_r(r) = \frac{S}{1 + \frac{r^\alpha}{z^\alpha} (\varepsilon_r(z) - 1)^2}$$, where $$S = 80$$, $$z = 0.25$$ nm (Hingerty et al. 1985, Karshikoff 2006).

The exponent $$\alpha$$ and the strength $$\varepsilon_i$$ of the charged amino acids are the free parameters of the model and are calibrated against experimental Stokes radii of FG-nup segments (Yamada et al. 2010). The parametrization is carried out in two steps. In the first step the values of $$\varepsilon_{rep}$$ and $$\varepsilon_{hp}$$ are chosen to be 10.0 and 13.0 kJ.mol$$^{-1}$$ in order to set the minimum energy for the interaction of the most hydrophobic amino acids to -5.2 kJ.mol$$^{-1}$$ as suggested in (Zhang and Kim 2000). The exponent $$\alpha$$ is then chosen such that the model can reproduce the experimental Stokes radius of one of the low-charge FG-nup segments (i.e., Nup42 (AA 1-212))(Yamada et al. 2010)). With $$\alpha = 0.27$$, the $$\varepsilon_i$$ (hydrophilicity) of the charged amino acids (i.e., D, E, K and R) are fine tuned in order to capture the Stokes radius of a highly charged FG-Nup segment (i.e., Nup116s (AA 765-960) (Yamada et al. 2010)).

It should be noted that the hydrophobicity scale for Proline is fine-tuned to reproduce the experimental end-to-end distances of poly-Proline segments reported in (Schuler et al. 2005). The hydrophobicity scale for Glutamine is corrected to reproduce the experimental data regarding the dimension of Polyglutamine with different lengths (Walters and Murphy 2009). Finally, the hydrophobicity index of Glycine is modified to reproduce the experimental gyration radius of polyglycine chains (Ohnishi et al. 2006).
### 3.2. Methods

| Amino acid | A | R | N | D | C | Q | E | G | H | I | L | K | M | F | P | S | T | W | Y | V |
| $\varepsilon_i$ | 0.7 | 0 | 0.33 | 0.0005 | 0.68 | 0.64 | 0.0005 | 0.41 | 0.53 | 0.98 | 1 | 0 | 0.0005 | 0.78 | 1 | 0.65 | 0.45 | 0.51 | 0.96 | 0.82 | 0.94 |
| Charge ($e$) | 0 | 1 | 0 | -1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Histidine is not charged in physiological conditions but upon a small shift in the pH, it will have a net charge. In this study we do not consider Histidine as a charged amino acid because it rarely appears in the sequence of the FG-nups and we do not expect a considerable change in the results by neglecting this effect.

**Table 3.1:** The relative hydrophobic strength $\varepsilon_i$ and charge for each amino acid used in Eq. 1 and 2.

<table>
<thead>
<tr>
<th>FG-nup segment</th>
<th>length</th>
<th>$R_s$ experiment</th>
<th>$R_s$ predict</th>
<th>Errors %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nsp1n lc (AA 1-172)</td>
<td>172</td>
<td>27.1</td>
<td>32.28</td>
<td>16.05</td>
</tr>
<tr>
<td>Nup116m lc (AA 165-715)</td>
<td>551</td>
<td>46.5</td>
<td>45.97</td>
<td>1.15</td>
</tr>
<tr>
<td>Nup100n lc (AA 2-610)</td>
<td>625</td>
<td>48.7</td>
<td>49.43</td>
<td>1.48</td>
</tr>
<tr>
<td>Nup49 lc (AA 1-215)</td>
<td>215</td>
<td>26.9</td>
<td>33.52</td>
<td>19.74</td>
</tr>
<tr>
<td>Nup42 lc (AA 1-212)</td>
<td>212</td>
<td>28.4</td>
<td>29.29</td>
<td>3.04</td>
</tr>
<tr>
<td>Nup57 lc (AA 1-255)</td>
<td>255</td>
<td>31.9</td>
<td>34.66</td>
<td>7.97</td>
</tr>
<tr>
<td>Nup155 lc (AA 1-242)</td>
<td>242</td>
<td>28.2</td>
<td>32.22</td>
<td>12.47</td>
</tr>
<tr>
<td>Nup1 lc lc (AA 798-1078)</td>
<td>279</td>
<td>32.4</td>
<td>36.31</td>
<td>10.77</td>
</tr>
<tr>
<td>Nup159 hc (AA 441-881)</td>
<td>441</td>
<td>55.4</td>
<td>61.13</td>
<td>9.37</td>
</tr>
<tr>
<td>Nup60 hc (AA 389-539)</td>
<td>151</td>
<td>31.3</td>
<td>34.81</td>
<td>10.09</td>
</tr>
<tr>
<td>Nup1m hc (AA 220-797)</td>
<td>578</td>
<td>67.9</td>
<td>72.73</td>
<td>6.64</td>
</tr>
<tr>
<td>Nup2 hc (AA 186-561)</td>
<td>376</td>
<td>59.8</td>
<td>55.84</td>
<td>7.1</td>
</tr>
<tr>
<td>Nsp1m hc (AA 173-603)</td>
<td>431</td>
<td>65.3</td>
<td>64.95</td>
<td>0.54</td>
</tr>
<tr>
<td>Nup155ns (AA 243-433)</td>
<td>191</td>
<td>29.8</td>
<td>36.47</td>
<td>18.29</td>
</tr>
<tr>
<td>Nup100s (AA 611-800)</td>
<td>190</td>
<td>36.6</td>
<td>40.91</td>
<td>10.53</td>
</tr>
<tr>
<td>Nup116s (AA 765-960)</td>
<td>196</td>
<td>39.1</td>
<td>42.28</td>
<td>7.52</td>
</tr>
</tbody>
</table>

