Title

Computational Biophysical Model of the Nuclear Pore Complex: Insights from Statistical and Big-Data Analysis

Author

Moussavi-Baygi, Ruhollah

Publication Date

2017

Peer reviewed|Thesis/dissertation

Computational Biophysical Model of the Nuclear Pore Complex: Insights from Statistical and Big-Data Analysis

By
Ruhollah Moussavi-Baygi

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy
in

Applied Science and Technology in the

Graduate Division of the

University of California, Berkeley

with Designated Emphasis on

Computational Science and Engineering

Committee in charge:
Professor Mohammad Mofrad, Chair
Professor David Attwood
Professor Ahmet Yildiz

Fall 2017

Abstract

Computational Biophysical Model of the Nuclear Pore Complex: Insights from Statistical and Big-Data Analysis

by
Ruhollah Mousavi-Baygi

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in

Applied Science and Technology
with Designated Emphasis on
Computational Science and Engineering

University of California, Berkeley

Professor Mohammad R.K. Mofrad, Chair

The nuclear pore complexes (NPCs) are the sole channels known on the nuclear envelope in eukaryotes through which nucleocytoplasmic traffic is selectively and efficiently conducted. While the NPC has been the subject of extensive research for the past six decades and many of its structural, biochemical, and biophysical details are revealed, mechanistics of selective transport yet to be known. The main game-player in conducting the selective transport is a class of the NPC proteins rich in Phe-Gly (FG)-repeat domains that are localized mainly to the channel interior. FG-repeat domains are natively unfolded and thus belong to the family of intrinsically disordered proteins (IDPs). The conformational behavior of FG-repeats remains unknown because of lack of detailed information about their microdynamics, leaving plenty of room for speculation on how the selectivity barrier forms inside the NPC.

To tackle the microdynamics of FG-repeats, here I used polymer physics' principles to develop a coarse-grained computational biophysical model, incorporating the full sequence of all amino-acids. The simulations were run under Brownian dynamics approach to generate long time-evolution of FG-repeat domains.

Under known physiological conditions and geometrical constraints, FG-repeats form a spatially nonuniform cohesive meshwork, percolating in radial and axial directions, with a dense hydrophobic zone in the middle and a low-density zone near the wall. The FG-meshwork is extremely dynamic, resembling a jerking plug with a fluctuating concentration in radial direction.

Being porous with the dominant pore sizes of 4 and 6 *nm*, this dynamic meshwork is permeable to the active cargos in a hydrophobic, and to a lesser extent, charge, stimuli-responsive manner, but strongly impermeable to inert cargos having the same size. An active cargo creates a big deformation inside the FG-meshwork, but because of rapid Brownian motions of the FG-repeats, it reconstructs itself in $<10\mu s$ in a cargo size and shape dependent manner, suggesting the individual FG-repeats undergo reversible collapse. Significantly, the reconstruction process follows a saturating exponential pattern with rapid and slow phases. The characteristic time of reconstruction is a function of cargo size and shape, and is generally smaller for the elongated cargos compared to globular cargos having the same surface chemistry.

Next, I used computational microrheology via many-particle tracking without external probe to investigate the full mechanical spectrum of FG-repeats under different physical and geometrical conditions, including FG-repeats' composition, FG-repeats' length, geometrical confinement, shuttling cargo, and end-tethering. The results reveal that FG-repeats show a non-Newtonian behavior as manifested in their shear-thinning viscosity. The viscoelastic response of FG-repeats is strongly frequency-dependent, and is consistent with the function of the permeability barrier at different frequencies, or equivalently, at different timescales. At low frequencies, equivalent to timescale of nucleocytoplasmic transport, FG-repeats form a pseudo solid-like meshwork. At high frequencies, equivalent to the timescale of thermal diffusion of small molecules, FG-repeats behave like a predominantly viscous liquid.

The end-tethering is determined to be the most influential factor in shaping the mechanical spectrum of the FG-repeats. When the end-tethering is lifted and FG-repeats get free in the space, they invariably behave as a non-Newtonian viscous liquid over all frequencies. The other factors investigated are geometrical confinement, FG-repeats composition, i.e. hydrophobicity and charge content, length of FG-repeats, and shuttling cargo. Although these factors might shift the value of viscoelastic response in the frequency domain, they do not change the physics of the response, as end-tethering does. Comparing the viscoelastic response of FG-repeats with a general polymer melt show a remarkable consistency between FG-repeats and polymer melt over the range of frequencies accessible to the current model. Ultimately, the answer to the question of "do FG-repeats form gel?" is discussed in detail.

Dedication

To:
The first pillar of my life, my father;
The first love-feeder of my life, my mom;
The first teacher of my life, my brother, Seyyed Mohammad;
and,
My forever-loved wife, without help of whom it was not possible to reach here.

Table of Contents

Acknowledgement

This work is the outcome of my exploratory journey at UC Berkeley, during which I have enjoyed fruitful discussions, comments, advice, and help from colleagues, friends, and professors, some of whom I have never met in person, but just communicated with over emails. I feel beholden to all of them, for which I want to express my sincere appreciation here.

In particular, I am so indebted to my research adviser, *Professor Mohammad Mofrad*, who has always been my go-to support for literally everything. I also thank my dear friends *Yousef Jamali* for all his great technical help and guidance, and *Reza Karimi* for his deep views and super-useful discussions. Fruitful discussions with all of the past and the current members of the Molecular Cell Biomechanics lab is highly appreciated. Particularly, I am so grateful to the following lab members:

Zeinab Jahed,
Ehsan Asgari,
Mohammad Azimi,
Hanif Mahboobi,
Hengameh Shams,
Mohammad Soheilypour,
Mohaddeseh Peyro,
Pritti Chandran,
Alex Javanpour,
Vimalier Reyes-Ortiz,
Javad Golji,
Amirhossein Arzani,
Amir Shamloo,
Allison Vuong,
Kiavash Garakani,

and, to professors who generously spent their times on answering my questions:
Faith A. Morrison,
David Wemmer,
John F Marko,
Ahmet Yildiz,
Karsten Weis,
Michael Rexach,
Ulrich Kubitscheck,
Thomas A. Waigh,
Ronald Larson,
Madeleine Djabourov,
David Steigmann,
Thomas Mason,
John Kuriyan,
Jan Ellenberg,
Anna Szymborska,
Matthew Tirrell,
Iwao Teraoka,
Michael P. Rout,
and *Ting Xu.*

Chapter 1. The nuclear pore complex: an overview of six decades

1.1 Introduction

In eukaryotic cells, the nucleus regulates gene expression and protects genetic material via the nuclear envelope (NE). Regulated exchange of information between the nucleus and the cytoplasm is therefore essential for eukaryotic cells, and transport of macromolecules across the NE is of crucial importance to cell function and viability. Nucleocytoplasmic transport (NCT) is effectively conducted through nuclear pore complexes (NPCs). The NPC is a large proteinaceous channel that spans and perforates the NE and regulates molecular traffic between the nucleus and cytoplasm (Fig. 1.1). As the unique gateway to the nucleus, NPC selectively facilitates the transport of large cargo while offering a relatively unobstructed pathway for small molecules and ions.

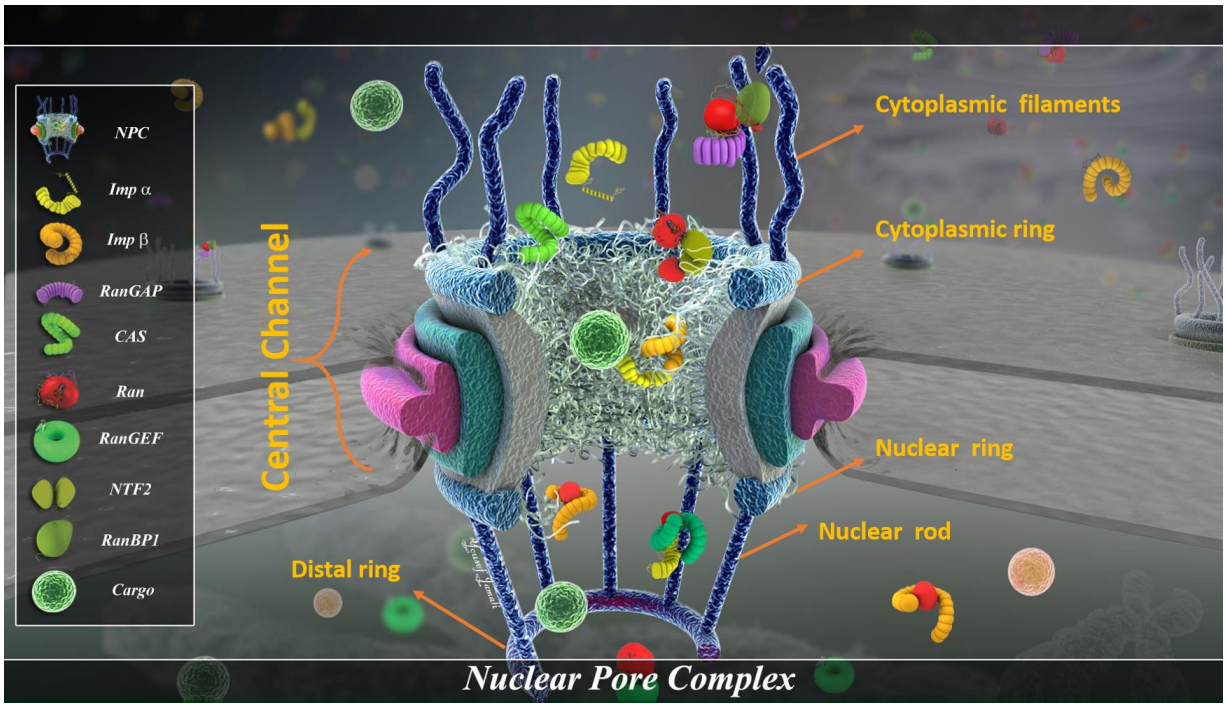


Figure 1.1 The Nuclear pore complex along with different transport-related factors (see the text) Adopted from (Moussavi-Baygi2011a).

Since its discovery in fifties (Callan1950), the NPC has been the subject of numerous studies from different viewpoints. In old days of the discovery of the NPC via electron microscopy, it was imagined only as a passive circular pore (Callan1950), and termed the 'nuclear pore' (Watson1955). It took a decade to find out the octagonality of the pores—then called 'pore complex' (Gall1967), augmented by the following studies where the term 'nuclear pore complex'

appeared (Maul1971). A decent chronological pr ogress in understanding the NPC is depicted in the Fig.1.2.

The NPC shows a plethora of intriguing phenomena from different viewpoints, including structural biology, biochemistry, biophysics, bioinformatics, disease-related, gateway of the gene therapy, and transfection. Recent studies progressively prove that the NPC is not exclusively a channel for the nucleocytoplasmic transport. Instead, it is involved in a wide variety of cellular processes from gene activation and cell cycle regulation to transcriptional regulation, and nuclear envelope integrity (Martínez1999, Sexton2007, Kotwaliwale2009, Capelson2010, Strambio-De-Castillia2010a, Strambio-De-Castillia2010b, Wente2010, Jamali2011, Hurt2015, Kabachinski2015, Jahed2016). Nevertheless, conducting nucleocytoplasmic transport is considered the foremost function of the NPC.

Here I would give a brief review on the NPC mainly from a mechanical perspective to pave the path for the next chapters in dissecting the NPC via computational modeling.

1.2 The structure of the NPC and nucleocytoplasmic transport

The building blocks of the NPC are about 30 distinct proteins collectively called nucleoporins (Nups) (Jamali2011). Each Nups has its own unique name, which is usually the three letters of 'Nup' followed by the molecular weight of the Nup, e.g., Nup62, Nup315, etc. There are indeed exceptions to this rule.

From a structural point of view, Nups can be classified into three groups (Fig 1.3): i) Poms proteins, which anchor the NPC to the NE, ii) Structural Nups, which form and shape the NPC as a whole, and iii) Nups with intrinsically disordered domains rich in Phe-Gly (FG)-repeats, collectively called FG-Nups.

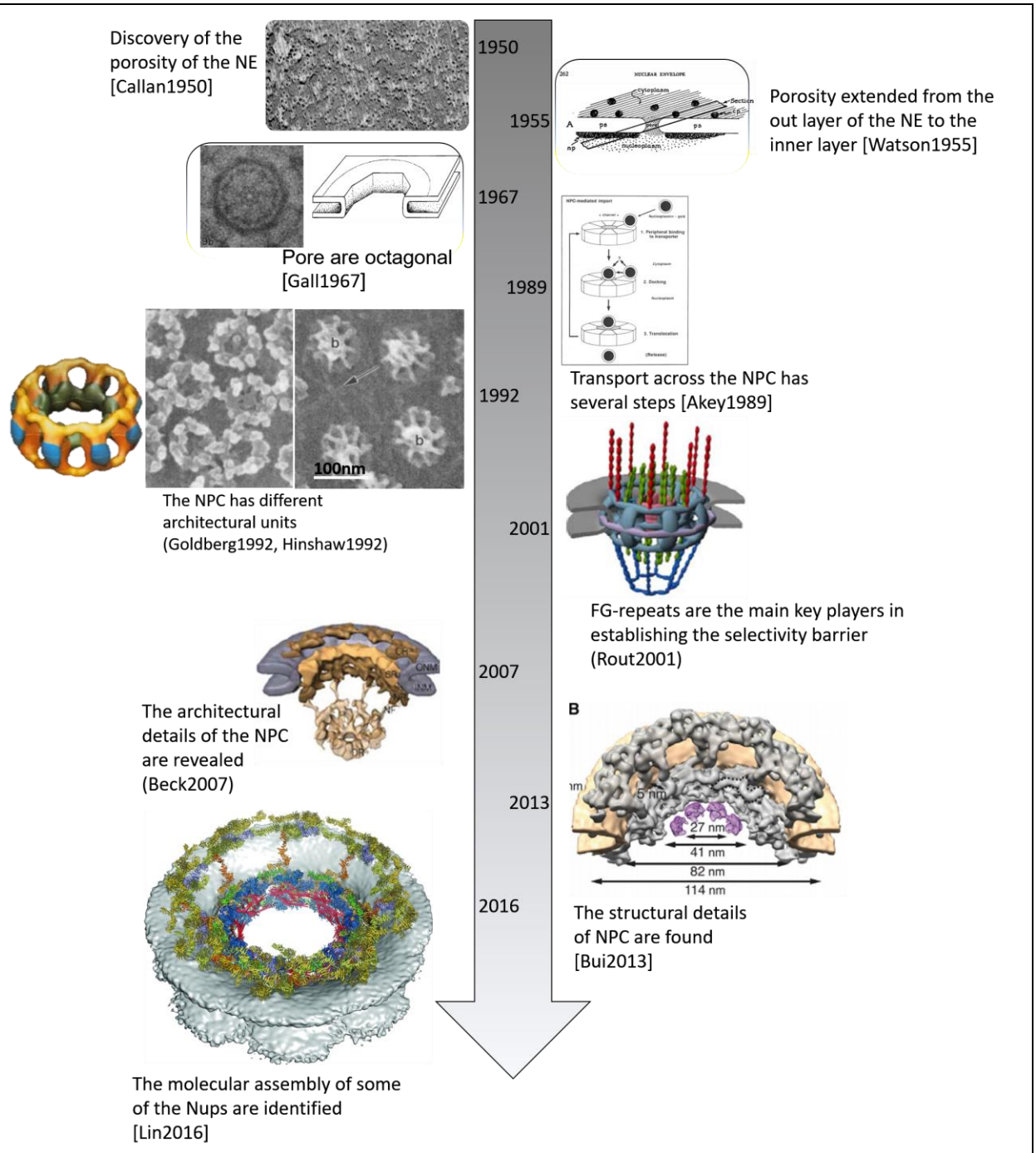


Figure 1.2 A chronological progress in our understanding of the NPC and its structural components since its discovery.

There are 11 different FG-Nups per NPC, constituting about 30% of the total mass of the NPC (Denning2003b, Stewart2007, D'Angelo2008). The total mass of the NPC is species-dependent and is about 44 MDa and 120 MDa in the yeast and vertebrates, respectively (Cronshaw2002, Fahrenkrog2002). The dimensions of the NPC also depend on the species. It is known that the dimensions of the yeast NPC, for example, are smaller than vertebrate NPCs (Fahrenkrog1998).

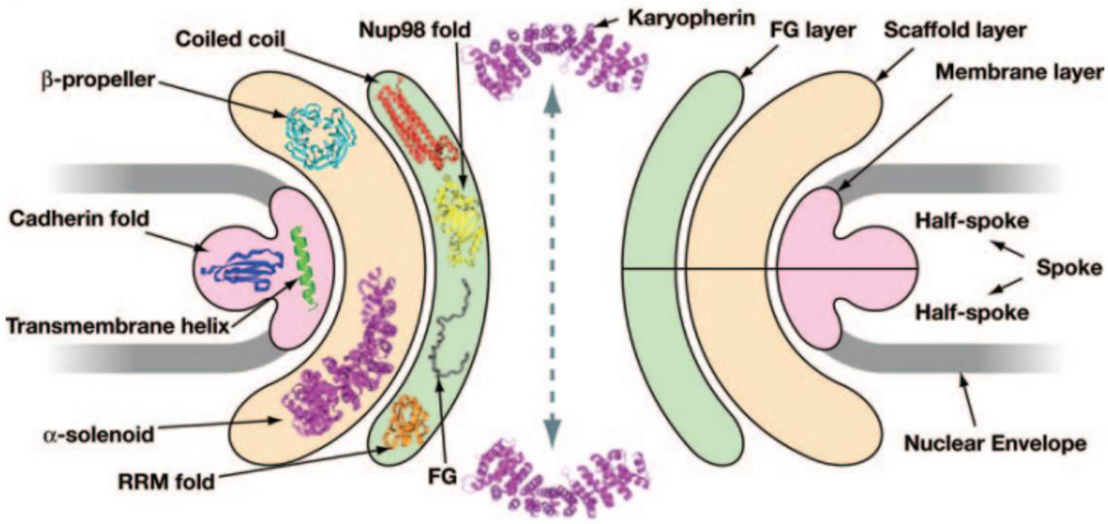


Figure 1.3. The three layers of Nups in the NPC structure. Adopted from (Devos2006).

The NPC has an octagonal, symmetric structure with eight identical segments, each of which called a spoke domain. Thus, Nups exist in a multiple of 8, 16, or 32 within the NPC. As depicted in the Fig. 1.1, the architectural elements of the NPC structure are cytoplasmic filaments, a cytoplasmic ring, a central channel that includes the spoke domains, a nuclear ring, and a nuclear basket that is composed of nuclear rods and a distal ring.

The NPC acts as a freeway for passive diffusion of molecules and ions smaller than ~5 nm in diameter (Yang2006, Mohr2009), while actively controls and facilitates the transport of larger cargos up to about 39-40 nm (Ribbeck2001, Pante2002). A single NPC can accommodate a remarkable rate of transport on the order of ~1000 translocations/sec, corresponding to a mass flow of ~100 MDa/sec (Ribbeck2001). The facilitated transport proceeds by discriminating between inert, i.e. nonspecific, and active, i.e. karyopherin-bound cargos. Karyopherins (kaps) are a family of soluble proteins that broadly are referred to as nuclear transport receptor (NTR) and act as chaperons for cargos across the NPC. The best-known kaps are importins α and β, also called kap α and β (Ben-Efraim2001a, Chook2001, Bayliss2002b, Weis2002, Isgro2005b, Yang2006).

A large cargo that is going to be transported across the NPC, first has to bind to an appropriate NTR via nuclear localization/export signal (NLS/NES) depending on import/export to form the

so-called cargo-complex (Macara2001a, Rout2003). Interacting with FG-repeats, the NTR escorts the cargo along the pore all the way to the destination compartment.

In the case of kap family, for example, once the cargo complex reaches the destination, the kap is dissociated from the cargo by RanGTP/GAP (import/export) through a conformational change process in the kap (Macara2001a). Similar to the other members of Ras superfamily, Ran acts in nucleocytoplasmic transport as a molecular binary with GTP- and GDP-bound states (Macara2001a, Terry2009). If it is an import process into the nucleus, the kap β is recycled back to the cytoplasm in complex with RanGTP. In the cytoplasm, the kap is released from the RanGTP via RanGAP (Ran GTPase activating protein)-mediated hydrolysis. RanGAP is localized to the cytoplasmic side of the NPC and simulates the hydrolysis of the RanGTP into RanGDP (Stewart2007). Finally, the RanGDP is imported back to the nucleus by the small homodimeric nuclear transport factor 2 (NTF2)/p10 (Weis2002, Stewart2007). Once in the nucleus, the guanine nucleotide exchange factor (RanGEF, also called RCC1) (Stewart2007) triggers the disassembly of the RanGDP-NTF2 complex via replacing GDP by GTP (Macara2001b). The small size of the NTF2 terminates the transport cycle. NTF2 has a diameter of about 5 nm (Ribbeck2001), and is small enough that freely diffuses across the NPC, meaning it does not need any other factor for its own transport.

It is believed that the key feature in nucleocytoplasmic transport is the interaction of NTR with these intrinsically disordered domains of FG-Nups, i.e. FG-repeat domains, or FG-repeats for short. NTRs have certain number of hydrophobic patches on their surface, and thus, have affinity with FG-repeats. This hydrophobic affinity is the key in letting a cargo-complex, i.e. a cargo that is bound to an NTR, favorably interact with FG-repeats.

1.3 Intrinsically disordered proteins and FG-repeats

Being ubiquitously common in all living organisms, intrinsically disordered proteins (IDPs) can be classified as important extensions of the protein kingdom (Tompa2002). Under physiological conditions, the structural dynamics, extremely intrinsic flexibility, and the presence of diverse functional motifs along their backbone enable IDPs to rapidly and promiscuously interact with different binding partners (Babu2012). As such, the lack of a folded structure confers a unique and diverse functionality to IDPs featured in a variety of cellular processes including signaling, cycle regulation, transcription, translation, and pathways control (Bah2015). Remarkably, IDPs also exhibit a stimulus-dependent response under different physical and geometrical conditions, opening new doors for exploring them as stimuli-responsive biopolymers. Thus, in recent years, a great deal of interest has been generated in fundamental understanding of their peculiar biophysics (Quiroz2015).

The aggregation of natively unfolded FG-repeat domains inside the NPC (Denning2003a) is one the impressive examples of IDPs that has provided inspirations for biomimetic materials and processes (Kabachinski2015, Kim2015). As per high degree of symmetry in the octagonal architecture of the NPC structure, each of 11 distinct FG-Nups is present in a multiple of 8, 16, or 32, similar to other Nups (von Appen2015). The FG-repeats are confined in, and tethered to the

interior of the NPC channel to form a ‘sticky soup’ of end-tethered, delicate biopolymers dangling inside the cramped space of the channel.

There is a wide consensus that it is this ‘sticky soup’ of IDPs that renders the NPC a sorting machinery (Koh2015). Heterogeneous mixture of FG-repeats establishes a robust selectivity barrier inside the NPC channel (Milles2015). Despite more than two decades of dedicated research, the mechanistic understanding of such a robust permeability barrier has not yet been reached, and is still under debate. This is mainly due to the lack of a comprehensive understanding of how FG-repeats precisely behave, and which conformations they adopt within the NPC channel in vivo. The concentrated, heterogeneous mixture of delicate FG-repeats, featuring a large number of hydrophobic and charged residues, densely end-tethered to and confined within the nanometer-sized NPC channel, have concertedly made it impractical to inspect their precise conformational behavior not only in vivo, but even in vitro as well. Today, there is no imaging technique capable of visualizing FG-repeats within the NPC channel in their functional state (Milles2015).

The lack of detailed information about the conformational behavior of FG-repeats has left plenty of room for speculation on the characteristics and formation of the FG-repeats-driven permeability barrier. Strikingly, depending on one’s interpretation of FG-repeats conformation inside the NPC channel, one can propose different models for nucleocytoplasmic transport. Consequently, over the past 20 years more than 15 hypothetical models have been proposed for the transport across the NPC (Simon1992, Shulga2000, Ben-Efraim2001b, Macara2001b, Ribbeck2002, Pyhtila2003a, Rout2003, Peters2005, Lim2006, Timney2006, Patel2007, Strambio-De-Castillia2010a, Yamada2010), many of which are no longer of interest. Indeed, distinctions between the different models originate from different speculations about the conformational behavior of FG-repeats inside the NPC.

It is thus crucial to first develop a clear understanding of the conformational behavior of FG-Nups per se without the mass flux bias across the NPC (Eibauer2015). Although insightful, the available conformational studies on FG-repeats have mainly focused on the single-type FG-Nup in an unconfined geometry, non-tethered state, where the results are extrapolated to the heterogeneous mixture of all FG-Nups, end-tethered to the nanometer-sized confined geometry of the NPC channel (Milles2015). There are other valuable studies on the conformation of heterogeneous mixture of FG-Nups (Yamada2010), or the energy landscape within the pore (Tagliazucchi2013), but they lack dynamical aspects and the Brownian motion of the FG-repeats.

As I will review in detail in the next chapter, experimentally-driven computational biophysical modeling approaches can contribute toward filling the gap by providing unique microdynamical insights at high spatiotemporal resolutions.

1.4 Emerging trends in further dissecting the role of the NPC in cellular processes

Importantly, individual Nups are directly related to several human diseases including influenza, cancers such as leukemia and inflammatory myofibroblastic tumors, and less frequent diseases like triple-A syndrome and primary biliary cirrhosis (Kasper1999, Gustin2001, Cronshaw2004, Satterly2007). Nups also play an important role in viral infections by providing docking sites for viral capsids, and also by blocking host cell mRNA export or inhibiting import of antiviral signals (Whittaker2000, Buehler2009). As such, more recent trends in the NPC-related research field

suggest that Nups are arbitrarily defined as they might have different functions and mobile nature. Interestingly, however, the main function of FG-Nups is still believed to be establishing the selectivity barrier across the NPC and regulating the transport (Aitchison2012b).

Interestingly, a more recent proteomic study performed on all of the Nups in the human NPCs of different cell types reported that the abundances of peripheral Nups are rearranged as a function of specific cell type, while the scaffold Nups are mainly static (Ori2013). The study suggests that the NPC Nups are fine tuned to comply with the 'particularities' arising from a specific cell type (Ori2013). Significantly, the study found that disease-specific cells, like cancer and inflammatory disorders, also show different abundance of dynamic Nups in the NPC periphery, supporting the notion of the NPC having different fine-tuned structure depending on the cell pathogenesis (Ori2013). The findings of this study further confirm that peripherical Nups are much more dynamic than the central Nups, consistent with another report where Rabut et. al. (2004) showed that the turnover times of Nups cover five orders of magnitude, ranging from several seconds to several days (Rabut2004). But, more importantly, Ori et. al. (2013) study also implies that the peripheral Nups are the key Nups in relation of the NPC to different diseases and cell types.

Emerging trends in the NPC research also show growingly a substantial role of this sorting machinery and it components, i.e. Nups, in transcriptional regulation (Sexton2007, Strambio-De-Castillia2010b). As far of the role of Nups in the gene regulation is concerned, to the best of my knowledge, for the first time it was Nup60 along with Nup145 in the yeast NPC that found to be required for the full repression of specific locus in the yeast genome, called HMR locus (Feuerbach2002). This and other studies triggered the notion that the NPC was actively associated with transcriptional regulation, and thus numerous studies conducted to reveal the role of different Nups in suppression or activation of genes in different species (Casolari2004, Mendjan2006, Capelson2010, Vaquerizas2010). It is, for example, known that the localization, i.e. spatial arrangement, of individual genes within the nucleus affects their expression (Light2010), and thus, the gene proximity to the nuclear periphery. The gene proximity to the nuclear periphery, wherein lies the nuclear side of the NPC, has been heavily associated with transcriptional repression (Takizawa2008). On the other hand, nucleoporins Sec13, Nup98, and Nup88 along with some of FG-Nups are shown to be required for the active expression of 'developmetnally regulated genes'—genes that are restricted in their expression to a specific developmental timepoint—by binding to the chromatin in *Droshophila* (Capelson2010). More specifically, Nup98, Sec13, and FG-Nups bind to active transcription sites, while Nup88 binds to silent loci on chromatin (Fig. 1.4).

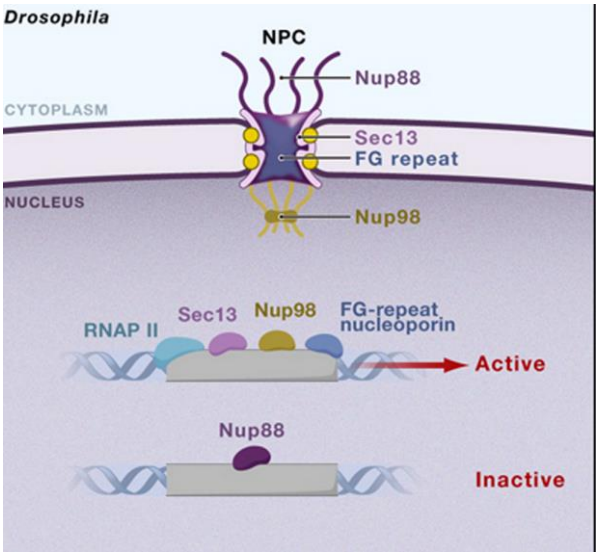


Figure 1.4 Four Nups identified in *Drosophila* to bind to different loci on chromatin, showing the role of Nups in gene regulation (Capelson2010).

The research on this aspect of the NPC has witnessed an ongoing progress and some Nups in the yeast, human, and metazoan have been shown to be involved in a phenomenon known as 'transcriptional memory' (Light2013, Pascual-Garcia2017). Transcriptional memory is a mechanism through which a gene is activated and transcribed much more rapidly and more robust if it has been expressed previously (Francis2001, Brickner2010). The mechanistic understanding of how a gene can 'remember' its past generations is of central interest in dissecting transcriptional memory. Most recently, Nup98 was found to be actively involved in the mechanism of transcriptional memory through mediating the enhancer-promoter looping of inducible genes (Pascual-Garcia2017).

These aspects of the NPC are less explored and much more work has to be done, both computationally and experimentally.

Chapter 2: Different numerical approaches to study the NPC and nucleocytoplasmic transport

During the past two decades there have been conducted numerous computational studies on the NPC and nucleocytoplasmic transport to elucidate parts of this big puzzle, that are not otherwise readily accessible to bench-based approaches. These computational studies can be categorized based on their approaches as follows: i) Analytical ii) Probabilistic iii) Structural modeling iv) Sequenced-based analysis and v) Biophysical modeling.

Here, I briefly talk about the first four approaches along with their accomplishments to date. Then I will go in depth to Biophysical modeling with a focus on the coarse-grained modeling in Brownian dynamics as a landmark in the NPC real-time dynamical studies.

2.1 Analytical approaches

Analytical approaches are mainly based on the mathematical analysis of the NPC and the transport process. Perhaps one of the earliest in-silico methodologies used in this context was based on the systems analysis, where the Ran transport across the NPC was modeled by a reaction-diffusion type of equation (PDE) (Smith2002). By merging the ideas of diffusion and rate equations, the reaction-diffusion equation simultaneously accounts for the space and time dependence under the assumption that various reactants and products can diffuse freely and fully explore the space (Phillips2012). This is indeed an idealistic assumption. Nevertheless, using this method, authors found that the crucial parameter for the steady-state Ran flux between the nuclear and cytoplasm is the Ran exchange factor RCC1 (Fig. 2.1). The model also provided the first estimate for the transport rate across the NPC as 520 molecule per second (Smith2002).

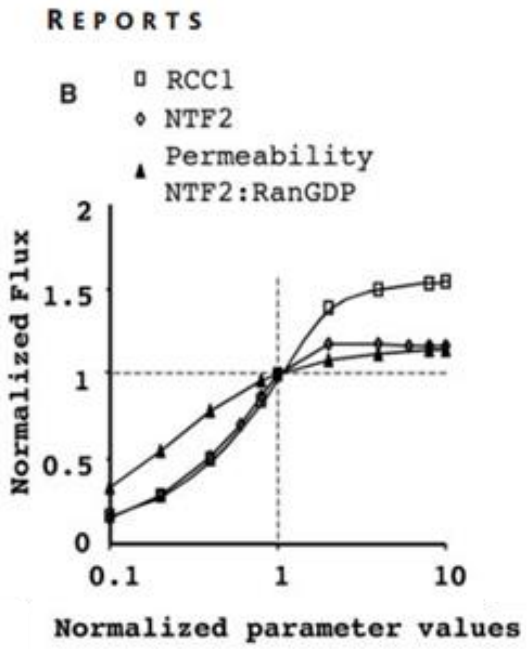


Figure 2.1. The effect of RCC1, NTF2, and RanGDP concentration on Ran flux across the NPC. RCC1 has the most dominant effect both upon decreasing and increasing its concentration. Adopted from (Smith2002)

In a later study (Görlich2003), the systems analysis approach was used to investigate how the

concentration gradient of the Ran interplayed with the NTR Impß. This simulation predicted a high concentration gradient of RanGTP across the NPC (~1000-fold increase in the nucleus compared to cytoplasm) and that the limiting factor for this steepness was cytoplasmic RanGAP, but not RCC1. The systems analysis approach was also used to shed light on the limiting rates in the process of cargo import. It was found that the cargo import was predominantly limited by the NTR Impα, Ran, and also NTF2 (Riddick2005).

More recently, ordinary differential equation (ODE)-based mathematical analyses (rate equations) have been used to reveal more details of the nucleocytoplasmic kinetics (Tetenbaum-Novatt2012). Authors concluded that NTR-FG interactions were strongly affected by a large collection of weakly-associated intra-NPC medium competitors. This would lead to reducing the affinity of NTR-FG several orders of magnitude, which is usually ignored in vitro. Without this consideration, the in-vitro reported values of NTR-FG avidities are too tight to be able to account for the rapid transport kinetics across the NPC in vivo (Fig. 2.2).

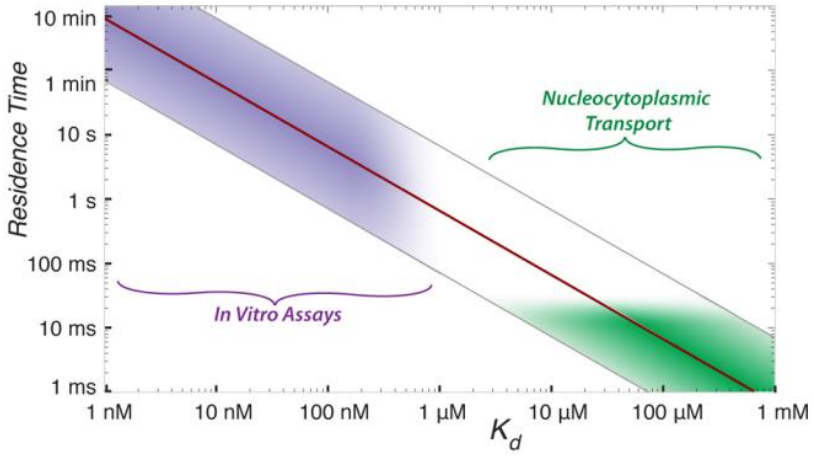


Figure 2.2. The alleged inconsistency between in-vitro and in-vivo NTR-FG affinities. Adopted from (Tetenbaum-Novatt2012)

Along with analytical approached, a diffusion-based theory was also developed to explain the efficiency, selectivity, and robustness of the transport across the NPC (Zilman2007), which attracted major attentions in the field. In this study, the authors attempted to merge a rigorous mathematical model of nucleocytoplasmic transport with the physics of diffusion in a channel. Base on their theoretical model, authors conduced that the selectivity of the NPC stems from a balance between efficiency and transport speed. Their model also suggests that there should be an optimal binding affinity between NTR and FG-repeats that maximizes the transport efficiency (Zilman2007).

In a follow-up study, within the framework of general kinetic theory, these authors used a set of three master equations (ODEs) to model the transport of specific and nonspecific particles across a selective narrow channel, mimicking the NPC (Zilman2010). This minimal kinetic model has only two essential elements: i) trapping of particles inside the channel is transient, and ii) there is competition for the limited space inside the channel (Zilman2010). The model predicts that inter-species competition would boost the selectivity in the channel, and that non-specific particles are effectively prohibited from entering the game of translocation through the channel. This is, in part,

because of the longer trapped time of specific species in the channel. Similarly, in the case of the mixture of two different specific particles, one with longer trapped time in the channel, the particles that are more weakly associated with the channel (and thus shorter trapped time) is strongly inhibited from transport (Zilman2010). These are indeed consistent with other findings on the NPC and nucleocytoplasmic transport. For example, we know inert cargo (non-specific in their model) are effectively rejected from entering the nuclear pore. Moreover, a cargo-complex with larger number of NTRs (in their model, species with longer trapped time) is more likely to overcome the selectively barrier compared to a cargo-complex with one NTR bound. Furthermore, there exists experimental evidence that in the presence of NTRs (specific species), the inhibition of inert cargos is significantly enhanced (Jovanovic-Talisman2009a), suggesting a boost in the selectivity, consistent with this analytical model.

In the analytical landscape of polymer study, people have also used the combination of the statistical mechanical methods and polymer physics principles to develop a semi molecular theory framework in which the most probable conformations of hydrophobic end-tethered polymers can be calculated (Carignano1994, Carignano1995, Szleifer2000, Nap2006). This approach is based on searching the phase space of the polymer conformations and calculating the probability of every conformation through the functional minimization of the free energy. Different physical and geometrical factors, including shape, size, conformational degrees of freedom, solvent nature (i.e. good solvent vs. bad solvent (Rubinstein2003)) quality, and intraparticle van der Waals (vdW) interactions are taken into account in this theoretical framework. The framework has been implemented for more complex geometry of a cylindrical channel with polymers end-tethered to the interior, resembling the NPC (Peleg2011). In this study, the theoretical framework suggests that the conformations of polymer chains heavily depend on the solvent quality and the grafting density. The prediction of this study (Fig. 2.9) is that in a poor solvent grafted polymers collapse on the channel wall, while in a good solvent the grafted chains stretch out of the nuclear pore, relaxing the internal osmotic pressure and maximizing the entropy (Peleg2011).

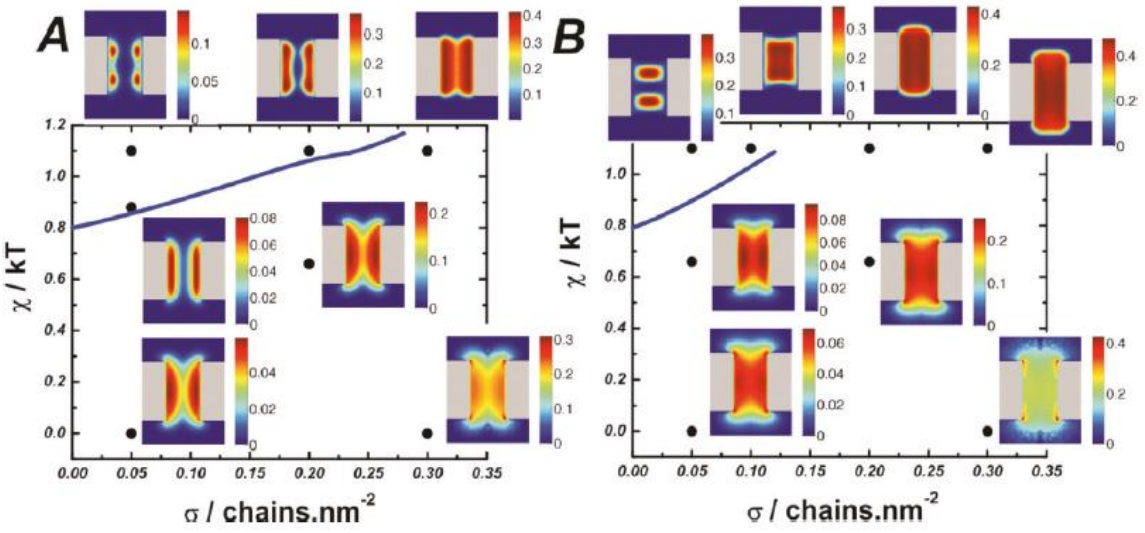


Figure 2.9. The morphology phase diagram of the grafted chain polymers in the cylindrical channel calculated by the molecular theory framework for grafted polymers (Peleg2011). The parameter χ is the monomer-monomer interaction strength, which indicates the solvent quality. Parameter σ is the grafting density. Adopted from (Peleg2011)

Under the analytical approaches, there are also theoretical studies that employ pure mathematical formalisms to derive different formulations for nucleocytoplasmic transport. For example, a recent study developed a mathematical framework for the diffusion of particles in the elastic medium (Chakrabarti2014), modeling the NPC as a continuum elastic polymer meshwork with cylindrical symmetry. For a particle to enter this medium, it must overcome the elastic energy that is assumed to be quadratic in the number of binding proteins. By adjusting the elastic constants in this model, authors could retrieve the behavior of a polymer gel and brush medium. Moreover, an analytical expression for the diffusion coefficient of particles moving in elastic medium is derived, which is suggested to govern the diffusion of macromolecule in the NPC (Chakrabarti2014). Another study, used the mathematical formalism of the generalized Smoluchowski equation in the cylindrical domain in a viscoelastic medium subject to appropriate boundary conditions (Chatterjee2011). Particularly, for a given cargo traversing the viscoelastic medium across the NPC, authors derive complex mathematical expressions to predict the mean-square-displacement, the distribution of the mean first-passage time, and the correlation of intensity fluctuations (Chatterjee2011). As these models are based on pure mathematics there are too many idealized assumptions, which perhaps make those models less attractive. Nevertheless, a single complicated mathematical equation based on pure theoretical assumptions does not seem to be able to predict the nucleocytoplasmic transport in the crowded, heterogenous environment of the NPC.

Overall, the analytical approach can be useful in dissecting nucleocytoplasmic transport, particularly when it is related to measure the kinetics of the transport, so long as there are corroborating bench validations (Smith2002, Tetenbaum-Novatt2012).

2.2 Probabilistic approaches

Probabilistic computational modeling approaches are generally based on random moves in the phase space, with each movement can either be favored or disfavored, depending on its local and/or global effects on the entire system configuration. One of the oldest and widely-used probabilistic modeling is the Monte Carlo (MC) method. The foundation of the MC is based on a simple idea of random sampling from the ensemble in equilibrium (Binder1993). Instead of visiting all microstates in the phase space, one can randomly choose points (microstates) from the phase space, and by averaging over them, one can hope to get a good estimation of a desired observable of the system (Pathria). This is, however, the simplest way of sampling, 'simple sampling', which has certain limitations and is not so useful for evaluating complicated systems beyond free random walk. Metropolis et. al. laid the foundation of the modern MC by proposing a probabilistic approach using the Boltzmann factor to randomly choose the microstates: "instead of choosing configurations randomly, then weighting them with $\exp(-\beta E)$, we choose configurations with a probability $\exp(-\beta E)$ and weight them evenly" (Metropolis1953). In each stepwise move, the change in the energy of the system, ΔE, is the defining factor to accept the move (if $\Delta E < 0$) or to conditionally accept the move (if $\Delta E > 0$). This 'importance sampling' is still widely used within the MC modeling.

In the context of the NPC and nucleocytoplasmic transport, the MC methodology has been employed at two different levels: first, as an on-the-side supporting computational framework, accompanying and corroborating the main experiment in studying the kinetics of transport or finding some local probability densities (Beck2007). (Beck2007, Dange2008b, Herrmann2009). The second level of the MC modeling is when scientists use it as a holistic, standalone model to mimic the entire NPC and the nucleocytoplasmic transport. Here, I review the latter level of the MC modeling for the NPC.

In one of the initial studies that used the MC approach to explore the NPC, FG-repeats were modeled as flexible polymers represented by connecting beads (Nielsen2006). Authors implemented two versions of the model: a torus-shaped pore in a place and a cylindrical-shaped pore (Fig. 2.3). Authors concluded that in both implementations, the present of an affinity, i.e. probability in the model, between FG-repeats the cargo would reduce the energy barrier and facilitates the entry of the cargo into the pore.

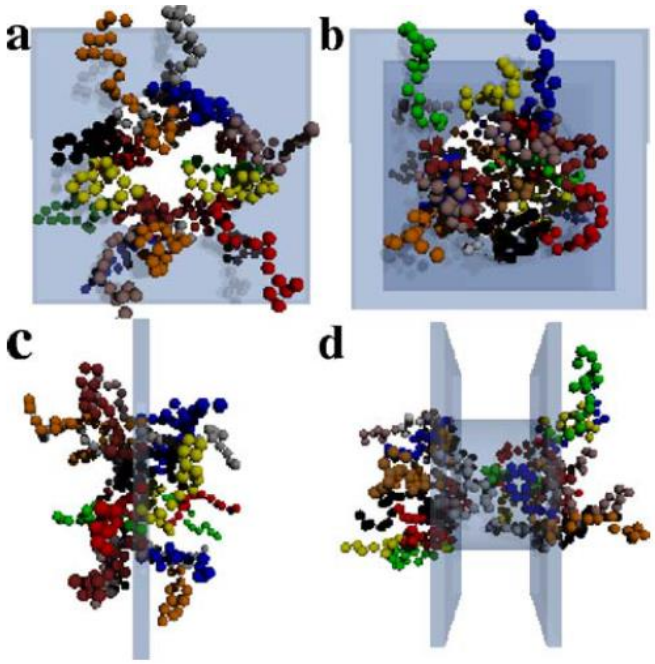


Figure 2.3. Two possible representations of the NPC and FG-repeats in an initial MC model of the NPC. a and c show the hypothetical torus-shaped NPC (top and side views), while b and d show cylindrical-shaped NPC. Adapted from (Nielsen2006).