* “lc” stands for low charge, “hc” stands for high charge content and “s” refers to the stalk region of the Nup (Yamada et al. 2010).

**Table 3.2:** The predicted $R_s$ values for FG-nup segments compared to the experimental values (Yamada et al. 2010).

Although the Stokes radii of only two FG-nup segments are used to obtain $\alpha$ and $\varepsilon_i$, the Stokes radii of the rest of the FG-nup segments are predicted within 20% error with respect to the experimental values (Table 3.2 of the Appendix). It should be noted that this accuracy cannot be achieved without considering repulsive interactions between hydrophilic amino acids. The performance of the force-field for the collective interaction of FG-nups is verified by simulating an array of Nsp1, end-grafted on a flat surface (see Appendix). The computed brush height of $h = 37 \pm 1$ nm compares well with the experimental value, $h = 34 \pm 4$ (Eisele et al. 2010).
Figure 3.1: (A) Simplified geometry of the core scaffold of the yeast nuclear pore complex reconstructed based on the model of Alber et al. (Alber et al. 2007b, Alber et al. 2007a). The outer radius of the scaffold changes from 30 nm at the center to 33.5 nm at the peripheries. The inner blobs which are decorated in 8 fold rotational symmetry, represent Nup188 of the inner rings. The blobs at the cytoplasmic ring represent Nup82 and Nic96, while the ones at the nuclear ring represent Nic96. (B) The circumferential projection of radial and axial positions of the anchor points of the FG-nups. (C) The distribution of charged amino acids (red sticks) and FG-repeats (green sticks) in the sequence of the FG-regions of the FG-nups as used in the model. All Nups are anchored to the scaffold at their C-terminus.

A simplified geometrical model of the NPC is built based on the geometry of the core scaffold of the yeast NPC and the FG-nups are anchored at the predicted positions inside the pore (Alber et al. 2007a, Alber et al. 2007b), as shown in Fig. 3.1. Table 3.3 lists the FG-nups with the coordinates of the anchor points specified in Table 3.6. The location of the FG-repeats and charged amino acids in the sequence of the FG-nups is depicted in Fig. 3.1B. The scaffold is modeled using hard-sphere beads with a radius of 2.5 nm which are assumed to have no specific interaction with the FG-nups. The initialization procedure and details regarding the
molecular dynamics simulations are given in the Appendix.