Another minimalistic MC modeling study (Diesinger2010) tried to answer the question of whether the hydrophobic affinity between FG-repeats was sufficient to produce a continuous percolating meshwork hydrogel, as had been observed experimentally at macroscopic level (Frey2006b, Frey2007, Frey2009, Mohr2009, Ader2010). To do this, 40 FG-containing short peptides from the yeast Nsp1 were brought into contact in the simulation box and the hydrophobic 3D clustering was examined (Diesinger2010). To determine the critical density at which the gelation occurs and the sol-gel transition proceeds, authors used the theoretical framework invented by the pioneering polymer scientists Flory (Flory1941, Flory1953) and Stockmayer (Stockmayer1943). The prediction of the MC modeling was that the hydrophobic interactions between the short peptides were indeed sufficient to overcome the entropic forces from excluded volume effects to form a 3D cluster, i.e. a gel system. The critical density for this sol-gel transition was found to be 42 mg/mL (Diesinger2010). This simple modeling, however, did not take into account the electrostatic and/or interactions among the side chains that can potentially shift the critical density.

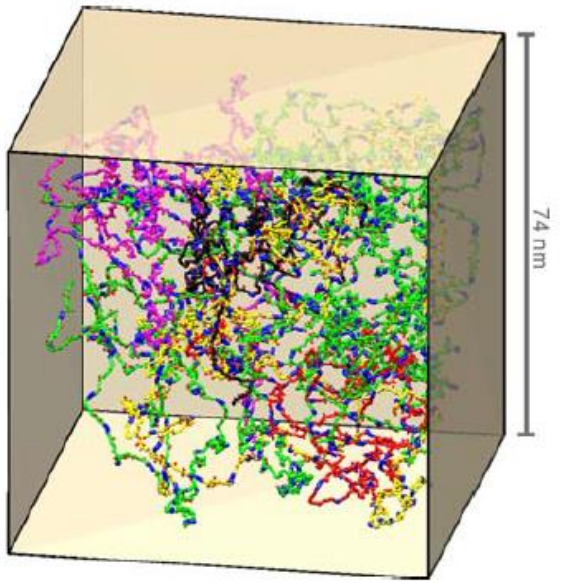


Figure 2.4. The MC simulation box consists of 40 short FG-containing peptides of the N-terminal Nsp1 from the yeast to study the gel formation capability of the FG-repeats in silico. Adopted from (Diesinger2010).

In a series of studies using the MC to study the NPC, FG-repeats were modeled as a simple monolayer composed of 40 polymers, each 100-nm in length and end-grafted in the middle of a 50-nm-diameter cylindrical channel in the presence or absence of large quantities of 8-nm spherical particles representing the NTRs (Osmanovic2012, Osmanović2013, Osmanović2015). All of the 40 polymers were identical and repeated in a rotational symmetrical manner inside the cylinder.

As one of their initial results, authors suggest that the NTRs appear to be a part of the selective barrier via forming a cross-linked network with FG-repeats. Their model also shows that the interactions of NTRs with FG-repeats can create considerable rearrangement in the monolayer polymer meshwork (Osmanović2013), which is perhaps expected given the entire FG-repeat domains are modeled as a single layer of 40 polymers. Moreover, this MC model suggests that depending on the model parameters, the polymers may adopt two different phases, namely wall phase or central phase (Osmanovic2012). While in the former the attractive monomer accumulate in the center of the cylinder, the latter scenario predicts those attractive monomers form local islands near the channel wall (Fig. 2.5).

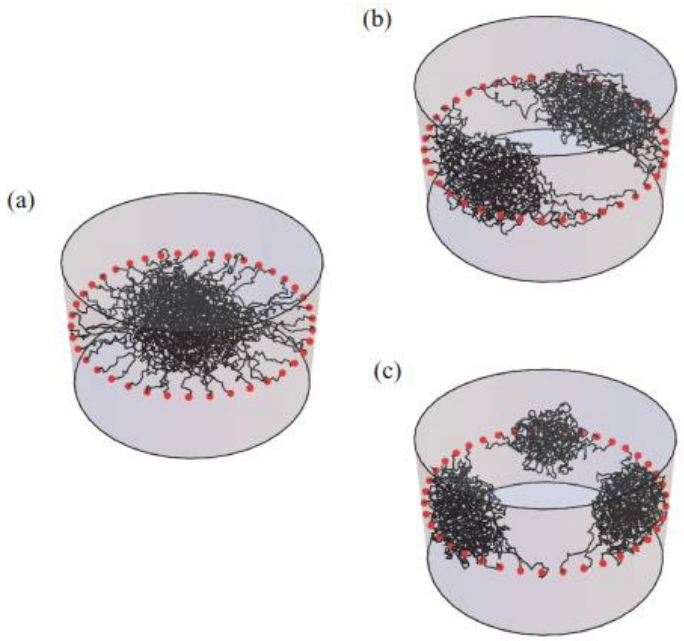


Figure 2.5. The formation of two different phases of polymers inside the cylindrical channel. (a) Central phase (b)&(c) wall phase. Adopted from (Osmanovic2012)

Interestingly, authors also used the predictions of the density functional theory[1] calculation under the assumption of rotational symmetry around the central axis of the cylinder. The polymers' equilibrium densities obtained from the density functional theory and the MC simulations under the perfect rotational symmetry assumption were in good agreement (Osmanović2013). However, in their following study, they removed the artificially imposed rotational symmetry in the polymeric monolayer and observed significant disparities in the equilibrium densities of rotationally symmetric and non-symmetric systems (Osmanović2015). This particularly becomes more obvious as the FG-FG affinity increases (Figure 2.7). The results of these two studies conducted by the same group suggest that theoretical methods like density functional theory are perhaps too simplistic to be able to predict the complicated dynamical physics of the NPC and FG-repeat domains. In contrast, even a simplistic MC simulation can produce more reliable view of the NPC dynamical behavior. It also highlights the importance of the dynamics of the polymers which plays a major role in the conformation of polymers.

[1] The density functional theory was originally developed to study a many-electron system by single-particle equations Grotendorst, J., N. Attig, S. Blügel and D. Marx (2009). "Multiscale simulation methods in molecular sciences." Lecture Notes, NIC Series **42**. But the idea has been adopted with certain modifications to study the many-polymer system with a single density equation Osmanovic, D., J. Bailey, A. H. Harker, A. Fassati, B. W. Hoogenboom and I. J. Ford (2012). "Bistable collective behavior of polymers tethered in a nanopore." Physical Review E **85**(6): 061917..

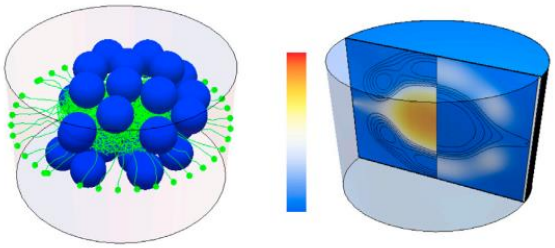


Figure 2.6. *Right panel:* The snapshot of the MC model of a monolayer of FG-repeats (green) along with NTRs (blue spheres). *Left panel:* The prediction of the density functional theory for the distribution of FG-beads and NTRs around the central axis of the channel. Adopted from (Osmanović2013)

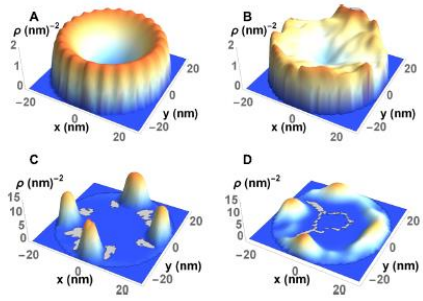


Figure 2.7 Disparities in equilibrium density of polymeric FG-monolayers predicted by the density functional theory (right) and the MC simulation (left) for weak (up) and strong (down) FG-FG affinities. Adopted from (Osmanović2015).

Agent-based modeling is another powerful probabilistic approach in which the spatial domain of the system of interest is discretized into lattices called 'cells' and the constituting particles of the systems are referred to as 'agents' (Azimi2011). Agents are, in effect, autonomous decision-making entities that collectively govern the ultimate behavior of the system on the basis of certain rules (Bonabeau2002). In each timestep, the agents are able to interact with each other and to hop between cells with certain probabilities under the condition that two agents cannot occupy the same cell at the same time. By letting the decision-making agents interact and run for statistically sufficient number of timesteps, ABM can capture the emergent phenomena that cannot be seen or predicted immediately in the system, but rather are the outcomes of the long-term dynamical collective behavior (Bonabeau2002). This methodology was first developed in the context of the social sciences in studying the collective and individual response of population of human under different conditions, like building fire (Still1993), or panic evacuation (Helbing2000).

Perhaps the simplest idea of the ABM to simulate the bidirectional transport through the NPC was first adopted in a minimalistic 1D lattice model based on non-equilibrium statistical mechanical principles, in which the NPC was treated as an open system (Kapon2008). In that model, authors introduced three sets of probabilities that derive the dynamics of the cargo movement in the 1D channel. Each set has two independent probabilities: probabilities of the cargo import from either sides of the channel, probabilities of the cargo export from either sides of the channel, probabilities of appearance and disappearance of a cargo at either sides of the channel (Kapon2008). The model had also a rather non-realistic assumption that once a cargo enters the pore, the probability of the cargo diffusing back is negligible, i.e. the cargo is forced to move forward. Nevertheless, using this simple model it was found that the dynamics of transport is dictated to a large extent by the rate of release of cargo from the NPC, which highlights the role of RanGTP availability at the nuclear entry and the RanGAP/RanBP1 at the cytoplasmic entry, which is in agreement with experimental observations. This underlines the strength and capacity of the probabilistic methodology in predicting features of nucleocytoplasmic transport, even when the model is so simple.

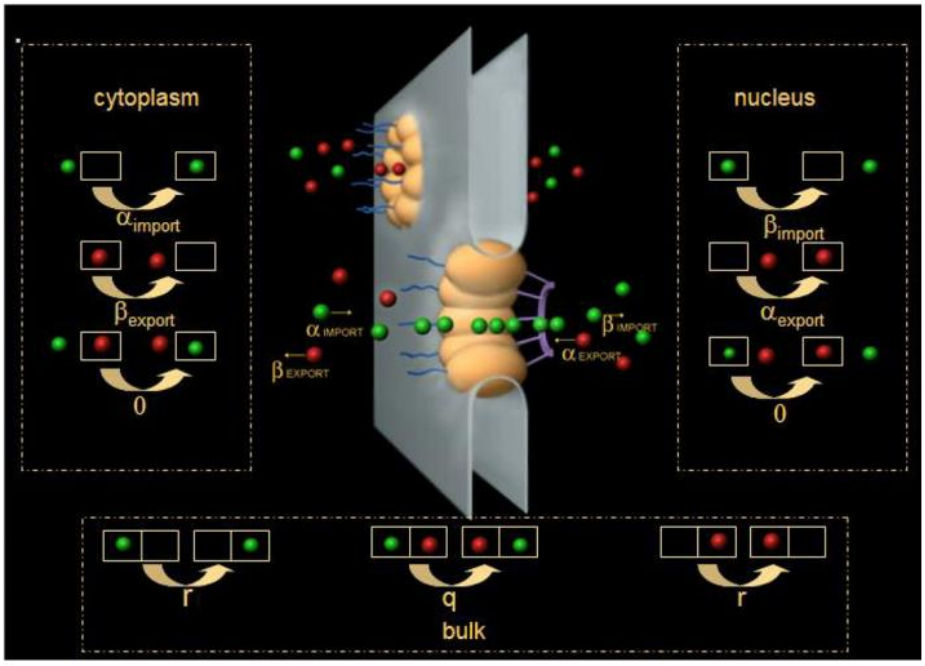


Figure 2.8. The single-value probabilistic lattice-based (ABM) model for importing (green) and exporting (red) particles across the NPC. Adopted from (Kapon2008).

In applying ABM to the biological subcellular systems, Azimi et. al. (2011) developed a biomimetic ABM methodology to consider different physiologically relevant properties in the model. In their ABM study of cytoskeletal diffusion, they consider properties like particle size, diffusion coefficients, affinity, viscosity, geometry, and crowdedness (Azimi2011). On this basis, they proposed a novel equation where the probabilities of movement are related to the diffusion coefficients of the agents in the crowded environment of the cell:

$$p_i = \frac{D_i \delta t}{\delta l^2} \qquad (2.1)$$

In this equation, p_i is the probability of movement of the agent i, D_i is its diffusion coefficient, δt is the timestep, and δl is the lattice discretization (Azimi2011). In this study, they also extended the traditional single-agent-per-cell ABM algorithm to a multi-particle per cell diffusion (see Fig. 2.10).

Using this approach in a sequel study, they shed light on the effects of affinity gradient in nucleocytoplasmic transport (Azimi2013). Indeed, the affinity gradient has been one of the known hypothetical models for the nucleocytoplasmic transport, and is based on the in-vitro observation that the affinity of Impß-Impα-bound cargo to three major FG-Nups—Nup42, Nup100, Nup1—increases as a function of the Nup's localization within the NPC. The complex of the Importin-bound cargo has a higher affinity to FG-Nups in the central channel than in the cytoplasmic entry, likewise it has higher affinity to FG-Nups localized to the nuclear basket than to the central channel (Ben-Efraim2001b, Pyhtila2003b). However, it was not clear, quantitatively, how this affinity

gradient affects the efficacy of transport, and what the sweet spot would be, if any. Moreover, following the affinity gradient model, some doubts were raised about the plausibility of this hypothesis for supporting the transport across the NPC (Strawn2004). The ABM study (Azimi2013) was the first computational simulation predicting that the existence of affinity gradient would indeed increase the rate of transport. In this study, the optimized macroscopic affinity gradient was quantified for the first time, and it was found to be 2000 µM, 200 µM, and 10 µM in cytoplasmic, central, and nuclear regions of the NPC, respectively (Azimi2013).

The ABM framework is also used to study the mRNA export across the NPC (Azimi2014). Despite the fundamental role of mRNA export in the cell biology, there are no experimental data on how the number and distribution of the export receptors on the mRNA backbone affect the transport process. By quantifying this process, the ABM study predicts that the mRNA export was sensitive to both numbers and distribution of receptors along the mRNA backbone. It was also found that during the export process, the rate-limiting step was the nuclear basket, corroborating the idea that mRNA undergo a threading process at the nuclear basket (Azimi2014).

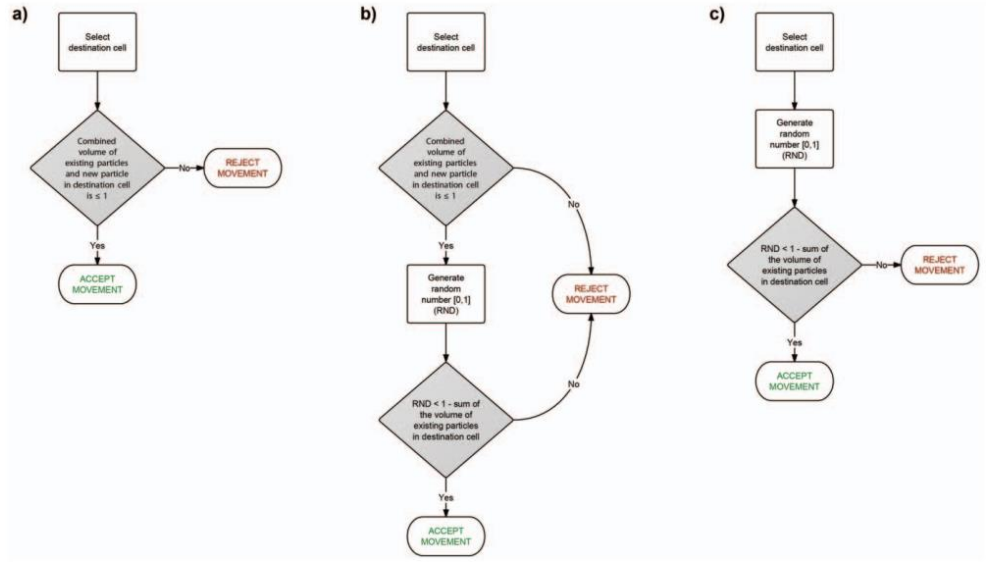


Figure 2. 10. Three different flowcharts of the algorithms of multi-particle per cell in the ABM study conducted by Azimi et. al. Adopted from (Azimi2011)

More recently, using the ABM Soheilypour et. al. (2016) were able to establish a minimal model of the mRNA quality control mechanism (Soheilypour2016). Using this model, the minimum binary requirements were identified for the quality control of mRNA prior to export: i) affinity of RNA-binding proteins to export receptor, and ii) interactions of RNA-binding proteins with the nuclear bask proteins (Tpr/Mlp1). One of the main findings was that the lower-than-normal affinity between RNA-binding proteins and the nuclear basket protein Tpr is the key factor in keeping the abnormal mRNA from exporting (Soheilypour2016). This was a significant finding of probabilistic pure computational modeling, showing that the level of affinity gauges the quality control in the mRNA export (Fig. 2.11).

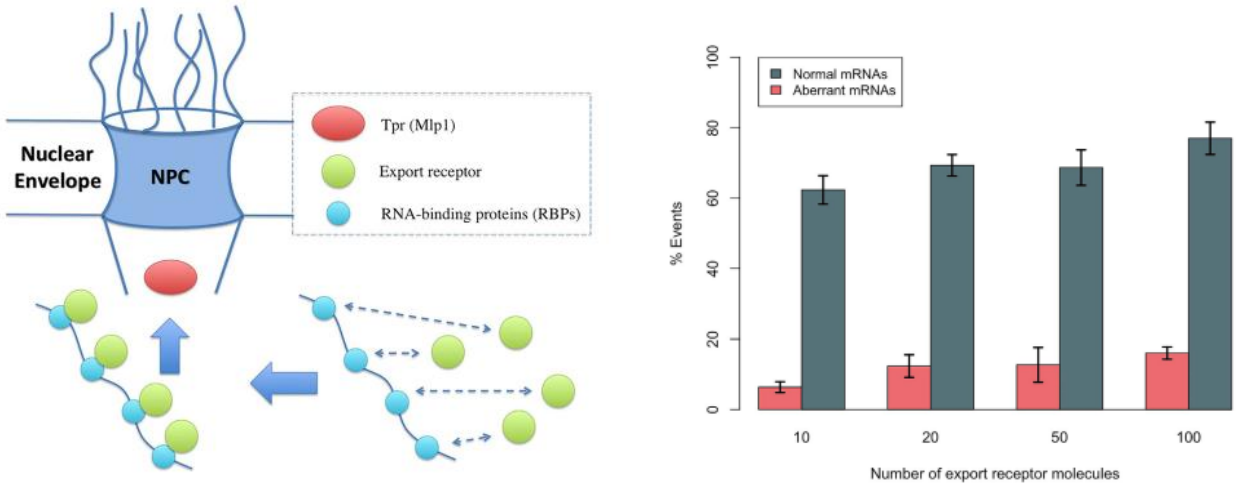


Figure 2.11. Left panel: The schematics of agent-based modeling of the minimal model for the quality control in the mRNA export process. Right panel: the export rate of mRNA when RNA-binding proteins have normal affinity with Tpr (grey) vs lower affinity (red). The data clearly show that the level of Tpr affinity plays a key role in holding mRNA in the nucleus or letting them pass through the pore. Adopted from (Soheilypour2016).

All in all, the probabilistic computational simulations of the entire NPC structure with the whole transport cycle using MC and ABM methodologies show a powerful and promising approach in dissecting the underlying factors in nucleocytoplasmic transport. While probabilistic approaches are not optimal for capturing the microdynamical details of nucleocytoplasmic transport and/or the NPC structural elements, they are well capable to point, for example, the key role-players in an efficient transport. Moreover, due to the coarse granularity of these approaches, which is a direct consequence of the absence of physical force-fields, both ABM and MC can span a very large temporal domain in simulating the NPC at minimal computational cost, without need to a high-performance computing system. For example, it is common for an optimized ABM of the entire NPC to simulate tens of minutes of the transport behavior. Additionally, since the dynamics of structural elements are absent, it is a straightforward task to simulate a large population of NPCs with ABM, albeit with reasonable computational resources.

2.3 Structural modeling

Computational structural modeling of the NPC aims to model the entire NPC structure, and more importantly the central channel, as an elastic material. The approach is based on a mainly static viewpoint to shed light on the stability and/or collective motions of the NPC in response to different mechanical stimuli, in particular stress. There have not been many computational works done on the NPC using this framework.

The finite element method (FEM) is perhaps one of the oldest and most profound computational modeling in structural analysis and the material properties calculation under different mechanical loads. As far as the NPC is of interest, the only pure FEM study on the NPC that I am aware of is conducted by Wolf et. al. (2008), in which the octagonality of the NPC structure was thoroughly examined (Wolf2008). Through a set of FEM analysis of the NPC central channel, this pioneering

work found that there was a mechanical rationale behind octagonality of the central channel in that the eightfold symmetry maximizes the bending stiffness of the spoke domains (Wolf2008). Interestingly, the authors also found that for the experimentally available values of the NPC dimensions in both yeast and vertebrate, the eight-fold symmetry in the FEM model optimizes the bending stiffness, as opposed to seven-, nine-, and 10-fold symmetries (Fig. 2.12).