### Table 3.3: The stoichiometry data of the FG-nups and their properties.

<table>
<thead>
<tr>
<th>Name</th>
<th>modellled length *</th>
<th># per NPC</th>
<th># of FG-repeats</th>
<th>%charge †</th>
<th>%hydrophobic ‡</th>
<th>( R_g; NPC / R_g; isolated ) §</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nup1</td>
<td>AA 1-520 (823)</td>
<td>32</td>
<td>43</td>
<td>22.42</td>
<td>11.77</td>
<td>6.51</td>
</tr>
<tr>
<td>Nup49</td>
<td>AA 1-472 (472)</td>
<td>16</td>
<td>18</td>
<td>12.71</td>
<td>24.15</td>
<td>5.51</td>
</tr>
<tr>
<td>Nup57</td>
<td>AA 1-541 (541)</td>
<td>16</td>
<td>16</td>
<td>14.60</td>
<td>22.37</td>
<td>3.91</td>
</tr>
<tr>
<td>Nup145</td>
<td>AA 1-866 (1317)</td>
<td>16</td>
<td>15</td>
<td>21.52</td>
<td>26.09</td>
<td>5.97</td>
</tr>
<tr>
<td>Nup116</td>
<td>AA 1-726 (1113)</td>
<td>8</td>
<td>47</td>
<td>3.44</td>
<td>17.36</td>
<td>4.01</td>
</tr>
<tr>
<td>Nup100</td>
<td>AA 1-818 (959)</td>
<td>8</td>
<td>44</td>
<td>8.21</td>
<td>19.00</td>
<td>4.52</td>
</tr>
<tr>
<td>Nup60</td>
<td>AA 1-539 (539)</td>
<td>8</td>
<td>0</td>
<td>24.68</td>
<td>22.63</td>
<td>5.78</td>
</tr>
<tr>
<td>Nup42</td>
<td>AA 1-934 (1076)</td>
<td>8</td>
<td>14</td>
<td>22.06</td>
<td>18.42</td>
<td>5.20</td>
</tr>
<tr>
<td>Nup159</td>
<td>390-1460 (1460)</td>
<td>8</td>
<td>25</td>
<td>6.51</td>
<td>18.60</td>
<td>4.41</td>
</tr>
</tbody>
</table>

* The anchor point of Nup53 and Nup59 are located at the outer face of the core scaffold (Alber et al. 2007b) and were not included in the simulations. Also, Nup2 was not included since there was no data regarding the location of its anchor point and it was categorized as one of the non-essential FG-nups (Strawn et al. 2004). Our predictions for the Stokes radius of the full Nup49 (\( AA 1-472, R_s; exp = 41 \ \text{Å}, R_s; sim = 44.9 \ \text{Å} \)) and Nup57 (\( AA 1-541, R_s; exp = 47 \ \text{Å}, R_s; exp = 51.6 \ \text{Å} \)) show good correspondence to the experimental values (Yamada et al. 2010). Therefore, we included the full FG-nup sequence of these Nups. The structural domain of Nup42 is small compared to the FG domain, therefore, we included the full sequence of Nup42 (Denning et al. 2003). Nup60 is also predicted to be fully unstructured, therefore the full sequence is included (Denning and Rexach 2007). The C-terminal domain of Nup1 (Sistla et al. 2007), Nup116 (Ho et al. 2000), Nup1 (Ballot et al. 2001), Nup100 (Devos et al. 2006) and N-terminal domain of Nup159 (Devos et al. 2006) are not included in the simulations. The FG domain of Nup145 is shorter than 896 residues, but since it has low amount of secondary structure up to \( AA850 \), that part is also included (Devos et al. 2006, Devos et al. 2004).

† The charge content is obtained based on the number of D, E, K and R amino acids in the sequence of the FG-nups.

‡ The hydrophobicity is obtained based on the number of the V, Y, W, F, M, L and I amino acids in the sequence of the FG-nups.

§ The \( R_g; NPC \) is obtained by calculating the gyration radius of all the copies of the same FG-nup inside the NPC (this study). The \( R_g; isolated \) is obtained from the single FG-nup simulations of the same Nup (this study).

### 3.3 Results

#### 3.3.1 Distribution of the FG-nups inside the NPC

First, the distribution of the FG-nups inside the NPC is investigated. The three-dimensional density distribution of the FG-nups is obtained by calculating the average number density of the amino acids over the simulation time. An iso-surface plot corresponding to \( 140 \ \text{mg/ml} = 0.7 \ \text{nm}^{-3} \) number density is shown in Fig. 3.2A. The results show a low density region up to \( r \sim 5 \ \text{nm} \) from the central axis of the NPC surrounded by a high density region.

The distribution is further analyzed by studying the distribution of the charged residues and FG-repeats inside the NPC. The charged amino acids (i.e. D, E, K and R, cf. Tab. 3.1) are accumulated near the scaffold in between the inner ring blobs (Nup188 (Alber et al. 2007b)) (see Fig. 3.2B). The FG-repeats are distributed in the form of a doughnut between the central low-density region and the highly-charged layer near the scaffold (see Fig. 3.2C). Since the distribution of the FG-nups is relatively homogeneous in the circumferential direction, it can be averaged to generate a two-dimensional (2D), \( r-z \) density map (see Fig. 3.3A). It shows that the FG-doughnut is located slightly above the central plane of the NPC (\( z \sim 4.5 \ \text{nm} \)) towards the cytoplasmic side.