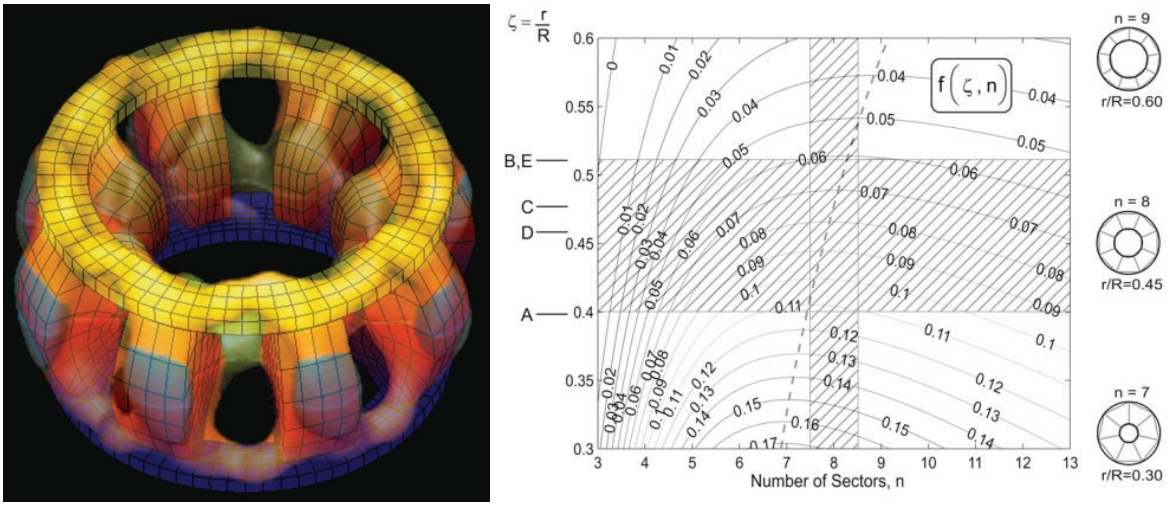


Figure 2.12 Left panel: the FEM model of the NPC. Right panel: the moment of inertia of a single spoke domain in the FEM model for different seven-, eight-, and nine-fold symmetries. Adopted from (Wolf 2008).

Another pioneering computational modeling under the structural approach used the elastic network model formalism to mimic the central channel of the NPC as a toroidal structure (Lezon2009). Independently from the FEM study, this work also found that eight-fold symmetry enables the NPC to undergo some certain modes of collective motions that are not accessible to other symmetrical structure (Lezon2009).

These two fully structural studies are overall consistent in terms of their predictions of the NPC symmetry and structural stability. It appears, however, that the extent of structural approaches is limited and there are few aspects of the NPC that can be addressed. This predominantly stems from the fact that structural approaches are primarily based on statistics, while the NPC and nucleocytoplasmic transport have extremely dynamic nature. Therefore, beyond the symmetrical structure investigation or stress distribution inside the structure, there are not too much that can be investigated by structural approach.

2.4 Sequence analysis.

Sequence analysis of different Nups, and in particular disordered domains, is another approach that has emerged more recently. In a broad sense and as far as related to the NPC, this approach takes the sequence composition of a certain Nup or a family of Nups and search for a common thread and link it to the functional state of the NPC and nucleocytoplasmic transport.

To the best of my knowledge, the first thorough sequence-analysis study on the NPC was conducted in 2010 to investigate the role of charge vs hydrophobicity in transport process

(Colwell2010b). In this study, all of the yeast Nups, including FG-Nups and scaffold Nups, were analyzed to find the charge and hydrophobicity distributions. The study found that FG-Nups are majorly characterized by a net positive charge and low hydrophobicity, while the scaffold Nups are predominantly negative and hydrophobic. The authors also analyzed the sequence of a number of NTRs and cargo proteins. Interestingly the finding was the NTRs are hydrophobic and negatively charged. This led the authors to propose a model in which charge also plays a role in transport across the NPC (Fig. 2.13).

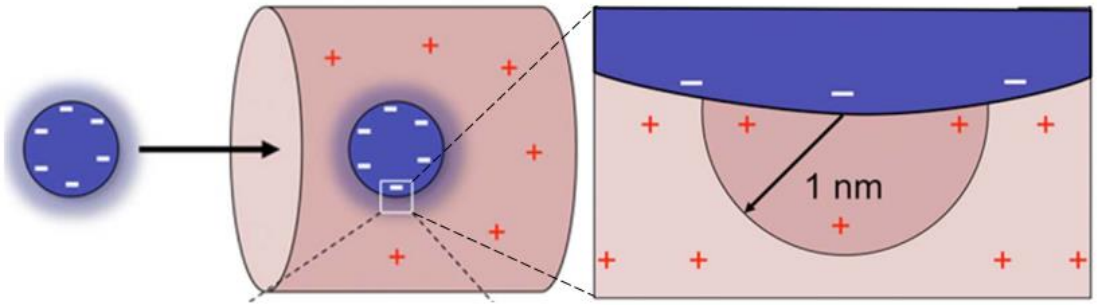


Figure 2.13 The hypothesized role of charge in nucleocytoplasmic transport, based on the sequence analysis of Nups and NTRs. Adopted from (Colwell2010b).

In a more comprehensive study, combining the sequence analysis with biophysical modeling of induvial FG-Nups, the amino-acid sequence of FG-Nups in 10 diverse eukaryotic species were analyzed (Ando2014) to shed light on the distribution of charge vs FG-motif density. The sequencing result was that FG-Nups share a common feature in that they all possess two distinct domains: one rich in FG-motifs, and the other high in charged residues. It was noted that these two domains are separated from each other in the sequence (Ando2014).

Along the same idea, a more extensive sequence-analysis study was conducted on sequences of more than 1,000 FG-Nups across 252 different species (Peyro2015c). This study extracted a comprehensive map of the charge distribution within FG-Nups (Fig. 2.14). The results revealed that polar residues and some specific sequence of charged residues, called like-charge regions, play a regulatory role in the formation of FG network inside the NPC (Peyro2015c). The study also suggests that like-charge regions with positive charges are favorable for negatively charged cargo complexes, and also regulate the FG-network size. Interestingly, their sequence analysis of FG-Nups and other intrinsically disordered proteins showed that FG-Nups do not share the same feature that other disordered proteins do. More specifically, while FG-Nups have significantly lower net charge compared to other disordered proteins, they share a similar hydrophobicity pattern with other disordered domains (Peyro2015c).

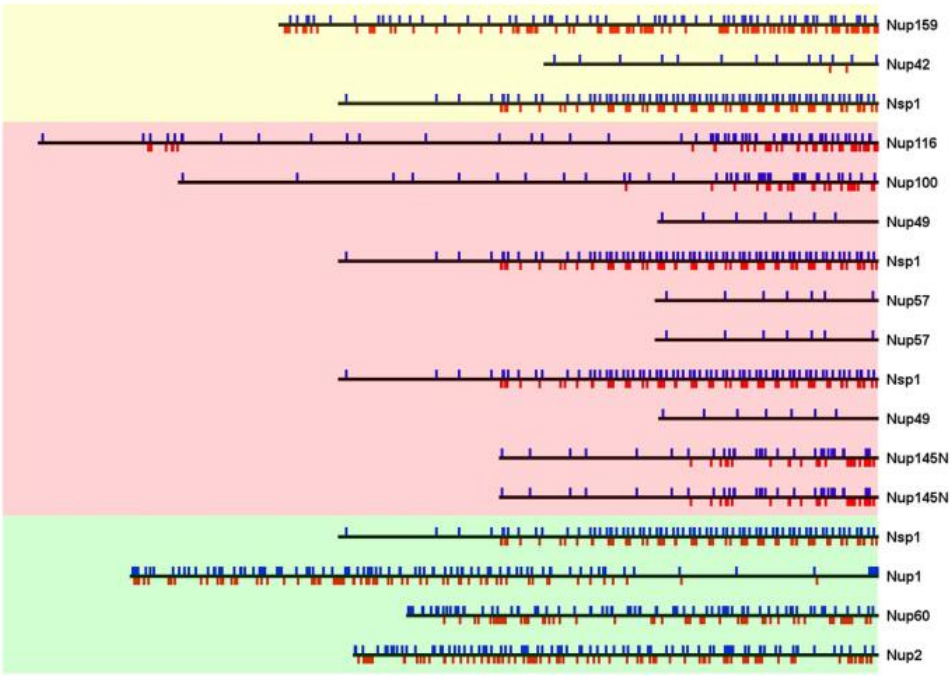


Figure 2.14. As an example, the yeast NPC is chosen in this figure to reveal the charge distribution along FG-Nups. Positive charges are in blue while negative charges red. Adopted from (Peyro2015c).

In a follow-up study, authors did a more in-depth analysis on the role of the like-charge regions by combining the bioinformatic sequence analysis with biophysical modeling (molecular dynamics) of specific domain of an FG-Nup (Peyro2015a). This strategy enabled the authors to include the dynamics of the FG-Nups into the game. In their biophysical modeling, the authors simulated the individual FG-Nups and by mutating the like-charge regions along the FG-Nups, they explored the role of those regions in the dynamic behavior of FG-Nups. The results showed that the like-charge regions in different FG-Nups are responsible for the larger RMSD and end-to-end lengths in wildtype FG-Nups when compared to mutated FG-Nups. As a result, the results suggest that the distribution of wildtype FG-Nups and mutated FG-Nups is governed, to a large extent, by the like-charge regions. For example, the collective behavior of the wildtype FG-Nups in terms of the time-averaged density map was different from mutated FG-Nups (mutation: charge residues > Alanine). The wildtype FG-Nups showed a more dynamic profile and a larger spread inside the pore compared to mutated FG-Nups (Fig. 2.15). This is, however, expected as all of the charged residues in the charge-like regions were replaced by a hydrophobic residue, i.e. Ala, which favors the collapsed conformation. Nevertheless, this combination of bioinformatics and biophysics can relate the collective behavior of FG-Nups to their residue composition.

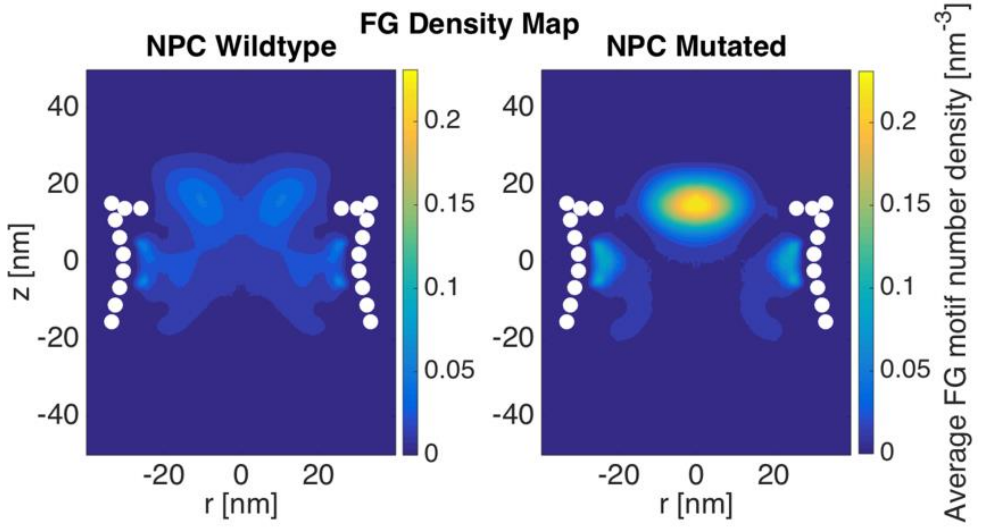


Figure 2.15. The effects of mutating the like-charge regions within the FG-Nups. The wildtype FG-Nups can cover a larger area within the pore over time compared to mutated FG-Nups. Adopted from (Peyro2015a).

Taken together, the bioinformatic approaches have recently appeared in analyzing different Nups, in particular FG-Nups, across a wide range of species. It can produce invaluable insights into the FG-Nups in terms of discovering the functional motifs along the amino-acid sequence by finding the evolutionary conservation analysis. This can pinpoint, with a good accuracy, the important motifs, like FG-motifs, that play key role the functional behavior of the selectivity barrier. Nevertheless, this approach cannot account for the dynamics of the FG-repeat domains, which is crucial in understanding how selectivity barrier regulates the transport across the NPC. Therefore, the more recent trend emerged in which the bioinformatic approach is conducted in combination with dynamical biophysical modeling (Ando2014, Peyro2015a). This kind of merging can indeed provide an unprecedentedly powerful analysis method in dissecting the selectivity barrier and identify dynamical behavior of the disordered domains down to specific functional regions inside the FG-Nups.

2.5 Biophysical modeling.

Computational biophysical modeling approaches aim to study the biological phenomena by integrating the equations of motion of constituting particles to find the time-evolved behavior of the system. This methodology has its deep roots in statistical mechanics. In particular, the coarse-grained modeling is employed to study the larger domains in time and space. In the next chapter, I will go over this approach in more details.

As far as the NPC and nucleocytoplasmic transport is of interest, collectively, the field has witnessed a much larger number of studies with this approach compared to other computational approaches discussed earlier. This can be best understood by the fact that the NPC and nucleocytoplasmic transport are extremely dynamic in nature, and therefore, time evolution is a central factor to dissect it. Within this approach, different levels of granularity, from all-atom MD,

to coarse-grained MD, to Langevin dynamics and Brownian dynamics exist. Thus far all of these scales have been performed to model the NPC or its subdomains to shed light on the dynamics of transport (Kustanovich2004, Isgro2005a, Isgro2007, Isgro2007a, Miao2009a, Dolker2010, Miao2010a, Mincer2011, Moussavi-Baygi2011a, Moussavi-Baygi2011b, Tagliazucchi2013, Ando2014, Gamini2014, Ghavami2014, Zhao2014, Coalson2015, Peyro2015b, Ghavami2016, Moussavi-Baygi2016, Osmanovic2017, Pulupa2017). Some simulations are combined with experiments as well (Yamada2010, Popken2015b). I will go over only some of these studies in this section.

One of the earliest and pioneering all-atom MD study on the NPC was conducted on the interactions of small-chopped FG-repeat domains with Importin-ß (Isgro2005a). Before this computational study, there had been performed many experimental studies both in vitro (Radu1995, Rexach1995, Chi1997, Shah1998, Kehlenbach1999, Kose1999, Seedorf1999) and in vivo (Iovine1995, Damelin2000), showing the chaperon proteins from the Importin-ß family were interacting with FG-Nups and Tpr. In particular, several hydrophobic patches on Importin-ß had been proposed based on analyzing the crystal structure of Importin-beta complexed with FG-Nups (Bayliss2000b, Bayliss2002b) as well biochemical analysis of the interaction between Importin-beta and FG-Nups (Bednenko2003). Nevertheless, those experimental studies fell short of analyzing the dynamics of interactions in terms of biding stability and atomistic details.

The results of the MD simulations mainly through energy landscape analysis predicted 10 hydrophobic patches on the surface of Importin-beta served as binding spots for FG-Nups. Seven of those were consistent with known experimental observations, while three were predicted as new binding sites (Isgro2005a).

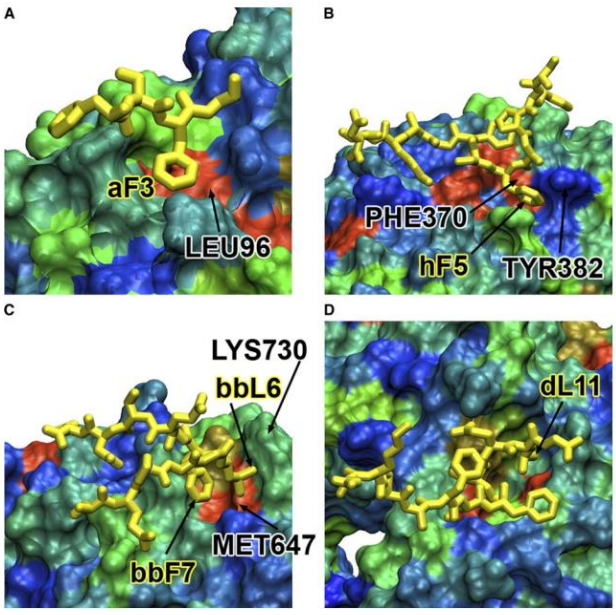


Figure 2.16. The view of four binding spots of Importin-Beta for FG-Nups obtained in a series of MD simulations. Adopted from (Isgro2005a).

Although it suffers from small chops of specific domains of a single FG-Nup artificially arranged around the Importin-ß, this study proved the power of MD approach in analyzing the dynamics of NPC and transient interactions therein at atomistic details.

Using this framework, the same group performed a similar study on investigating the binding of

small-chopped FG-repeats to the NTF2 surface, confirmed four known binding spots on the NTF2 surface and predicted two novel binding sites (Isgro2007a).

Following this approach, we investigated the interaction of Crm1, a major exportin, with C-terminal domains of the Tpr, a nuclear basket nucleoporin that dangles out inside the nucleus (Zhao2014). Before our MD study, Tpr had been hypothesized to have a role in mRNA export (Zimowska1997, Bangs1998). This notion was later corroborated by independent observations that Tpr-deficient nuclei block the export of poly(A)+RNA (Shibata2002) as well as other cargo proteins (Frosst2002), suggesting a vital role of the Tpr in the export process of RNA family as well as proteins. Because the Crm1 (also called Xpo1p or Exportin1) is known as a major exportin protein from the kap family (Hakata2003), it was naturally a sought-after candidate for interaction with Tpr. In an experimental setup with solid phase binding assays, it was found that Tpr favorably bound to the Crm1 in an export complex (Ben-Efraim2009).

To clarify the molecular details of this binding process, we carried out all-atom MD simulations of Crm1 with chopped domains of the unfolded regions of C-terminal Tpr, C-Tpr. To make sure that in dividing the C-Tpr we won't lose any functional domain, the segmentations were chosen to have 10 residues overlap between the consecutive fragments at each end. The C-Tpr then was divided into a total of 33 fragments, and subsequently, those segmentations were arranged randomly around the Crm1-RanGT-Snurportin export complex (Fig. 2.17). Two independent simulations were carried out, each of which was 200 ns long. To maximize the orthogonality, the initial random arrangements of C-Tpr segments around the export complex were different in two simulations (Zhao2014).

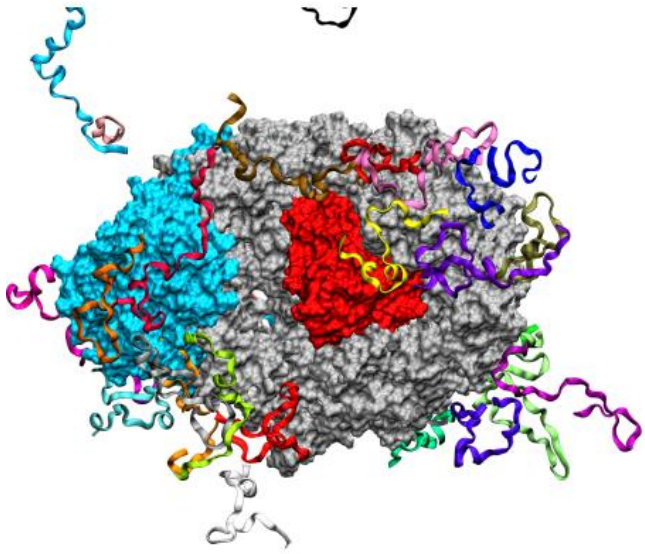


Figure 2.17. The initial random arrangement of chopped C-Tpr around the Crm1-RanGTP-Snurportin export complex. C-Tpr segments are shown by ribbons with assorted colors, Crm1 in gray, RanGTP in red and Snurportin in cyan. Adopted from (Zhao2014).

In our MD study, we identified 11 favorable binding sites on the export complex surface (Fig. 2.18), 9 of which were exclusively on the Crm1 surface, while two were found to be on the Snurportin (Zhao2014). To provide insight into the importance of binding regions predicted by the MD simulations, we also performed conservation analysis on different binding regions. The results revealed that the most conserved region located on the Crm1 surface and served as a binding site for a cargo NES. We also identified three binding regions on the concave surface of the Crm1 for RanGTP. While the binding regions on the Crm1 surface have generally similar conservational

levels as other portions of this exportin protein, most critical residues within the binding regions show much stronger conservation than the other domains in Crm1 (Zhao2014). This implies the importance of specific residues in transient proteins binding, which can be explored to a very high accuracy and atomistic details with MD simulations—an achievement that is not easily accessible to experimental approaches.

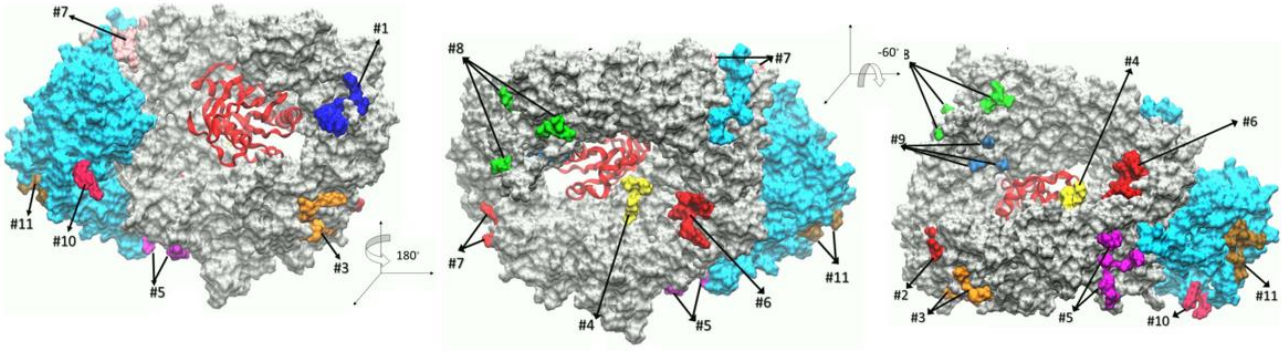


Figure 2.18. Identified binding spots to the C-Tpr on the surface of the export complex through two independent 200 ns -long all-atom MD simulations. Adopted from (Zhao2014).

Another independent set of all-atom MD simulations was performed to investigate the stability and conformational dynamics of Cse1p, an exportin protein, at atomic scale (Zachariae2006). While it was already known based on studying the x-ray structures that Cse1p had an open conformation in the cargo-bound form, Kap60p (Importin-α in yeast)-RanGTP-Cse1p complex, and a closed conformation in the free form (Matsuura2004, Cook2005), it was not achievable for those experimental approaches to find the dynamics of this conformational change. The MD simulations, however, could shed light on the details of conformational dynamics change and reveal that Cse1p had a strikingly fast transition from open conformation to the close conformation, which occurred within 10 ns (Zachariae2006). Two specific domains in the HEAT repeats were identified to be responsible for this rapid, switch-like conformational change: the flexible structure of the N-terminal HEAT repeats 1-3 and the electrostatic properties of the middle HEAT repeats 14-17 (Fig. 2.19).

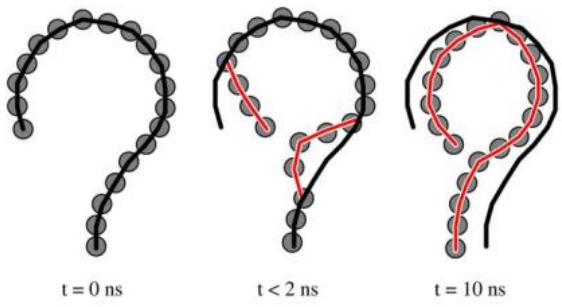


Figure 2.19. The schematic representation of conformational dynamics in Cse1p from an open conformation to a closed conformation in 10 ns. The flexibility of three HEAT repeats in the N-terminal along with the electrostatic properties of HEAT repeats 14-17 were found responsible for this fast-conformational change. Adopted from (Zachariae2006).

The first, and, at the time the most comprehensive, MD study specifically on FG-Nups was performed systematically on subdomains of 100 residue-long of Nsp1 end-tethered on a planar surface in a 5x5 array (Miao2009b). In this study, authors combined all-atom MD with the coarse-grained MD to reach a longer simulation time of 4 μs in CGMD and about 50 ns with all-atom MD.

Simulations results revealed that chopped FG-repeats collectively form a dynamic brush structure, composed of several bundles of FG-repeats (Miao2009b). However, a single chopped FG-repeat shows an entirely different globular conformation (Miao2009b), which is indeed expected from polymer physics viewpoint. A following study conducted by the same group on chopped FG-repeats of Nsp1, but this end-tethered on a torus-shape ring (Gamini2014). The authors got similar results as that in planar surface (Gamini2014), presumably because of the same vertical arrangement of chopped FG-repeats in both setups (see Fig. 2.20).

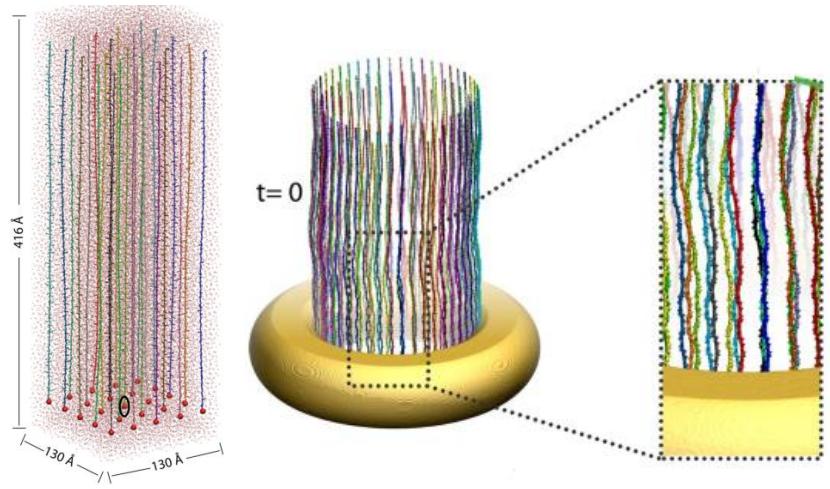


Figure 2.20 CGMD simulations of chopped segments of Nsp1 end-tethered on planar surface (left panel) (Miao2009b) and a torus-like ring (right panel) (Gamini2014). Results from these two setups are qualitatively in agreement with each other. Adopted from (Miao2009b).

Overall, computational biophysical modeling approach is an exceedingly powerful framework to study the microdynamical details of the NPC and the role-playing factors in transport. In particular, mutation studies can be conducted to find, with atomistic details, the importance of specific residues in certain binding processes and conformational dynamics.

A more promising approach emerges when the MD modeling is combined with supporting experimental studies. Indeed, this approach was conducted successfully in a deep analysis of FG-Nups conformation (Yamada2010) and also to elucidate the functional domains of FG-Nups in the permeability barrier formation (Popken2015a). Another combinative area is to marry the MD

modeling with bioinformatics-based analysis, which has emerged more recently (Ando2014, Peyro2015a).

An unexplored area, however, is the combination of molecular mechanics modeling with probabilistic approaches and genetic analysis to reveal a multi-scale view of how, mechanistically, the NPC and its components play roles in the gene regulation. MD simulations can perhaps provide the best framework to explore the atomistic details of plausible interactions between specific domains of Nups with different genetic factors at the nuclear periphery. On the other side of the spectrum are probabilistic approaches, which can portray the big picture of the NPC capacity in gene regulation, as long as the model is fed by reliable input from a robust genetic analysis. To the best of my knowledge, there is no study to date with this multiscale approach.

Chapter 3. Brownian Dynamic Modeling

3.1 Introduction: Brownian motion

The focus of this chapter is to discuss the details of coarse-grained Brownian dynamics and how it has been applied, in combination with the known biophysical assumptions about the NPC, to develop a comprehensive model for the NPC. To so, I begin with a brief description on the foundation of Brownian motion and fluctuation dissipation theorem.

The ubiquity of random motions of microscopic particles in liquids (Stachel2005) was first investigated by the botanist Robert Brown (Brown1828), which, in fact, was come into his attention after his voyage to Australia (Nelson1967). Later on, Fick used a phenomenological approach to address the diffusion of liquids (Klafter2005), and derived the famous diffusion equation for the transport of water and ions across the membrane (Fick1855), i.e. $\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$.

To further investigate the motion of Brownian particles, Einstein continued the phenomenological approach (Doi1988), but pursued this problem in more details by using the molecular theory of heat. He proved that "the microscopic particles suspended in liquids must already carry out an observable random movement, which is generated by thermal motion" (Stachel2005). From the first principles, Einstein derived the diffusion coefficient of small spherical particles undergoing random motion in suspension, and showed that at any temperature it depends solely on the liquid viscosity and the particle size (Einstein1905). He derived his famous equation of:

$$D = \frac{RT}{N}\frac{1}{6\pi a\eta}, \tag{3.1}$$

with R being the universal gas constant, a the radius of the particle, and η the viscosity of the medium. Surprisingly, the same year Sutherland independently derived similar, but more general, equation for diffusion coefficient of molecules under the assumption presence of slip between the diffusing molecules and the solution (Sutherland1905):

$$D = \frac{RT}{N}\frac{1}{6\pi a\eta}\frac{1+3\eta/\beta\alpha}{1+2\eta/\beta\alpha}, \tag{3.2}$$

Where β is the sliding friction coefficient if any slip exists between the diffusing molecules and the solution. This can happen in the extreme case of small diffusing molecules among large solvent molecules. However, in the case of no slipping of solution at the molecular surface, when $\beta = \infty$ in the Sutherland's definition, the Einstein's equation is recovered.

The following year, Smoluchowski profoundly attributed diffusive motion to the random walk, and used a combinatoric approach in analyzing the Brownian motion and obtained similar results (Von Smoluchowski1906). Interestingly, Einstein acknowledged that the approach used by Smoluchowski in Brownian motion analysis was better than that of himself (Stachel1990).

Using the Stokes law for the drag coefficient of a spherical particle, $\zeta = 6\pi a\eta$, the so-called Stokes-Einstein relation is reached:

$$D = \frac{k_B T}{\zeta} \tag{3.3}$$

This is perhaps one of the most elegant equations in physics, which relates the thermally-driven microscopic motion (diffusion) to the macroscopic external force (friction). It simply states that for a given particle these two quantities are interrelated solely through the temperature.

Following Smoluchowski, using a pure statistical analysis of a free particle's random walk, say in 1D, one can show that, in the absence of elasticity, the mean square displacement of the particle increases linearly in time, i.e. $\langle x^2 \rangle \propto t$. As this is a 1D free diffusion, i.e. a 1D Brownian motion, the proportionally constant, λ, must be related to the diffusion coefficient of the particle. It can be shown that $\lambda = 2D$ and thus $\langle x^2 \rangle = 2Dt$ (Berg1983). This result is easily extendable to 2D and 3D, as the free diffusion is fundamentally independent in each dimension. Thus, the dynamics of the particle's motion is the same in each dimension and can be summed over:

$$\langle r^2 \rangle = 6Dt\,, \tag{3.4}$$

which is the general form of the particle's free diffusion in 3D.

Mathematically speaking, the laws of Brownian motion result from the central limit theorem (Bouchaud1990). This theorem states that the probability distribution of the sum of any set of random variables —"subject to certain very general conditions" (Cramer1950)—approaches the Gaussian distribution for a large number of variables (Reichl2016). The constraint "certain very general conditions" only requires that: the random variables must be statistically independent, with finite means and variance (Lemons2002). (These simple conditions can probably explain why the physical world is dominated by the *normal* distribution, rather than any other distribution!).

Much more details on the extension of relativity to the time and space in the Brownian motion of macromolecules in liquid can be found elsewhere (Mezzasalma2008), which are beyond the scope of the current dissertation.

3.2 Brownian dynamic simulation: A coarse-grained approach

In a many-body system, every single configuration (= a set of 3N coordinates for an N-particle system) at each timestep is a 'point' in the phase space. By the end of the computational biophysical simulation, there are recorded statistically sufficient number of 'points' in the phase space, through which the statistical analysis, based on statistical mechanical principles, is performed to find the desired propertied of the system. The foundation of this approach is based on the assumption that the Hamiltonian of the system is ergodic. The validity of ergodic theory for a given system assures that the time average of any classical property in the system would be the same as that in ensemble average (Pathria). Thus, fundamentally speaking, computational

biophysical modeling can mimic the reality and predict reliable results by analyzing the time-evolved trajectory of a system, at least for classical treatment of the system. To this end, however, the *model* must be physically accurate and the *simulation* must be long enough to confidently explore the phase space to be statistically reliable.

Perhaps, the best classic example of this approach is the molecular dynamics (MD) framework, in which the system of interest in considered at atomistic resolution. Depending on the system of interest, the MD simulation can be conducted in different statistical ensembles, i.e. NVT, NPT, etc.

In a very broad sense, the core of an MD simulation is to solve, iteratively in time, the Newton's equations of motion simultaneously for a many-particle system. Thus, the following equation is solved iteratively in time for every single particle i in the system:

$$m_i \frac{d\dot{\mathbf{r}}_i(t)}{dt} = \mathbf{F}_i(\mathbf{r}_i, t), \tag{3.5}$$

where $\mathbf{F_i}$ is the vector sum of all conservative forces on particle i and comes from the force-field:

$$\mathbf{F}_i(\mathbf{r}_i, t) = -\frac{\partial \Phi(\mathbf{r}_i, t)}{\partial r_i}, \tag{3.6}$$

with $\Phi(\mathbf{r}_i, t)$ being the potential energy.

In practice, there are several finite difference algorithms that are used to solve the RHS of this equation (described in detail in (Allen1987)). The general idea is that given the particles' positions, velocities, and other dynamical properties at time t, we look to calculate those properties at time $t + \delta t$.

The heart of an MD simulation in particular, and every computational biophysical modeling in general, is the force-filed (LHS of eq. 3.5): how, *quantitatively*, constituting particles interact with each other. At atomic scale, developing a molecular force-field as well as validating the force-field are deeply rooted in quantum mechanical analysis (Wang2001, Hornak2006). For coarse-grained modeling, however, the force-filed becomes, to some extent, system-dependent, as is the case in the current work. As there are several excellent references for the MD simulation approach with deep statistical mechanical along with mathematical details, here I don't go to the details of MD. Perhaps one of the best references which serves as a canon of MD simulations and still is widely used by the MD developers is the book by Allen and Tidesley (Allen1987). It covers the foundations of equilibrium as well as nonequilibrium MD simulations along with many introductory and advanced simulation techniques that have been elaborated in an algorithmic and practical way. There are indeed other strong reference books, covering from beginners (Haile1992, Leach2001) to more advanced topics in MD (Frenkel2001, Berendsen2007).

To increase the computational efficacy and the ability to simulate larger systems, the MD can be made coarse-grained (CG) by averaging over some of the fine details through new paradigms. One of the main paradigm is to encapsulate several atoms into a new 'atom' and introduce new force-

fields for participle-particle, particle-solvent, and solvent-solvent interactions. Perhaps the first CGMD was conducted by Levitt and Warshel (Levitt1975) in an effort to model the folding of a small protein, on the basis of two main assumptions: i) fine details of protein structure can be replaced by a single average variable, and ii) to obtain an overall folding, it is sufficient to consider only most effective variables in protein folding (Levitt1975).

The process of coarse graining in the MD was continued in more depth by others and new force-fields specific to coarse-grained systems were developed. One of the successful force fields in CGMD is the MARTINI which was initially developed and parametrized for simulating lipids and surfactants (Marrink2004, Marrink2007). It subsequently extended to simulate a broad range of biomolecules (Monticelli2008). This approach is widely utilized by GROMACS (Hess2008). Yet, the equations of motions are still Newtonian equations.

The process of coarse-graining can be further continued to achieve larger domains both in time and space, indeed at the price of losing atomistic-level information. While there is no clear boundary of the definition of different levels of coarse-graining, there is a general guideline mainly based on the equations of motions that are solved in the simulation.

Following this guideline, the next level of coarse-graining is Langevin dynamics, where the solvent molecules are set as a 'bath' in the background and their detailed dynamics is neglected (Snook2006). Instead, the effects of the solvent on the large Brownian particle suspended in the solvent are taken into account by introducing a white noise, i.e. a random force, and a dissipative frictional force (Schlick2010). In his succinct paper, Paul Langevin (1908) was the first that devised this approach for investigating the Brownian motion of spherical particles in the viscous fluid. He applied the Newton's second law to a Brownian particle and proposed the following stochastic equation of motion (Lemons1997):

$$\mathbf{F}_i(\mathbf{r}_i, t) + \mathbf{R}_i(t) = m_i \frac{d\dot{\mathbf{r}}_i(t)}{dt} + \zeta_i \dot{\mathbf{r}}_i(t), \qquad (3.7)$$

where $\mathbf{R}_i(t)$ is the random force on particle i and ζ_i is its drag coefficient, which depends on the particle's size and the solvent viscosity η.

This equation is generally referred to as the Langevin equation and lays the foundation of what is now called the Langevin dynamics. Equation (3.7) has been solved by stochastic integration and the consequently numerical equations relating the positions of particle at consecutive timesteps have been derived. This approach would introduce a significant computational boost and can simulate systems that are several orders of magnitude larger than the capability of all-atom MD simulation (Scott1999).

In Langevin dynamics, the solvent exists 'implicitly' through random and frictional forces on Brownian particles. The dissipative frictional force manifests itself linearly dependent on the particle's velocity $\dot{\mathbf{r}}_i(t)$, with the particle's drag coefficient, ζ_i, being the proportionality constant. The random force $\mathbf{R}_i(t)$ which represents the molecular collisions of the solvent, however, does not have any analytical relationship; instead, it can be described by the following statistical properties (Kubo1966, Eastman2015):

0. The nature of random force is independent of the presence of the force field.
1. It has a mean value of zero: $\langle \mathbf{R}_i(t) \rangle = 0$.
2. It does not depend on the Brownian particle's location, $\mathbf{r}_i$.
3. It varies so chaotically, and thus is uncorrelated with itself over times larger than a characteristic time τ: $\langle \mathbf{R}_i(t)\mathbf{R}_i(t+T) \rangle = 0$, $T \gg \tau$

The characteristic time τ is the time beyond which a random force is uncorrelated with itself. In effect, τ is the time over which frictional force operates, $\tau = \frac{m}{\zeta}$ (Eastman2015). Analogous argument about neglecting the inertial effects of the probe applies when one is deriving the microrheology equations in the frequency domain (Mason2000). It will be mentioned it in the next chapter.

Kubo (1966) argued that because the random force and the frictional force both arise from the impacts of surrounding small molecules on the Brownian particle, they have the same origin, and thus, they must be related. In fact, he compellingly argued that molecular impacts have two different types of effects on the suspended Brownian particle. The first effect comes from random forces, manifested as the nonstop irregular motion of the Brownian particle. The second effect is the *systematic* part, manifested as the resistance, i.e. friction, against the forced motion of the Brownian particle (Kubo1966). For a system in contact with a heat bath the dissipative friction and random fluctuations are fundamentally linked together; you cannot have one without the other (Eastman2015). The internal relationship between thermal random force and frictional force is elegantly illustrated in the so-called *fluctuation-dissipation theorem (FDT)*. This theorem represents that spontaneous thermally-driven fluctuations are in essence related to the characteristics of the system's response to an external field (Doi1988). As such, the Einstein's relation is a special case of the FDT. The theorem can be stated as follows (Kubo1966, Schlick2010):

$$\langle \mathbf{R}(t)\mathbf{R}(t') \rangle = 2k_B T \zeta \delta(t - t') \qquad (3.8)$$

This relation shows how the particle's drag coefficient ζ, and thus the solvent viscosity η, control the autocorrelation of random force, i.e. for how long a random force remembers its history. In other words, not only the solvent viscosity modulates the magnitude of the *frictional* force, but it also controls how *random* force correlates with itself.

At very short time scales, i.e. $t << \tau$, the suspended particle basically does not yet feel the frictional force from the solvent random molecular collisions, and thus, it behaves as a *ballistic* particle rather than a Brownian particle. In this regime, the movement of the particle is deterministic and the classical Newtonian equations of motion (eq. 3.5) are used to obtain the time evolution of the system. Within this domain the molecular dynamics methodology, either all-atom MD or CGMD, is utilized where the atomistic details of the system are captured at very short timescales. At longer time scales $t >> \tau$, however, the random force becomes uncorrelated with itself and the particle loses its memory, and thus, show pure stochastic behaviors. The particle's motion is rapidly damped by the solvent molecules. Different names have been coined for this regime, i.e., *diffusive*, *Brownian*, or *overdamped* regime (Leach2001, Fall2002, Snook2006). In

the diffusive regime, the inertial term in eq. (3.7) is negligible compared to viscous force, and can be dropped, leading to the following equation of motion for a particle i:

$$\mathbf{F}_i(\mathbf{r}_i,t)+\mathbf{R}_i(t)=\zeta_i\dot{\mathbf{r}}_i(t) \qquad (3.9)$$

This is the next level of coarse-graining, and is appropriate to model the macromolecules swimming in the intracellular viscous environment under a very small Reynolds number (Purcell1977). This equation is the basis of what is called the Brownian dynamic approach, which is in the heart of my work on modeling the NPC in the current dissertation.

In fact, for particles suspended in liquids, the solvent viscosity determines dynamical regime of the particle's motion as manifested by the Reynolds number—the ratio of the inertial force to the viscous force acting on the particle, $\mathrm{Re}=\frac{\rho vL}{\eta}$. Because the drag coefficient is directly related to the solvent viscosity, for a given suspended particle the characteristic time τ is solely determined by the solvent viscosity, i.e. $\tau=\frac{m}{\zeta}\propto\frac{1}{\eta}$. In the limit of small viscosity (nonadhesive medium), the characteristic time of a suspended particle gets large enough so that practically for any observation time t the motion of the particle can be considered ballistic. The opposite is true for very viscous medium, where τ is so small that within the time-window of interest, the particle's motion is primarily governed by the friction and random collisions, rather than the particle's momentum.

For example, in the current dissertation, the mass of a monomer is about 0.5 kDa ($\approx 0.5 \times 10^{-21}$ gr). Considering an average protein density of $1.35\ \frac{gr}{cm^3}$ (Fischer2004), an equivalent hydration diameter of about 1.2 nm is obtained for a single bead suspended in cytoplasmic viscosity, with a drag coefficient of about $5.7\times10^{-8}\ \frac{pN.s}{nm}$. Therefore, the characteristic time would be $\tau\approx10^{-14}s$. This time is so small compared to the time-window of interest where the protein global conformation occurs, and is orders of magnitude smaller than the diffusive motion of proteins across the NPC. This justifies the use of Brownian dynamics in modeling the NPC and any other macromolecular assembly or motor proteins.

Taking integration of both sides of eq. (3.5) in only x-direction, we would have:

$$\int_t^{t+\Delta t}F_i^x(x,t')dt'+\int_t^{t+\Delta t}R_i(t')dt'=\zeta_i\left[x_i(t+\Delta t)-x_i(t)\right]$$

or (3.10)

$$x_i(t+\Delta t)\approx x_i(t)+\frac{1}{\zeta_i}F_i^x(x,t)\Delta t+\frac{1}{\zeta_i}\int_t^{t+\Delta t}R_i(t')dt'$$

The last term on the RHS of the latter equation requires stochastic integration of the random force, which would lead to a random displacement during the timestep Δt. From the central limit theorem, compellingly, the random force, and thus the resulting random displacement, should

follow a normal distribution. Moreover, the physics of the pure diffusive motion dictates that $\langle R_i(t)\rangle = 0$ and $\langle x_i(t)\rangle = 0$. Without losing the generality, we start from $(x,t) = (0,0)$ and drop subscript i. In the absence of any external force, it follows from (3.9) that a pure random displacement is:

$$x(t) = \frac{1}{\zeta}\int_0^t R(t')dt', \tag{3.11}$$

which indeed satisfies $\langle x(t)\rangle = 0$.

From (3.7) we can deduce the covariance of the pure random displacement as $(t > s)$:

$$\mathrm{cov}[x(t)x(s)] = \frac{1}{\zeta^2}\left\langle \left(\int_0^t R(t')dt'\right)\left(\int_0^s R(t'')dt''\right)\right\rangle \tag{3.12}$$

We now use the FDT (eq. (3.8)) to constitute the random force in (3.12). Before that, however, we have to consider a property of the Dirac delta function (in eq. (3.8)). Dirac delta function is a mathematical concept with interesting properties (Ghatak2004) that can be considered as a normal distribution in the limit of zero variance (Fall2002):

$$\delta(x-a) = \lim_{\sigma\to 0}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-a)^2}{2\sigma^2}\right], \quad \sigma > 0 \tag{3.13}$$

One of its properties is that the integration of the product of any given function with Dirac delta renders the function itself:

$$\int_{-\infty}^{\infty} g(t)\delta(t-s)dt = g(s), \tag{3.14}$$

that is true for any function $g(t)$. It then follows from the theory of Dirac delta function that:

$$\begin{aligned}\mathrm{cov}[x(t)x(s)] &= \frac{1}{\zeta^2}\int_0^t\int_0^s \langle R(t')R(t'')dt''dt'\rangle \\ &= \frac{2k_BT}{\zeta}\int_0^t\int_0^s \delta(t'-t'')dt'dt''' = \frac{2k_BT}{\zeta}s = 2Ds\end{aligned}, \tag{3.15}$$

from which we can recover the 1D mean-square displacement in a free diffusion motion:

$$\langle x(t)^2\rangle = 2Dt \tag{3.16}$$

Eq. (3.15) lays the foundation of stochastic integration from the last term on the RHS of eq. (3.10) to find the random displacement due to the Brownian force within the timescale of Δt. It then appears that:

$$x(t+\Delta t) \approx x(t) + \frac{1}{\zeta} F^x(x,t)\Delta t + \left(2D\Delta t\right)^{\frac{1}{2}} Z \,, \qquad (3.17)$$

where Z is a random variable from a Gaussian distribution with the mean value of 0.0 and the variance of 1.0. This guarantees that $\left(2D\Delta t\right)^{\frac{1}{2}} Z$ is a random number that follows a normal distribution with the mean value of 0.0 and the variance of $2D\Delta t$. The random displacement $\left(2D\Delta t\right)^{\frac{1}{2}} Z$ encapsulates two effects on the Brownian particle: i) the dissipative friction force and ii) the molecular random collisions from the solvent.