Furthermore, the 2D density maps of different residues can be averaged in the vertical direction to obtain the radial density profiles. The radial density distribution of the FG-repeats and negatively and positively charged amino acids are plotted in Fig. 3.4. The FG-repeat distribution shows a peak value at \( r = 13 \ \text{nm} \) from the central axis of the NPC while the peak values for the negative and positive amino acids are at \( r = 20 \ \text{nm} \) and \( r = 21 \ \text{nm} \), respectively. More-
over, the charge distribution shows that the central region of the pore has a net positive charge which is consistent with the simulation results by Tagliazucchi et al. (Tagliazucchi et al. 2013). The conformation of the FG-nups inside the pore is found to be notably different from that of the isolated FG-nups: the ratio of the $R_g$ values for the FG-nups inside the NPC to the gyration radius of the isolated FG-nups ranges from 3.9 to 6.5 (see Table 3.3).

Figure 3.2: The 3-dimensional density distribution of different amino acids inside the NPC obtained from wildtype-1 simulation. (A) The distribution of all amino acids. The iso-surface plot corresponds to a mass density of 140 mg/ml. (B) The density distribution of charged amino acids, (C) The density distribution of FG-repeats inside the transport channel of the NPC. The iso-surface plot corresponds to an average distance of 2.7 nm between the FG-repeats.

A second simulation (wildtype-2) with the same FG-nups but different starting configuration and initial velocity distribution is performed to study the sensitivity to the initial conditions (the method to generate starting configurations is discussed in the Appendix). The 2D density distribution (Fig. 3.3B) shows the same characteristics as that of the first simulation (wildtype-1, Fig. 3.3A). In addition, the localization of the individual FG-nups is similar in both simulations which confirms that the obtained results are not biased by the simulation setup (see Figs. 3.9 and 3.10). Furthermore, there are uncertainties regarding the exact position of the anchor points of the FG-nups according to the experimental data (Alber et al. 2007a). To investigate the sensitivity to the anchor point locations, we performed a new simulation (wildtype-3) in which the anchor points are displaced in a random direction. The coordinates of the anchor points for wildtype-3 (see Table 3.3) are displaced by an average distance of $2.1 \pm 0.9$ nm relative to wildtype-1, which corresponds to the experimental uncertainties reported in (Alber et al. 2007a). The obtained distribution has the same pattern as that of the first and second simulations as shown in Fig. 3.8. We conclude that small deviations in the anchor point locations of the FG-nups do not affect the overall distribution of the FG-nups.

In order to study the localization of the individual FG-nups inside the pore, the 2D ($r$-$z$) density distribution is calculated for all FG-nups that are anchored at the same ($r$, $z$) coordinates.
3.3. Results

The results reveal that some of the FG-nups (i.e., Nup49, Nup57, Nup116, Nup100 and Nup42) are confined inside the pore and contribute in forming the central high density region. The rest of the FG-nups (i.e., Nsp1, Nup159, Nup145, Nup60 and Nup1) have more conformational freedom and spread out over a larger volume at both sides of the NPC.

Figure 3.3: 2D density plots of the FG-nups in the simulated NPCs. (A) wildtype-1 NPC (B) wildtype-2 NPC simulated with a different starting configuration and initial velocity distribution compared to wildtype-1. (C) no charge NPC, in which all charged residues are replaced by neutral beads in the sequence of the FG-nups. (D) “Denatured” NPC where all residues are replaced by neutral beads. (E) “Reversed” NPC where the FG-nups are anchored from their N-terminus. (F) “Uniform” NPC in which the sequence of the FG-nups is modified such that they have a uniform distribution of charged and hydrophobic amino acids along their length.

3.3.2 Effect of hydrophobic and electrostatic interactions on the distribution of the FG-nups

The contribution of the hydrophobic and electrostatic interactions in shaping the distribution of the FG-nups inside the wildtype NPC is systematically studied. First, the wildtype NPC is modified by removing the charged residues, D, E, K and R, and replacing them with neutral beads (i.e. residues of radius 0.6 nm with no specific attraction, repulsion or charge interaction). The resulting density plot in this “no charge” NPC shows that by removing charged amino
acids, all FG-nups aggregate in a wide, high density region, leaving an empty region at the
center of the NPC (Fig. 3.3C). In a next step, the rest of the amino acids are also replaced with
neutral beads resulting in a “denatured” NPC. The obtained distribution for the denatured
NPC shows a uniform distribution of low density throughout the pore (Fig. 3.3D). The results
indicate that repulsive interactions between the charged amino acids in the wildtype NPC serve
as a bumper to push the dense FG-nup clusters, formed through hydrophobic interactions,
towards the center of the pore. Therefore, the “doughnut” structure in Fig. 3.2 is a direct result
of the balance between the electrostatic and hydrophobic interactions. This balance is strongly
related to the amino acid sequence of the FG-nups, showing a high number of charged residues
near the C-terminal domain (Fig. 3.1C).

![Figure 3.4: The radial density distribution of the positively and the negatively charged amino acids and FG-repeats. The radial distribution is obtained by averaging the 2D density plot for each residue in the vertical direction, \( -70 < z < 70 \) nm.](image)

3.3.3 Effect of amino acid sequence

Next, we explore the effect of the FG-nup amino acid sequence on the density distribution in the
wildtype NPC. To do so, we anchored the FG-nups from their N-terminus to the same anchor-
ing points. The distribution for this “reversed” NPC is shown in Fig. 3.3E. Finally, a “uniform”
NPC is simulated where the residues in the sequence of the FG-nups are reshuffled such that
the charged and hydrophobic amino acids become uniformly distributed along their length
(Fig. 3.3F). The density distributions for both the “reversed” and “uniform” NPC are different
from the wildtype NPC. In the reversed case, the high density region has shifted towards the
scaffold on the cytoplasmic side. For the “uniform” NPC, a wide, low density region is de-
tected at the center of the pore. A density concentration can be observed near the cytoplasmic
entrance of the pore but with a lower peak density compared to the wildtype NPC. In these
two cases, the amino acid composition of the FG-nups (percentage of charged and hydropho-
bic amino acids) has not changed compared to the wildtype NPC, suggesting that the amino
acid sequence of the FG-nups plays a key role in the density distribution of the FG-nups in the
wildtype NPC.
3.3. Results

3.3.4 Viable versus inviable NPCs

Experiments have shown that cells can survive despite deletion of more than half of the mass of the FG-repeats of the FG-nups (Strawn et al. 2004). More specifically, Wente and co-workers (Strawn et al. 2004) showed that the presence of Nup100 or Nup116 is necessary but not sufficient. They proposed that if Nup116 is present and Nup100 is not, any combination of two of the four nups, Nsp1, Nup49, Nup57 or Nup145, are required, and conversely, if Nup100 is present and Nup116 is not, three of the mentioned four FG-nups must be present. We now explore a possible relation between cell viability and FG-nup density distribution by studying two extremes: a viable NPC with the least amount of essential FG-nups (minimal mass) and inviable NPCs with maximal mass of the FG-nups (see Table 3.4). The “viable” NPC includes all the FG-nups anchored near the central plane of the NPC (see also Fig. 3.1B). The importance of Nup116 and Nup100 is investigated by removing only those two Nups (inviable-2). For the other four simulated inviable cases, Nup116 and two of the four Nups, Nup49, Nup57, Nup145 or Nsp1 are removed from the pore.

![Figure 3.5: The 2D density plots for viable (A) and inviable NPCs (B-E) corresponding to Table 3.4.](image)

The 2D density distributions for all these cases are shown in Fig. 3.5. Although the viable NPC contains less mass than the inviable cases, it has formed a high density ring region, similar to what was observed before in the wildtype NPCs (Figs. 3.3A and B), whereas the inviable cases do not exhibit this. This difference is summarized in Fig. 3.6 in terms of the radial density.
distribution at the \( z \)-location of the maximum density (excluding the anchor points) for each simulated NPC. The maximum density for the inviable NPCs does not exceed 185 mg/ml while for the viable and wildtype NPCs a maximum density of 300 mg/ml is reached.