In the current model, equation (3.17) along with its y- and z-direction counterparts, are the main numerical displacement equations that are used in each timestep to determine the current position of the particle based on the previous position, the force-fields, and the random displacement.

Like an MD simulation, the main and the foremost output of a Brownian dynamic simulation is the time-evolved trajectory of Brownian particles. At each certain timesteps, the positions (x,y, and z in Cartesian coordinate) of Brownian particles are recorded. This trajectory then is analyzed according to statistical mechanical rules to produce different properties of the system.

The numerical derivation of the displacement equation for a system composed of N Brownian particles in the presence of hydrodynamic interactions (HI) was first analyzed in a pioneering work by Ermak and McCammon (Ermak1978). They began from the Langevin equations for a many-body system, and their work is more general than what presented here in that they have also incorporated the effects of HI. Their final displacement equation is similar to the eq. (3.17) above, except that in their equation there is an extra term for the position-dependent interparticle friction that accounts for hydrodynamic interactions.

3.2.1 Dimensionless units

One of the good practices in developing computational biophysical simulations is to reduce the units of parameters to make them dimensionless during the simulation. Although this process is not necessary and does not affect the physics of the system, it makes the numerical simulation smooth and standardized. In SI units, for example, the molecular parameters, including charge energy, force, etc. are extremely small and prone to error accumulation during the run.

In the current work, time, length, and energy units are chosen as three fundamental units: $\varepsilon_u = 1 k_B T$, $\sigma_u = 1 nm$, and $t_u = 10^{-10}\, s$. All other physical units are then expressed with respect to these three fundamental units as following:

Mass: $m_u = \dfrac{\varepsilon_u \times 4.2821\times10^{-21} \times \left(t_u\right)^2}{\left(\sigma_u\right)^2 \times 10^{-18}}\ kg$ (3.18-i)

Temperature: $T_u = \frac{\varepsilon_u}{k_B} \quad K$ (3.18-ii)

Force: $F_u = \frac{m_u \times 1.66053 \times 10^{-21} \times \sigma_u}{\left(t_u\right)^2} \; pN$ (3.18-iii)

Charge: $c_u = \left(\varepsilon_u \times 4.2821 \times 10^{-21} \times \sigma_u \times 10^{-9} \times 4\pi\varepsilon_0\right)^{1/2} C$, (3.18-iv)

in which $\varepsilon_0 = 8.8542 \times 10^{-12} \; Fm^{-1}$ is the vacuum permittivity.

Following the selection of units, the physical parameters are become dimensionless with respect to these units. For example, the dimensionless mass, $m^* = \frac{m}{m_u}$.

3.2.2 Random number generator

> *"Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin"—John von Neumann (1951)*

The final, but the vital, practical point on the displacement equation (3.17) is the choice of the random number Z (technically speaking, *pseudo*random number) which must be drawn from a normal distribution. Generally, cautions should be taken when using a random number generator (RNG) algorithm in different simulations (Hellekalek1998), from the viewpoints of both computational efficiency as well as statistical properties. This becomes a critical point in scientific commutating (Park1988). Equally important, is the random seed, which is the initial number to begin a sequence of random numbers. The details of the random number generator and the random seed are discussed in the Appendix A.

3.3 Development of a coarse-grained model of the NPC

In this section, I will describe the details of the process of developing the NPC coarse-grained model, along with biological, biophysical, and computational considerations that I have regarded during the process. Since this work initially started with developing a 2D model of the *Xenopus* oocyte NPC, and then continued to a more comprehensive 3D model of the yeast NPC central channel, there are lots of overlaps between 2D and 3D models from development point of view. Therefore, I would take a combined approach to describe a comprehensive development process.

In the current work, the force-field development lies under two broad categories of forces, i.e. bonded and nonbonded forces. There are three bonded forces in the model: i) **linear springs** that form the main structure of the NPC, ii) **nonlinear springs** that mimic the polymeric meshwork of FG-repeats, and iii) **angular harmonic springs** that account for bending potential energy in both the main NPC structure and FG-repeats. There are also three nonbonded forces in the model: i) **van der Waals**, ii) **electrostatic**, and ii) **hydrophobic** forces.

While bonded forces form the structure of the NPC, the nonbonded forces account for different types of interactions. A prerequisite for incorporating bonded forces, however, is to discretize the system into appropriate elements, i.e. beads and springs in the current model. Therefore, in the following I will start with introducing the bead-spring model and the rationales for using it in the computational biophysical model of the NPC. It will be followed by the discussion about bonded and nonbonded forces in the context of the current work. The physical and biological foundation of each case will be discussed in details.

3.3.1 Bead-spring model and the rationale behind it in computational biophysical model of the NPC.

Physical modeling and computer simulation of long molecular chains have been of interest since perhaps the early days of scientific computing in this field (Stellman1972, Gottlieb1976, Fixman1978a) to predict various physical quantities of polymeric systems, including different relaxation rates (Fixman1978b, Kranbuehl1979), dynamic (Fixman1978b) and intrinsic viscosity (Yamakawa1974), as well as other conformational and transport properties (Fixman1978a). As a general rule in the physical computational modeling space, one needs to discretize the system of interest into **subsystems** and solve the constitutive equations in each subsystem subject to appropriate boundary conditions. The very original idea of discretizing a polymeric chain into N submolecules and consequently describing the whole chain by N connecting vectors was first introduced by Rouse (Rouse Jr1953). The idea was widely appreciated and extensively developed in coming decades.

From the subsystems' movements continuity viewpoint, there are two general paradigms in modeling: **continuous** and **discrete**. The subsystems can either be anywhere in the 3D space—so long as their movements satisfy the model requirements—or they can only occupy specific locations. The former represents modeling in a continuous space as in MD or Brownian dynamics, while the latter is referred to as discrete space modeling, as in the Monte Carlo and Agent-Based. This is, among others, one of the fundamental differences between MD and MC simulation methods, for example.

For the computer simulation of polymers in the discrete space, perhaps the best-known model is the **lattice model**, which is widely used in the Monte Carlo and Agent Based simulations. Similarly, in the continuous space, there are several geometrical models that are employed to investigate the physical and dynamical properties of polymeric systems. Broadly speaking, three main models are used in the continuous space to represent individual polymers and/or their segments (Teraoka2002b): i) **bead-stick model**, ii) **pearl-necklace model**, and iii) **bead-spring model**, as illustrated in Fig. 3.2.

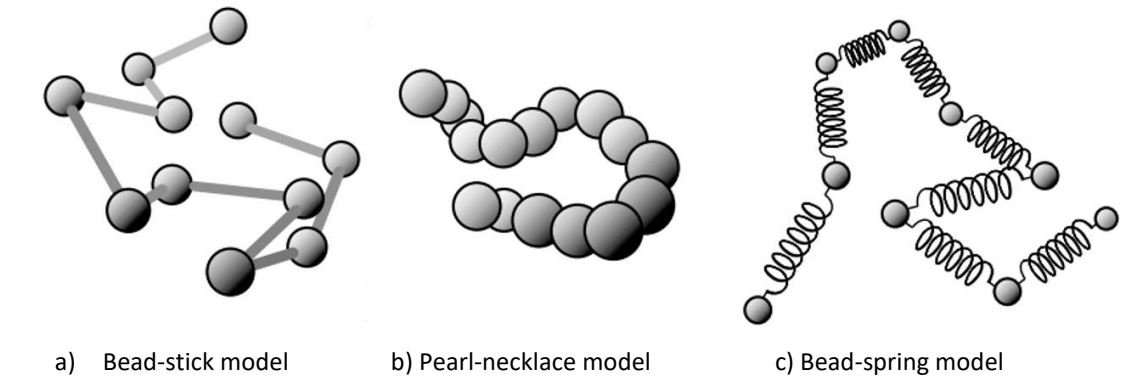
In the bead-stick model, the backbone of a polymer chain is imitated as a series of beads which are connected consecutively by rigid sticks or rods (sometime called **bead-rod model**). Many vibrations are possible. For example, the angle between two adjacent sticks can be fixed, free, or a determined by an angular potential, depending on the level of complexity in the model. The famous **freely jointed chain** is in fact a bead-stick model with no angular potential. Moreover, it is possible to introduce a torsional energy to account for energy-dependent dihedral angle (the angle between a stick and the second neighboring stick).

Pearl-necklace model can be considered an extreme extension of the bead-stick model, in which pearls (=beads) have diameters equal to the connecting sticks. The pearls are always connected,

but similar to bead-stick model we can define an angular potential between two adjacent pearls to account for the backbone's bending rigidity.

The bead-spring model is also similar to the bead-stick model, except that instead of a rigid rod an elastic spring connects two consecutive beads. Similar to the bead-stick model, here also one can define the bending rigidity or torsional potential energies.

From a granularity viewpoint, all of the above models can be tuned to demonstrate atomic or coarse-grained scale, depending on the entity that each bead and/or connector represents. For example, the backbone of a vinyl polymer can be modeled at atomistic scale by a freely jointed chain if the angular bond in the model is fixed to the tetrahedral angle of a carbon atom and the dihedral angle is fixed at corresponding values accordingly (Teraoka2002b).



a) Bead-stick model b) Pearl-necklace model c) Bead-spring model

Figure 3.2. Three main physical models for computer simulation of linear polymers in a continuous space. Adopted from (Teraoka2002b).

Among the above models, the bead-spring approach has the unique feature of the bond's reversible extensibility, i.e. elasticity, which distinguishes it from the other two models. As one of the main components of the nuclear envelope (NE), the proteinous, structured backbone of the NPC is clearly showing **Hookean elasticity**, consistent with the elasticity of the NE (Dahl2004, Rowat2005, Yokokawa2008). Moreover, the unstructured domains of Nups, i.e. FG-repeats, also reveal a strong **nonlinear elastic** behavior (Lim2006, Lim2007a). Both linear and nonlinear elasticity can be accurately modeled by the bead-spring approach. The bead-spring model is thus the pivotal point in this dissertation upon which I have built the discretized model of the FG-repeat domains and the NPC structure.

3.3.2 Different types of springs in the bead-spring model: the grounds for linear and nonlinear elasticity in the NPC modeling

The connecting springs in the bead-spring model respond to the applied force either linearly or nonlinearly. While a linear spring follows a Hookean force-extension, a nonlinear response shows finite extensibility.

To perform an accurate coarse-gained model of the NPC, we need to consider its structural components separately and attribute appropriate springs to them. The NPC structure is composed

of three different groups of Nups (Fig. 3.3) which, structurally, are on the top of each other (Rout2003, Peters2009):

1. Pore membrane proteins (Poms) that anchor the NPC scaffold to the NE and surround the central scaffold.
2. Structural proteins that give the NPC shape and strength and maintain the main scaffold.
3. FG-Nups that provide the binding sites for NTRs and are central to selectivity.

In the current work, the first two groups of Nups are modeled by **linear springs**, while the FG-Nups are modeled by **nonlinear springs**. There are also **angular springs** in structural proteins and FG-Nups to account for bending potential energy (Fig. 3.4).

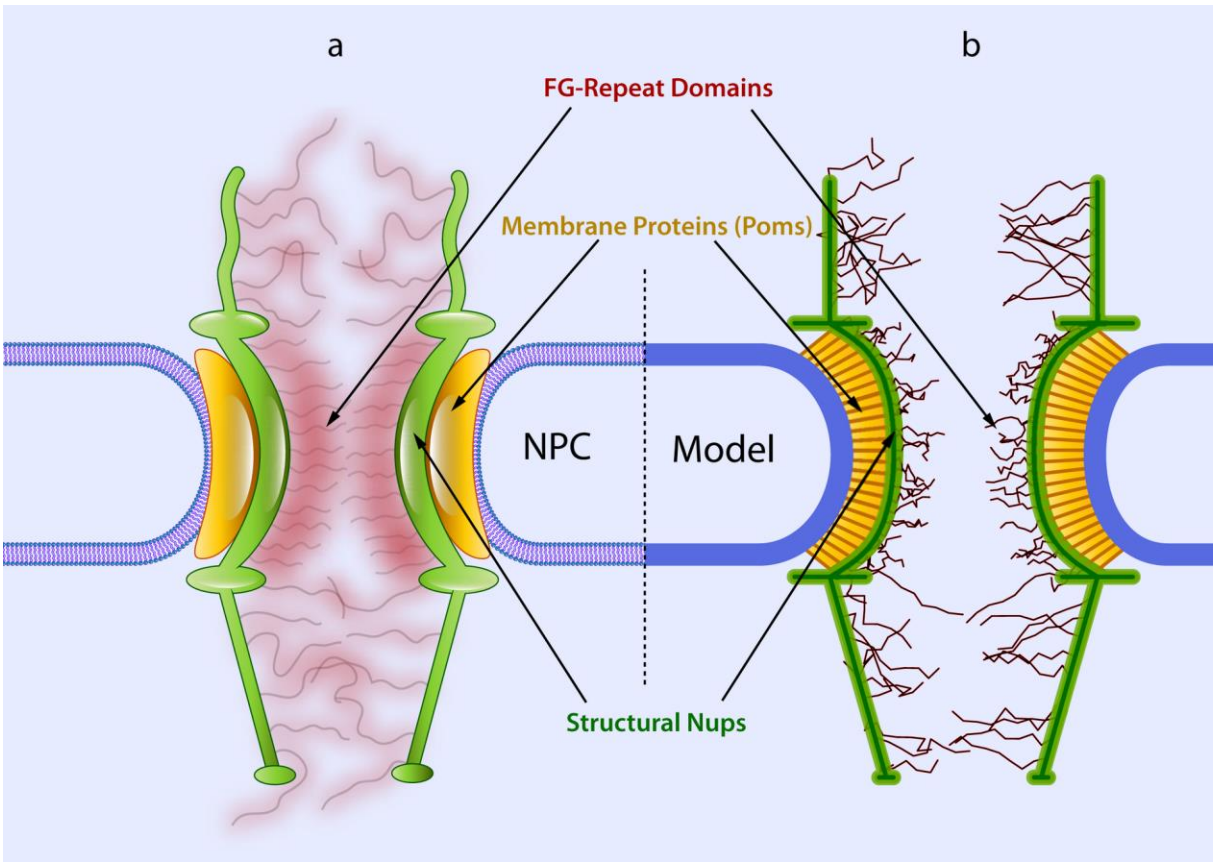


Figure 3.3. The three main layers structural elements of the NPC are Poms proteins, structural Nups, and FG-Nups. The mechanical properties of all of these elements are modeled by bead-spring approach. A) A 2D representation of the NPC, and B) The 2D coarse-grained model of the NPC with equivalent structural elements.

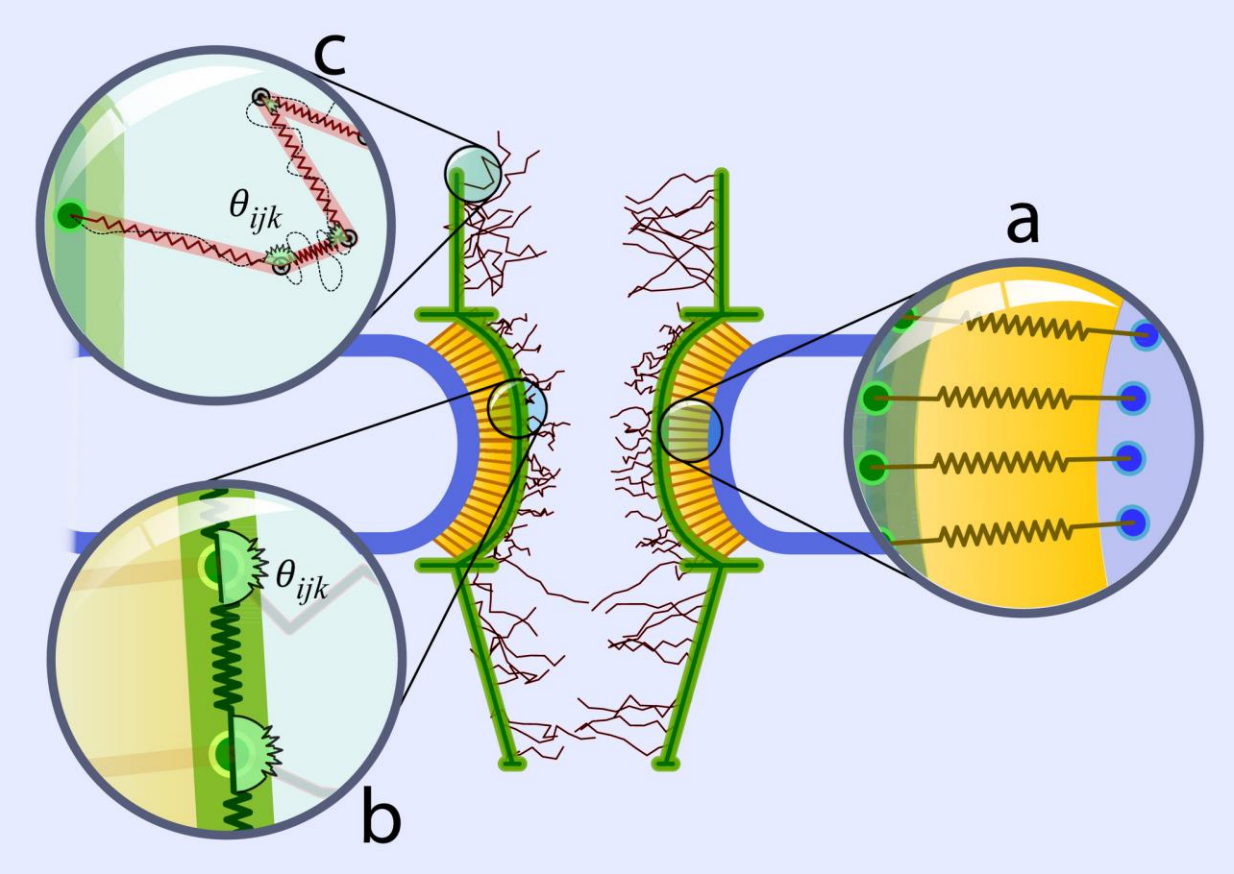


Figure 3.4. The three main structural elements of the NPC in the current work are modeled through bead-spring approach. There are three type of springs: linear, nonlinear, and angular. A) The Poms proteins are modeled merely by parallel linear springs. B) The structural Nups are modeled as a series of linear springs along with angular springs. C) FG-Nups are modeled by WLC along with angular springs. Adopted from (Moussavi-Baygi2011b).

3.3.2.1 Linear elasticity

Linear elasticity has a long history in modeling and simulation of biological materials and is frequently used to predict their mechanical properties under sustainable—small and, subject to some conditions, large—deformations with an overall structural stability (Chook2001, Redondo2004, Vaziri2007, Boal2012). This makes it an excellent choice for modeling the backbone of the NPC that acts as a channel wall, ornamented with cytoplasmic filaments and the nuclear basket, showing deformations and structural stability during the transport (Akey1995, Griffis2004, Bui2013, Sellés2017).

In the current dissertation, Poms are modeled as a set of parallel Hookean springs that on one end are fixed to the NE and on the other end are connected to the central channel, and therefore, anchor the NPC to the NE (Fig. 3.4 A). This accurately represents the function of the Poms in terms of anchoring the NPC to the NE. As such, beads in the Poms proteins are not assigned any diffusion coefficient, since they are stationary and their function is only to anchor the main structure to the NE.

Similarly, The NPC main scaffold, consisting of the cytoplasmic filaments, central channel, and nuclear basket, is also assumed elastic (Wolf2008)and in this model is mimicked by bead-spring elements (Figure 3-4 B). Each bead in the NPC backbone represents a lumped mass which is determined by the discretization scheme as following. The total NPC mass is proportionally distributed in each of these architectural elements according to the cry-electron tomography of the NPC conducted by Stoffler et. al. (Stoffler2003) as:

- cytoplasmic filaments $\sim$ 3.6%,
- cytoplasmic ring $\sim$ 25.6%
- central channel $\sim$ 44%
- nuclear ring $\sim$ 16.8%
- basket filaments $\sim$ 3.6%
- distal ring $\sim$ 6.4%

The corresponding mass of each architectural element is then distributed equally among beads in that element. Subsequently, the radius of each bead is calculated according to the protein's average density of $1.35\ \frac{gr}{cm^3}$ (Fischer2004) as:

$$r_b = \left(\frac{3}{4\pi\rho_{prot}} m_b \right)^{\frac{1}{3}}, \qquad (3.19)$$

From the bead radius we can calcuate drag and diffusion coefficients according to the Stokes and Stokes-Einstein equations:

$$\zeta_b = 6\pi r_b \eta_{cyt} \quad \& \quad D_b = \frac{k_B T}{\zeta_b} \qquad (3.20)$$

Here, $\eta_{cyt} = 5\ cP$ is considered the mean value for cytoplasmic viscosity as measured by independent studies (Swaminathan1997, Gorlich2003).

While there is no rotational diffusion is considered for the cargo in this work, the same equation as 3.20 is used to calculate the translational diffusion coefficient for globular cargo. For the elongated cargo, however, the translational diffusion coefficient has perpendicular and parallel components (Tirado1979):

$$D_{cyl} = (1/3)\left(2D_{cyl}^{\perp} + D_{cyl}^{\parallel}\right), \qquad (3.21)$$

where:

$$D_{cyl}^{\parallel} = \left(k_B T / 2\pi\eta h\right)\left(\ln p - 0.207 + 0.980\, p^{-1} - 0.133\, p^{-2}\right) \qquad (3.22)$$

and

$$D_{cyl}^{\perp} = \left(k_B T / 4\pi\eta h\right)\left(\ln p + 0.839 + 0.185\, p^{-1} + 0.233\, p^{-2}\right) \qquad (3.23)$$

with $p = h_{cyl} / 2R_{cyl}$.

This formulation has been successfully applied to elongated macromolecules in the context of Brownian dynamics (Klenin1998).

Then an elastic bond is defined between two consecutive beads i and j as a Hookean spring with the corresponding spring constant k_{ij} .