<table>
<thead>
<tr>
<th>Simulation name</th>
<th>Nups removed (Strawn et al. 2004)</th>
</tr>
</thead>
<tbody>
<tr>
<td>viable</td>
<td>nup42 nup159 nup1</td>
</tr>
<tr>
<td></td>
<td>nup60 nup100 nsp1 nup145 *</td>
</tr>
<tr>
<td>inviable-1</td>
<td>nup116 nsp1 nup49</td>
</tr>
<tr>
<td>inviable-2</td>
<td>nup116 nup100</td>
</tr>
<tr>
<td>inviable-3</td>
<td>nup116 nup145 nup49</td>
</tr>
<tr>
<td>inviable-4</td>
<td>nup116 nup57 nup49</td>
</tr>
<tr>
<td>inviable-5</td>
<td>nup116 nup145 nsp1</td>
</tr>
</tbody>
</table>

* For the viable case, only the peripheral copies of nsp1 (nsp1-1 and nsp1-2) and nup145 (nup145-1) are removed.

**Table 3.4:** The list of simulated viable and inviable NPCs and their composition

![Figure 3.6: Comparison of the radial density distribution at the \( z \)-location of maximal density in the 2D density plots for the simulated NPCs. The mass density is calculated using an average mass of 120 Da per residue.](image)

### 3.4 Discussion and conclusions

We have investigated the density distribution of the FG-nups inside the transfer conduit of the NPC. For this purpose a one-bead-per-amino-acid coarse-grained model is developed that accounts for hydrophobic/hydrophilic and Coulombic interactions. We have followed a bottom-up approach in which the force field is parametrized against experimental data of 16 isolated FG-nup segments (Yamada et al. 2010). Subsequently, the model is validated against FG-nup brushes and finally, the model is used to study the collective behavior of the FG-nups inside the three-dimensional geometry of the NPC.

We show that the collective interaction of the FG-nups results in the formation of different
zones inside the pore. A relatively low density region is observed along the central axis of the NPC which is surrounded by a coherent hydrophobic “doughnut”-like structure, rich in FG-repeats. The accumulation of charged residues is higher near the scaffold where the FG-nups are anchored. We show that the disruption of the FG-nup amino acid sequence considerably changes the density distribution of the FG-nups (see Fig. 3.3E and 3.3F). This is in contrast to the results of Tagliazucchi et al. (Tagliazucchi et al. 2013), and can be traced back to the different interaction energies used between hydrophilic amino acids. In (Tagliazucchi et al. 2013) an attractive interaction energy between hydrophilic amino acids on the order of $1.0 k_B T$ is used, while in our approach a repulsive interaction is defined between hydrophilic residues. The rationalization for the latter is that the favorable interaction between hydrophilic amino acids (including charged amino acids) and the solvent gives rise to a repulsive interaction between these amino acids. This is in accordance with the experimental observations of Yamada et al. (Yamada et al. 2010), which show that the Stokes radius of short disordered segments of FG-nups is correlated with the total number of charged amino acids and not with the net charge of these segments. As a result, it can be postulated that repulsive hydrophilic interactions are more important than Coulombic interactions in determining the dimensions of these proteins.

Comparison of the FG-nup distribution of viable and inviable NPCs reveals that the formation of a high density hydrophobic doughnut-like structure, rich in FG-repeats, is a characteristic footprint of viable NPCs. The contribution of the individual FG-nups to the total distribution is investigated through localization maps of all FG-nups inside the pore (see Figs. 3.9 and 3.10). The localization data suggest that FG-nups fall into two different categories; roughly half of the FG-nups contribute to the formation of the FG-repeat doughnut at the center of the NPC (i.e., Nup57, Nup49, Nup42, Nup116, Nup100 and Nup145-2), while the rest have more conformational freedom at the entrance and exit of the NPC (i.e., Nup60, Nup159, Nsp1, Nup1 and Nup145-1). The FG-repeats of the FG-nups in the latter category can facilitate the entrance of Kaps/cargo complexes to the pore, while most of the FG-nups in the former category are marked as essential FG-nups, necessary for the viability of the cell (Strawn et al. 2004). Since it has been shown that the deletion of the FG domains of the FG-nups increases the permeability of the NPC (Patel et al. 2007), a possible reason for the inviability of the cells in the experiments of Wente and co-workers (Strawn et al. 2004) could be the disruption of the permeability barrier of the NPC. Our results show that even though the total mass of the FG-nups in the inviable mutant NPCs is larger than the viable mutant NPC, the density of the FG-nups does not exceed a certain value (i.e., 185 mg/ml). This indicates that the mutual interaction of an essential set of FG-nups is key in forming a doughnut-like hydrophobic cluster, rich in FG-repeats, which might control the NPC’s permeability barrier.