Under the assumption of isotropy in the material, one can easily relate the Hookean spring constant, k, to the Young modulus E as (Fig. 3.5):

$$E = \frac{\sigma_{xx}}{u_{xx}} = \frac{F/A}{x/L} \quad \Rightarrow \quad F = \frac{EA}{L}x = kx \tag{3.24}$$

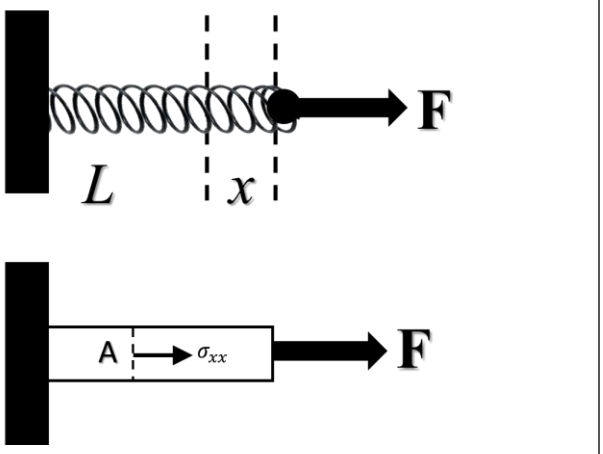


Figure 3.5. The linear spring constant is directly related to the Young's modulus in homogenous and isotropic material.

Therefore, to calculate the stiffness of each spring, three parameters are required: i) the initial length of the spring that is determined by the discretization scheme, ii) the Young modulus, and iii) the cross-sectional area of the element.

Because there is no specific data available on the elasticity of the Poms proteins or structural Nups, the elastic modulus of the NE, 17 kPa, as measured by Yokokawa (Yokokawa2008) is used to calculate the spring constant for Poms and the main structure. This seems reasonable because the NE is usually assumed uniform when its material properties are of interest (Guilak2000, Vaziri2007, Yokokawa2008). However, this value is much larger than the 1-2 kPa elasticity of Nsp1 FG-hydrogel obtained by AFM (Frey2006a), implying that the NE and hence the anchoring and structural Nups are far stiffer than FG-Nups.

The cross-sectional area of the relevant element is also required in eq. (3.24) to calculate the spring constants for that element; however, no such detailed data was available for different elements of the NPC. Therefore, I chose the most accurate cryoelectron-microscopy images available to calcuate the cross-sectional area of different elements of the NPC. In doing so, four criteria for the NPC image accuracy were i) acknowledgement by the NPC field, ii) number of citations, iii) construction of the succeeding NPC models based on the image, and iv) agreement with subsequent studies on the NPC.

Via this analysis, I found the following as the **diameters** for different elements of the NPC (the details are depicted in Fig. 3.6 with yellow dashed lines and arrows):

- cytoplasmic filaments ~ 25 nm (Kiseleva1998)
- cytoplasmic ring ~ 25 nm (Beck2004)

- central channel (spoke domain) ~ 36 nm (Yang1998)
- nuclear ring ~ 25 nm (Kiseleva2004)
- basket filaments ~ 15 nm (Kiseleva1996)
- distal ring ~ 15 nm (Beck2004)

Surprisingly, there is a strong balance in the elements' diameters. For example, cytoplasmic and nuclear rings have the same diameters. Spoke domain is the thickest element, contributing in the structural stiffness of the central channel, consistent with its functionality in the transport. Moreover, I found that the nuclear-face filaments, i.e., basket filaments and distal ring were the thinnest, and thus most flexible, elements. This becomes most remarkable in light of recent growing body of evidences on the role of nuclear basket in chromatic regulation, spindle assembly checkpoint, and mRNA biogenesis, which call for a dynamic and flexible structure of the basket (Gallardo2017).

The linear elastic force between two adjacent beads i and j at time t is then calculate by:

$$\mathbf{F}_{ij} = k_{ij}\left(\mathbf{r}_{ij} - \mathbf{r}_{ij}^{0}\right), \tag{3.25}$$

which is in fact the derivative of the harmonic elastic potential energy (Fig. 3.6 A):

$$U_{ij} = -\frac{1}{2}k_{ij}\left(\mathbf{r}_{ij} - \mathbf{r}_{ij}^{0}\right)^{2}, \tag{3.26}$$

Here $\mathbf{r}_{ij}^{0}$ refers to the bond's rest length.

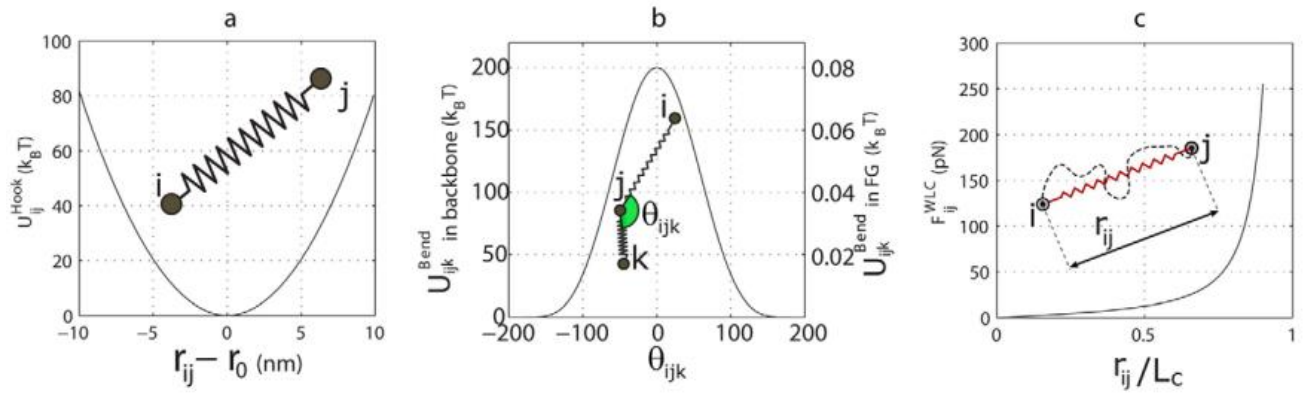


Figure 3.6. The energy and force-law of the bonded interactions in the NPC structural component. **A**) Extensional elastic potential energy between two consecutive beads in the Pom or main scaffold regions, representing the elastic extension. **B**) Angular potential energy between two consecutive springs representing bending rigidity of the main scaffold and the FG-repeat domains. Y-axis. on the left shows values of the angular potential energy for the main scaffold, while the right y-axis shows those values for FG-repeat domains. **C**) The force-law of the wormlike chain (WLC). Discrete WLC models the FG-repeat domains.

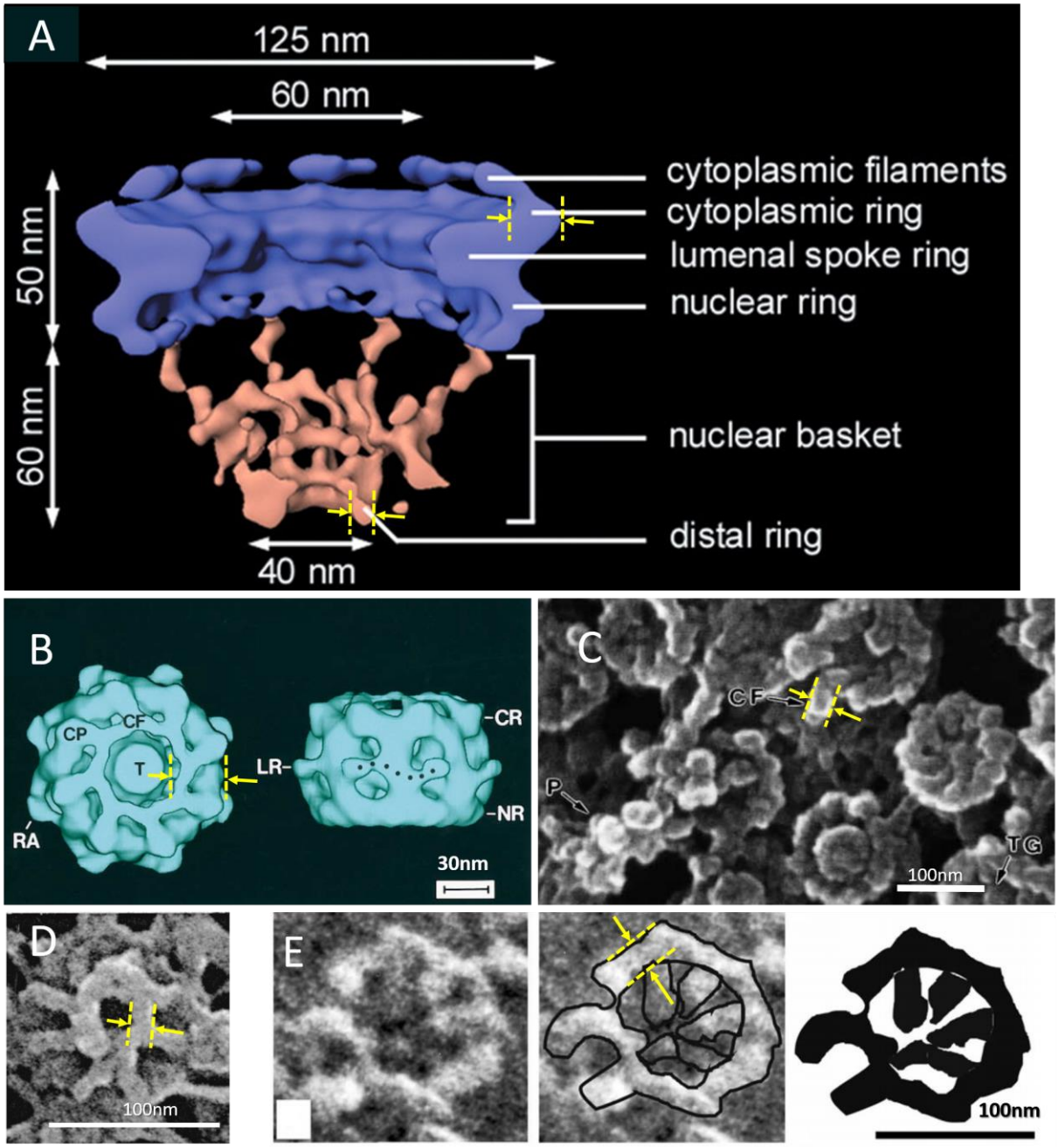


Figure 3.7. The cross-sectional area of all elements of the NPC were calculated from the available images of different sections of the NPC. In all figures, the yellow dashed line and arrows show the thickness of the corresponding elements in the NPC. The thickness calculated by comparing with the scalebar. **A**) The overall dimensions of the NPC from a cryoelectron tomography image. I used this figure to calculate diameters of the cytoplasmic and distal rings as 25 nm and 15 nm, respectively. Adopted from (Beck2004). **B**) The thickness of spoke domain is calculated from an early pioneering study on the 3D architectural analysis to be 36 nm. Adopted from (Yang1998). **C**) The diameter of the cytoplasmic filaments is calculated to be 25 nm. Adopted from (Kiseleva1998). **D**) The diameter of the basket filaments were calculated to be 15 nm. Adopted from (Kiseleva1996). **E**) The diameter of the nuclear ring was calculated to be 25 nm. Adopted from (Kiseleva2004).

3.3.2.2 Nonlinear elasticity

The linear spring represents an infinite extensibility, which is the source of its failure in modeling non-Newtonian fluids, like polymeric solutions (Bird1987). A nonlinear spring, on the other hand, shows finite extensibility, which arises from extremely high entropic cost when the chain' end-to-end length approaches its fully-extended length conformation.

It is well established that the finite-extensibility is an important characteristic in shaping the polymers dynamics and rheological properties (Wiest1989, Van den Brule1993, Van den Brule1995, Underhill2004b), as observed in literally every flexible and semiflexible (see below for their definitions) polymer (Berg1979, Van den Brule1993, Doyle1997, Klenin1998, Underhill2004b, Kneller2005). The absence of this characteristics in a simulation would leave results unphysical and somewhat misleading, so long as the nonlinearity is intrinsically part of the system (Prabhakar2004a). As a crucial mechanical factor, the backbone's entropic elasticity has also been shown to exists in various bio-macromolecular chains that are made up of amino-acid-based elements, like proteins, or nucleotide-based elements, like DNA and RNA, or the combination of both, like chromosome.

For example, fibronectin—which regulates many cellular functions (Bao2009)—was shown by an early study to have a dynamic and highly extensible structure (Ohashi1999). The extensibility of fibronectin region in the extracellular matrix protein tenascin was independently found to arise from an entropic elasticity and it was suggested to be crucial in the regulation of cell adhesion and migration (Oberhauser1998). Interestingly, the force-extension relation was fit to the wormlike chain (WLC) model with a persistence length of 0.42 nm (Fig. 3.8 A) (see following for definitions of the WLC and the persistence length).

Strikingly, around the same time titin, a giant protein that is responsible for the muscle elasticity (Maruyama1997), was also found to follow a WLC pattern with a persistence length of 0.4 nm in another groundbreaking, early study (Rief1997) (Fig. 3.8 B). Since then numerous studies conducted on the entropic elasticity of a wide variety of proteinous filaments, both experimentally or computationally (see for example (Block2015)).

Similarly, nucleic acids are known from early studies to demonstrate entropic elasticity in response to mechanical stimuli. Perhaps the groundbreaking study in the field was performed by Bustamante et. al. where the stretch force was applied to the free end of the end-tethered λ-DNA to provide a direct mechanical measurement (Smith1992). Consequently, the group confirmed that the force-extension relation followed an entropic elasticity that can be fit to a WLC model (Fig. 3.8 C) with a persistence length of 53.4 nm (Bustamante1994).

In a similar fashion, early works on the elasticity of mitotic chromosome confirmed that this DNA-protein-combined filamentous structure also responds nonlinearly to stretching force and the corresponding force-extension curve emerged as a WLC model (Fig. 3.8 D) with a persistence length of 2.7 μm (Houchmandzadeh1997, Houchmandzadeh1999).

FG-repeat domains, as the most important role-player in selectivity battier, are intrinsically disordered (Denning2003b) and adopt a string-like shape (Rout2003).The polymeric nature of the FG-repeat domains necessitates considering the nonlinearity for the connecting springs, which amounts to having finite-extensibility.

All in all, the nonlinear elasticity arises solely from the Brownian motion and is an entropic effect (Dealy2006). This lies in the fact that, entropically, it is extremely unfavorable for a polymer to reach its fully extended length, because the corresponding end-to-end distance would be left by

only one state. The entropic elasticity in the biopolymer backbone is a fundamental characteristic and it must be described by an appropriate model.

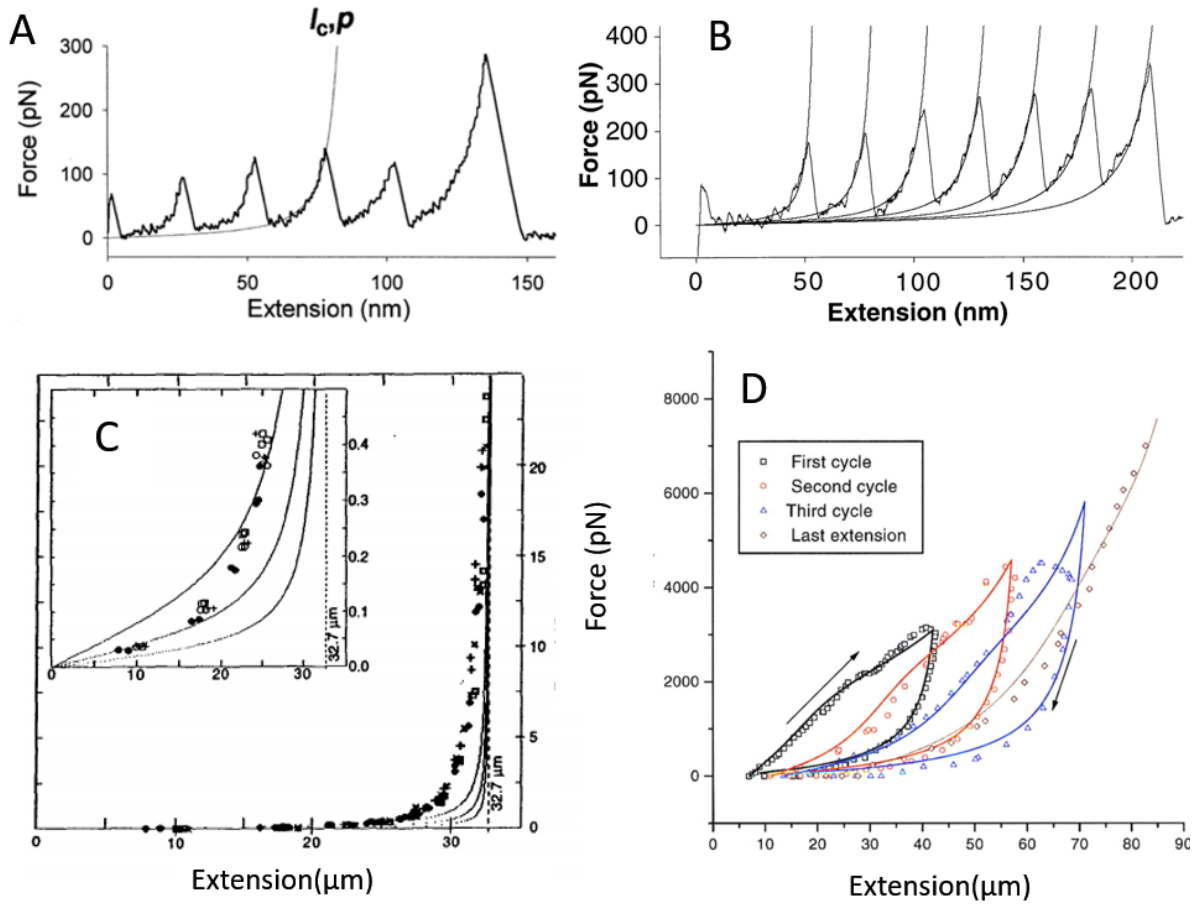


Figure 3.8. Early pioneering studies on the backbone's entropic elasticity of proteins, nucleic acids, and chromosomes. In all of these studies the emergence of the WLC model in the force-extension curve was confirmed. **A**) Force-extension curve of the extracellular matrix protein tenascin arises from the fibronectin region and is fit into a WLC model with the persistence length of 0.42 nm. Adopted from (Oberhauser1998). **B**) The force-extension curve for a single titin fragment. The persistence length was found to be 0.40 nm. Adopted from (Rief1997). **C**) Force-extension curves of four induvial λ-DNAs. Adopted from (Smith1992). Inset: the zoomed-in look into the low-extension part of the curve. When it is fit into WLC model, the persistence length is measured 53.4 nm. Adopted from (Bustamante1994). **D**) The force-extension of mitotic chromosome again follows a WLC model, with a persistence length of 2.7 µm. Adopted from (Houchmandzadeh1999).

There are several models proposed for the force-extension of finitely extensible nonlinear springs, including **finitely extensible nonlinear elastic (FENE)** model (Warner Jr1972), **Cohen-Pade**

approximation of the inverse Langevin function (Cohen1991), and the **wormlike chain (WLC)** (Marko1995). Nevertheless, all of the proposed models are qualitatively similar and predict comparable results for a wide range of biopolymers, implying the nonlinear elasticity in biological systems arises from the same origin (Broedersz2014).

From a molecular argument it can be shown that the elasticity of a polymer molecule composed of $N-1$ segments of length a is described by the inverse Langevin function (Bird1987):

$$F(r) = \frac{k_B T}{a} \mathcal{L}^{-1}\left[\frac{r}{(N-1)a}\right], \qquad (3.27)$$

where $\mathcal{L}^{-1}$ is the inverse to the Langevin function $\mathcal{L}(x) = \coth(x) - x^{-1}$. This equation was first derived by Kuhn and Grun (Kuhn1942) to show how a freely joined chain could describe the polymer's elasticity, given a probability distribution function of the segments follows a Gaussian function.

The inverse Langevin is not an analytic function, and its mathematical complexity makes it difficult to work with in theoretical and computational analysis of polymeric systems (Larson1999). To overcome this barrier, Warner (Warner Jr1972) proposed the FENE model as an empirical approximation of the inverse Langevin function to estimate the nonlinear elastic force of polymer molecules as:

$$\mathbf{F}_i^{FENE} = \frac{H\mathbf{r_i}}{1-\left(\frac{r_i}{r_0}\right)^2}, \qquad (3.28)$$

where $H = \frac{3k_B T}{(N-1)a^2}$ is the spring constant (Hur2000). $\mathbf{r_i}$ is the connecting vector of two ends of the segment i, and r_0 is the maximum extension, at which point the function becomes singular, representative of the finite extensibility of the polymer (Kröger2005).

As this approximation deviates from the inverse Langevin function, particularly at high extensions, Cohen applied the Pade's approximation to the inverse Langevin function (Cohen1991), and proposed a more accurate equation that is increasingly being used (Underhill2007) (Fig. 3.9 B):

$$F(\hat{r}) = \frac{k_B T}{b_K} \frac{\hat{r}\left(3-\hat{r}^2\right)}{1-\hat{r}^2}, \qquad (3.29)$$

where b_K is the chain's statistical Kuhn length (Hur2000). As illustrated in Fig. 3.9 A and B, the inverse Langevin function and its approximations, FENE and Pade, predict linear elasticity at small extensions (Kröger2005).

Finally, the WLC is a successful model, originally proposed to predict the micromechanics of the DNA nonlinear elasticity (Smith1992, Bustamante1994, Marko1995, Smith1996). This model is more powerful than inverse Langevin in that it *explicitly* incorporates the mechanical stiffness of the polymer through the persistence length, l_p.