It has been shown that Kaps simultaneously interact with up to four FG-repeats during transport (Naim et al. 2009). This suggests a correlation between the average distance of the FG-repeats inside the NPC and the distance between the binding sites on the surface of the Kaps. Experiments have revealed that the binding sites of NTF2 are separated by 3.5 nm (Bayliss et al. 2000) and simulation studies have shown that the distance between the binding spots of Kapβ-1 ranges from 2.5 to 4.0 nm (Isgro and Schulten 2005). We have used the density distribution of the FG-nups to estimate the average distance between the FG-repeats inside the pore (Fig. 3.11). The results show that inside the FG-doughnut the distance between the FG-repeats is lowest, reaching values as low as 2.7 nm. This suggests that throughout the FG-
doughnut, there is more than one FG-repeat available for each binding site so that the Kaps can easily find a nearby FG-repeat and translocate from one FG-repeat to the other (assuming that the presence of the Kaps does not considerably change the local conformation of the FG-nups). This is in accordance with the model of Bednenko et al. (Bednenko et al. 2003) which proposes that transport is accomplished through a series of binding and unbinding events between the FG-repeats and Kaps.

Single molecule tracking of small molecules, Imp-β1 and import complexes has revealed two different but not completely separated pathways for active and passive transport (Ma et al. 2012). The spatial distribution of small molecules along the radius of the pore shows a peak value at the center of the NPC, while the distribution of Imp-β1 exhibits a peak closer to the scaffold. This is in agreement with the density distributions obtained from our simulations which show a low density region of FG-nups at the center (which might allow passive diffusion of small molecules) and a peak density of FG-repeats at 13 nm along the radius of the NPC (which might mediate active transport).

Our model does not allow discriminating between the different transport models proposed in the literature. However, we do note that the shrubs and trees as described by the forest model (Yamada et al. 2010) were not reflected in the obtained density distributions. A possible reason is that the forest model is based on the conformation of isolated FG-nups which apparently is different from the conformation of the FG-nups inside the core of the NPC (see Table 3.3, last column). The density of the FG-nups inside the FG-doughnut is in correspondence to the densities (i.e., ≥ 200 mg/ml) at which a “saturated hydrogel” with selective barrier properties of the NPC can be formed (Frey and Görlich 2007, Hülsmann et al. 2012), while in the central region of the pore the density of the FG-nups would be too low (i.e., ≤ 150 mg/ml) (Frey and Görlich 2007).

To conclude, we have studied the collective behavior of FG-nups inside the transport conduit of the nuclear pore complex through coarse-grained molecular dynamics simulations. The obtained density distribution of the FG-nups reveals the presence of a low density region of diameter ≥ 10 nm at the center of the NPC, surrounded by a high density, FG-repeat rich region. We show that this unique doughnut-like distribution is encoded in the amino acid sequence of the FG-nups and is driven by hydrophobic, hydrophilic and electrostatic interactions. Our results indicate that the maximum FG-nup density correlates with the viability of the cells and is independent of the total mass of the FG-nups. Comparison of experimental observations with our results suggests that the low density region at the center of the NPC provides a permeable medium through which ions and small proteins can freely diffuse, while the doughnut-like FG-repeat-rich region provides the required interaction sites for facilitated transport of Kap-cargo complexes. Further simulation studies are needed to confirm this.
3.A Appendix: Simulation details

Langevin dynamics simulations are carried out at 300 K using Gromacs molecular dynamics simulation software (Hess et al. 2008). The cut-off distance for Van der Waals and Coulombic interactions are set to 2.5 nm and 5.0 nm, respectively. A time-step size of 0.02 ps is chosen and the Langevin friction coefficient is set to 50 ps\(^{-1}\) which is similar to the collision frequency of the water molecules (Izaguirre et al. 2001). The Stokes radii are calculated according to the method explained in (Ghavami et al. 2012), in which each FG-nup is simulated for \(10^7\) steps and the average Stokes radius for the generated conformations (excluding the first \(10^6\) steps) is calculated using the HYDRO computer program (Carrasco and Garcia de la Torre 1999, Garcia de la Torre et al. 1994).

3.B Appendix: Brush simulation

The brush simulation is carried out by making a \(10 \times 10\) array of Nup62 (AA 1-240) nups, grafted from the C-terminus to a flat surface. The simulation box is periodic in the two lateral directions but not in the direction along the height of the brush. The distance between the anchor points is set to 2.4 nm corresponding to the experimental set up of Schoch et al. (Schoch et al. 2012). In this experiment nonspecific proteins (i.e., BSA) are used as molecular probes to measure the height of the FG-nup brush by surface plasmon resonance (SPR) measurement. In order to simulate the experiments of Schoch et al. non-interacting BSA particles are included in the simulation box. BSA particles are modeled as spheres with a diameter of 7.4 nm which have no favorable interaction with the FG-nups and are free to move in the simulation box.

![Figure 3.7](image)

**Figure 3.7:** The density distribution of model BSA particles along the height of Nup62 brush. The inset shows the maximum height of the Nup62 brush as a function of simulation steps.

The simulation is carried out in two steps. In the first step, the end-grafted Nup62 brush in the absence of BSA particles are simulated for \(2.5 \times 10^6\) steps. The maximum height of the FG-Nups as a function of simulation steps is plotted in the Fig. 3.7 inset. The final structure...
of the brush is then used for brush simulation in the presence of BSA particles. The model BSA particles are placed in the simulation box and the system has been simulated for another $2.5 \times 10^6$ steps. The density profile of the model BSA particles in the direction along the height of the brush for the last $1.25 \times 10^6$ steps is calculated and plotted in Fig. 3.7. Since the presence of BSA particles will change the refractive index of the solution, the inflation point of the BSA density distribution is used as an indication of the brush height. Our results suggest a brush height of 13 nm for Nup62 brush while SPR measurements suggest a height of 14.1 nm for Nup62 brush.

### 3.C Appendix: NPC simulation

In order to perform the simulations on the NPC, single FG-nups are placed inside the transfer channel with an initial conformation taken from single FG-nup simulations. For each simulation, the system is first energy minimized in order to remove the possible clashes between the FG-nups. The starting configuration for the wildtype-1 NPC is obtained by replacing all amino acids of the system with neutral beads and running the system for $5 \times 10^5$ steps. The starting configuration for NPC-wildtype-2 is obtained by replacing all of the charged residues with neutral beads and running the system for $5 \times 10^5$ steps. Each subsequent NPC-simulation consists of at least $6 \times 10^7$ steps, and the first $5 \times 10^6$ steps are neglected in generating the results.

In order to obtain 3D number densities, the NPC is centered in a box of 100 nm by 100 nm along the sides and 140 nm along the axial axis of the NPC which then is discretized using $(0.5 \text{ nm})^3$ unit cells. The number of the desired beads (e.g., all residues, only FG-repeats or only charged residues) in each cell is counted over the total simulation time and a 3D number density profile $n(r, \theta, z)$ is generated based on the collected data (Beckstein and Sansom 2003). 2D density maps $n(r, z)$ are then calculated from

$$n(r, z) = \frac{1}{2\pi} \int_0^{2\pi} n(r, \theta, z) d\theta.$$   \hspace{1cm} (3.4)

Finally, the radial density distribution is obtained by averaging the obtained 2D density maps in the $z$ direction

$$n(r) = \frac{1}{L_z} \int_{\theta, z} n(r, z) dz,$$   \hspace{1cm} (3.5)

where $L_z = 140 \text{ nm}$ is the height of the box.

The number density of the FG-repeats is also used to estimate the average distance $d(r, z)$ between the FG-repeats inside the transfer channel through

$$d(r, z) \approx \left( \frac{1}{n(r, z)} \right)^{1/3}.$$   \hspace{1cm} (3.6)
Table 3.5: The hydrophobicity scales used in the model.

3.D Appendix: Figures and Tables

Table 3.6: The coordinates of the anchor points of the FG-nups extracted from the localization data of FG-nups (Alber et al. 2007b) used in the wildtype-1 & 2 simulations. In wildtype-3 the coordinates of the anchor points are displaced in a random direction. The coordinates are for one spoke of the NPC and are repeated periodically in eight-fold symmetry.

Figure 3.8: The 2D density plot of the NPC in which the anchor points of the FG-nups are randomly displaced (wildtype-3).
Figure 3.9: Localization of the individual FG-nups for wildtype-1 NPC.
Figure 3.10: Localization of the individual FG-nups for wildtype-2 NPC.
3. Probing the disordered domain of the nuclear pore complex

Figure 3.11: 2D contour plots showing the distance between the FG-repeats inside the NPC for wildtype-1 and wildtype-2.