Persistence length conveniently characterizes the polymer stiffness and is a measure of stiffness in response to thermal noise (Houchmandzadeh1999). As the polymer repeatedly snakes in response to thermal noise, the 'memory' of the tangent vector will be lost beyond the persistence length (Phillips2012). Thus, the persistence length is defined as the length over which the tangent vectors

are correlated (Netz2003). If $\mathbf{t}(s)$ is the tangent vector evaluated at the distance s along the polymer, then l_p determines how fast the tangent vector loses its memory (Phillips2012):

$$\langle \mathbf{t}(s).\mathbf{t}(u) \rangle = e^{-\frac{|s-u|}{l_p}} \qquad (3.30)$$

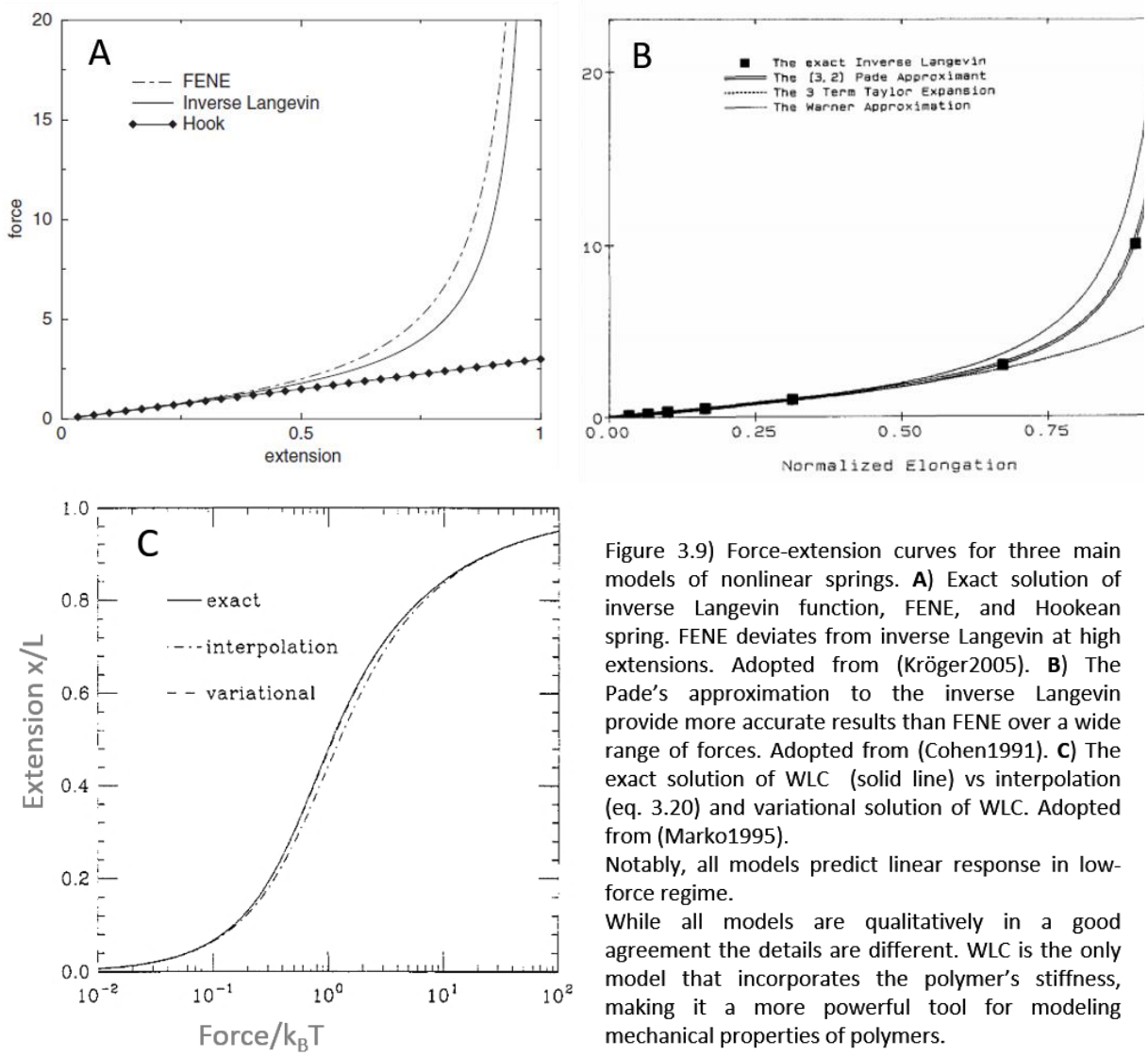


Figure 3.9) Force-extension curves for three main models of nonlinear springs. **A**) Exact solution of inverse Langevin function, FENE, and Hookean spring. FENE deviates from inverse Langevin at high extensions. Adopted from (Kröger2005). **B**) The Pade's approximation to the inverse Langevin provide more accurate results than FENE over a wide range of forces. Adopted from (Cohen1991). **C**) The exact solution of WLC (solid line) vs interpolation (eq. 3.20) and variational solution of WLC. Adopted from (Marko1995).
Notably, all models predict linear response in low-force regime.
While all models are qualitatively in a good agreement the details are different. WLC is the only model that incorporates the polymer's stiffness, making it a more powerful tool for modeling mechanical properties of polymers.

Depending on the relative length of l_p and L_c, polymers are classified into three groups in terms of their overall flexibility (Storm2005). The completely flexible polymers are those in which $l_p \ll L_c$; their elastic response is purely entropic. When l_p and L_c are comparable, the polymer is called semiflexible. These polymers are neither completely flexible nor are they rigid rods, but still

demonstrate a good degree of flexibility in response to thermal fluctuations (Broedersz2014). Their elastic response is perhaps more complicated that the first groups in that their backbone extensibility, arises from enthalpic effects, contributes in the overall response. The third class of polymers are on the other side of the spectrum, where l_p the is larger than the L_c, and thus the polymer essentially behaves as a rigid rod (Storm2005). Notably, the majority of biological filaments lie within the second class, i.e. semiflexible, yet their elastic response is treated merely from entropic considerations in literature.

FG-repeat domains are among rare biological filaments that are completely flexible. A list of a wide range of l_p's for different biological filaments and single-wall nanotube (SWNT) is provided in the Table 3.1. The value of $\frac{l_p}{l_{p,FG}}$ shows the stiffness of the filament relative to FG-repeat domains. It appears that FG-repeats are within the most flexible group of biological filaments.

Table 3.1 The length properties of several biopolymers and SWNT compared with FG-repeat domains.

Biopolymer	**Persistence length, l_p**	**$\frac{l_p}{l_{p,FG}}$ (order of magnitude)**	**Contour length, L_c**	**Approximate diameter (nm)**	**Reference**
FG-repeats (yeast NPC)	0.43 nm	1	80 - 260 nm	0.2	(Lim2006, Lim2007a, Nelson2013a)
Titin	0.40 nm	1	58 - 227 nm	0.33	(Rief1997) (Rodriguez2010)
Extra cellular matrix protein tenascin	0.42 nm	1	28.5 nm	5.2	(Oberhauser1998) (Spring1989)
Poly(U) RNA	0.79 nm	1	1 - 2 μm	2	(Lee1989, Vanzi2005)
DNA	50 nm	10^2	< 1 m	2	(Howard2001)
Chromosome	2.7 μm	10^4		800	(Houchmandzadeh1999)
Intermediate filament	0.3 - 1 μm	10^3 - 10^4	2 - 10 μm	9	(Mücke2004)
Single wall nanotube	10 μm	10^5	> 1 μm	< 1	(Fakhri2009)
F-actin	17 μm	10^5	2 - 20 μm	7	(Gittes1993, Dogic2004)
Microtubule	1 - 5 mm	10^7	10s of μm	25	(Gittes1993)

As the persistence length is an important parameter in analyzing mechanics of polymers, its explicit inclusion in the nonlinear elastic spring relation is of great importance. Marko and Siggia (Marko1995) performed a detailed statistical mechanical analysis of all force-extension DNA experiments available at the time, and suggested an elegant equation for the WLC model based on their interpolation:

$$F(r) = \frac{k_B T}{l_p}\left[\frac{1}{4}\left(1-\frac{r}{L_c}\right)^{-2} - \frac{1}{4} + \frac{r}{L_c}\right] \qquad (3.31)$$

Similar to the inverse Langevin, WLC model predicts a linear response in the low-force regime (Fig. 3.9 C). Remarkably, however, Marko and Siggia introduced the characteristic force of $\frac{k_B T}{l_p}$ as the criteria for low-force regime (Marko1995). If the extension force is less than the characteristic force, the response will be essentially linear.

Numerous independent studies emerged in following years, continued to date, proved the strength of this equation in predicting the mechanical behavior of literally all types of flexible and semiflexible polymers.

Later, Wang et. al. (Wang1997) modified the WLC model by taking into account the intrinsic enthalpic contributions of the chain's backbone. The enthalpic elasticity appears in the over-stretched regime as observed in double- and single-stranded DNA (Smith1996) as well as several stretching experiments of RNA (Vanzi2005). To accommodate this effect, Wang et. al. modified the WLC into an extensible WLC (E-WLC), which can be extended beyond its contour length:

$$F(r) = \frac{k_B T}{l_p}\left[\frac{1}{4}\left(1-\frac{r}{L_c}+\frac{F(r)}{K}\right)^{-2} - \frac{1}{4} + \frac{r}{L_c} - \frac{F(r)}{K}\right] \qquad (3.32)$$

This equation is no longer singular as the end-to-end extension approaches the contour length. Importantly, the extension force $F(r)$ has an implicit functionality in eq. (3.32), making it computationally intractable and less attractive. The key new parameter here is the stretch modulus K, representing the intrinsic resistance of the polymer backbone to longitudinal strain. In effects, K manifests the enthalpic contributions, and for an ideal WLC with uniform and isotropic elasticity it is related to the persistence length and the diameter of the chain as (Wang1997):

$$K = \left(\frac{16 k_B T}{d^2}\right) l_p \qquad (3.33)$$

Notably, the stretch modulus can be roughly estimated as the force required to stretch the polymer backbone twice its contour length (Vanzi2005).

It has been shown that this modified WLC mimics accurately the experimental data for the force-extension curve of poly(U) RNA, and predicts more accurate results, particularly at large force regime (Vanzi2005) as shown in Fig. 3.10. However, considering the implicit functionality of stretching force in the E-WLC model, the computational cost of calculating eq. (3.32) is much higher than the WLC, as the bond equations must be calculated in every single timestep for all of the bonds in the system.

After all, in regard to the FG-repeat domains, there is not such an over-stretched and large external force regime inside the NPC, which along with computational feasibility, strongly rationalizes the use of original WLC. In fact, the force-extension curves of FG-repeat domains have been shown

to closely follow the WLC model (Lim2006, Lim2007a). For the first time Lim et. al. (Lim2006, Lim2007a) performed an AFM experiment on an end-tethered FG-Nup, Nup153, and showed that the force-extension pattern follows a WLC model with the persistence length of 0.43±0.16 nm (Lim2007a) shown in Fig. 3.11.

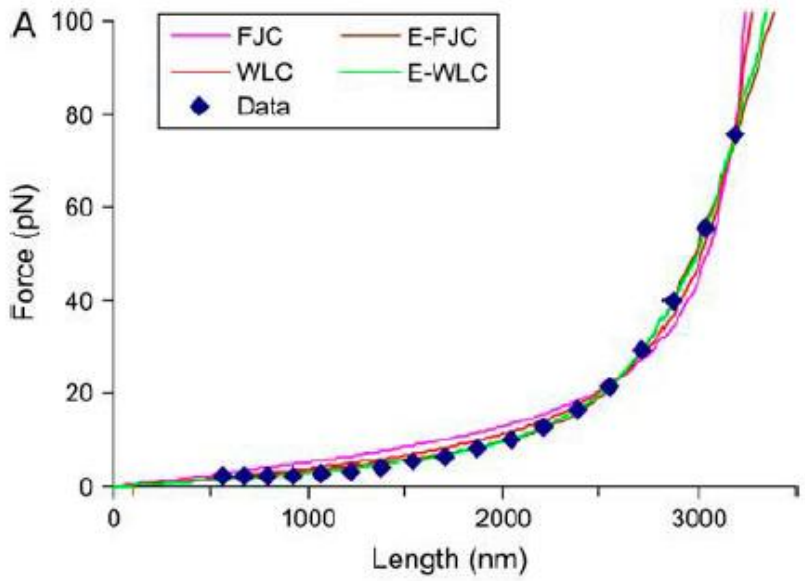


Figure 3.10. Comparison of different force-extension models in stretching ploy(U) RNA. The modified WLC (E-WLC) shows marginal accuracy over WLC at low-stretch regime. Howler, it becomes more evident at higher forces in the over stretched regime, where the RNA is extended beyond its contour length (after 3000 nm). From (Vanzi2005).

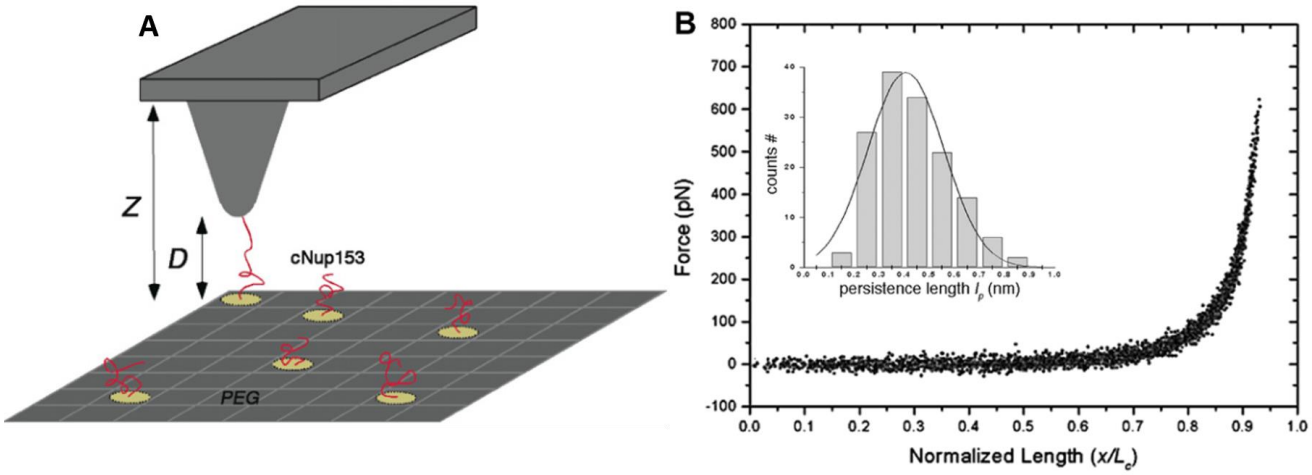


Figure 3.11. The pioneering Lim et. al. experiment on the tethered Nup153. A) The schematic of their experimental setup where Nup153 are end-tethered to the PEG surface. B) The mastercurve of an ensemble of 148 independent end-tethered Nup153's subject to the AFM force-extension experiment, showing a remarkable fit to the WLC model with the measured persistence length of 0.43 ± 0.16 nm (inset histogram). Adopted from (Lim2007a).

3.3.2.3 Angular

In addition to the elastic extension in the backbone, I also consider the bending potential energy in the model. For quite a long time, the backbone bending energy has been known to be important in both proteins (Mao1982) and DNA (Starr1995) structures. It may also be among the main features forming the structural stability of the NPC (Perez-Terzic1999, Cohen2003, Griffis2004). In its most basic form, a damped harmonic bending potential energy in a Brownian motion follows the Langevin dynamics (McCammon1976):

$$I\frac{d^2\theta}{dt^2}+\zeta\frac{d\theta}{dt}+k\theta=T_r(t), \qquad (3.34)$$

where θ is the angular displacement of a bond between two adjacent springs or rods. In this equation, I is the effective moment of inertia, ζ is the friction coefficient, k is the angular force constant, and $T_B(t)$ is the random Brownian torque.

Similar to translational Brownian motion discussed earlier, in the absence of hydrodynamic effects and in the limit of overdamped regime, eq. 3.33 simplifies to simple harmonic oscillator:

$$I\frac{d^2\theta}{dt^2}+k\theta=T_r(t), \qquad (3.35)$$

which would have, among other possible solutions, a cosine-based solution with the angular frequency of $\sqrt{\frac{k}{I}}$, an amplitude A and a phase angle of ϕ whose values depend on the specific problem.

In the computational biophysical modeling usually the following three different forms of functionalities are considered for the harmonic bending potential: function of $\left(\theta_{ijk}(t)-\theta_{ijk}^0\right)$ (Allison1986, Pei2009), function of $\cos\left(\theta_{ijk}(t)-\theta_{ijk}^0\right)$, or function of $\left(\cos\theta_{ijk}(t)-\cos\theta_{ijk}^0\right)$ (Schlick1989, Hakansson1994). In all of these, $\theta_{ijk}(t)$ is the angle between two consecutive bonds ij and jk, representing the angular displacement of the bond relative to the equilibrium angle θ_{ijk}^0.

Among these three main functionalities commonly used in computational simulations, the choice of $\Delta\cos\theta_{ijk}$ has several advantages, which were the base of my decision to use it in the current work. First, from calculation view point, it is much easier to calculate the cosine between two vectors (springs) $\mathbf{r_{ij}}$ and $\mathbf{r_{jk}}$ as:

$$\cos\theta_{ijk}(t)=\frac{x_{ij}x_{ij}+y_{ij}y_{jk}+z_{ij}z_{jk}}{\left(x_{ij}^2+y_{ij}^2+z_{ij}^2\right)^{\frac{1}{2}}\left(x_{jk}^2+y_{jk}^2+z_{jk}^2\right)^{\frac{1}{2}}}, \qquad (3.36)$$

than finding the angle between them.

Secondly, cosine-based potentials are computationally more efficient (Schlick1989) and numerically stable (Hakansson1994). Thirdly, a cosine-based potential works better for large fluctuations compared to the harmonic θ potential energy (Schlick1989). This is indeed essential for an accurate modeling of the fine structure of FG-repeats, which undergo high degrees of fluctuations due to interaction with cargo and thermal noises.

Thus, I have used the following potential energy to account for the bending energy of two adjacent bonds in the NPC main structure (Fig. 3.6 B):

$$U_{ijk}^{bend} = \frac{1}{2} b\left(\cos\theta_{ijk}(t) - \cos\theta_{ijk}^{0}\right)^2 \qquad (3.37)$$

In eq. (3.36), b is the bending potential constant is a material property, whereas the equilibrium angle θ_{ijk}^{0} depends on the NPC geometry. Because there is no such detailed information about the NPC material properties, I arbitrarily chose the bending force constant equal to the minimum value that guarantees structural stability during transport (100 $k_B T$).

As a requirement of the polymers' discretization scheme, it is also necessary to account for the bending potential energy between successive segments in the discrete WLC (Allison1986, Underhill2004a, Pei2009). Here also I use the same cosine-based bending potential energy (eq. 3.37). However, the bending force constant is no longer arbitrarily chosen. It is directly related to the chain flexibility, and thus is a function of persistence length and the bond length. Schellman (1974) analytically derived the functional form of the bending force constant for DNA as a flexible linear molecule (Schellman1974), which is widely used for flexible and semiflexible filaments. In this work, I have calculated the bending force constant for FG-repeat domains as following:

$$b = k_B T \frac{l_p}{dL_c} \qquad (3.38)$$

where dL_c is the contour length of the segment.

It appears that b is three orders of magnitude smaller for FG-repeats than the NPC main structure. This is expected as the chain curvature in the WLC is based on thermal fluctuations (Ogden2006) that are on the order of $k_B T$. Our model consistently predicts that, in the course of transport, FG-domains show a high degree of flexibility exhibiting the character of natively unfolded proteins (Denning2003b). We found in our simulations that these fluctuations are the main barrier preventing inert cargo from entering the nucleus and hence the main factor in selective rejection (Moussavi-Baygi2011a, Moussavi-Baygi2011b). Our model also suggests that the fluctuations are important to form the ultrafast reconstruction of the FG-meshwork within the NPC (Moussavi-Baygi2016).

3.3.3 Nonbonded forces

While bonded forces are permanent and account for the shape's integrity, the important class of transient, breakable nonbonded interactions underlie the dynamical properties and behavior of the system. Within a computational modeling, the bonded interactions' constants are taken from material properties if available, and can be guessed otherwise. This is not as straightforward for nonbonded interactions, however, because a small change in a force constant and form can lead to a totally different behavior of the system. Cares thus should be taken in designing the appropriate force-fields. In this section I will describe the nonbonded force-field development used in the current work.

3.3.3.1 Short-range repulsive force.

This is perhaps the most basic, yet the important, type of the nonbonded forces with a short-range nature that accounts for collisions and does not allow two beads occupy the same space at the same time. As the most fundamental property of any physical particle, interestingly, the short-range repulsions produce a long-range effect (Yamakawa1971) in polymer solutions termed excluded

volume (EV) effects. In its most basic form, the EV of a hard sphere with the *diameter* d is the volume of an imaginary sphere with the *radius* d (Teraoka2002b) as depicted in Fig. 3.12. In effect, the center of the sphere 2 in the Fig. 3.12 A is exclude by the imaginary sphere shaded in light red.

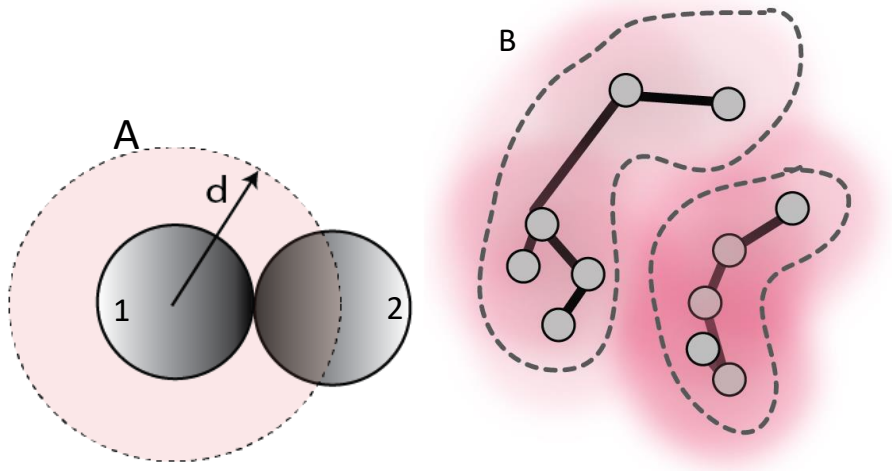


Figure 3.12. The excluded volume shown for A) a hard sphere, and B) for a polymer modeled as bead-spring.

In polymers, the EV is a claimed volume by a polymer or a segment that cannot be occupied by other polymers or segments. This property has a big contribution in long-range effects like the intrinsic viscosity and friction coefficient in a polymer solution (Yamakawa1971). There are numerous studies showing the importance of considering EV in modeling (bio)polymers and their solutions (Knudsen1996, Graessley1999, Li2000, Pierleoni2000, Xie2006).

There are indeed some simplistic models, like ideal chains (Meijer1991, Groot1998) or some probabilistic models of physical systems (Lee1973, Smit1995), that do not consider repulsions and therefore two particles can take up the same space with no EV. Although the predicted behavior of such simulations deviates from the real systems in several aspects, still those models can qualitatively predict some thermodynamic properties (Lee1973, Smit1995, Groot1998).

When it comes to the FG-repeat biopolymers inside the crowded environment of the NPC, however, the EV becomes more important because of the limited space inside the pore and different competitions (monomer-monomer and cargo-monomers). Therefore, as we want to obtain a detailed picture of the permeability barrier and the cargo transport across the NPC, we obviously need to consider this vital interaction.
To replicate the excluded volume effects of polymer aggregations in a bead-spring model, several potentials, namely Lennard-Jones, Morse, and other exponential forms are used (Prabhakar2004b).

The most basic potential energy for repulsion is the hard sphere:

$$U_{ij}^{HS} = \begin{cases} \infty, & r_{ij} < \sigma_{ij} \\ 0, & r_{ij} \geq \sigma_{ij} \end{cases}, \qquad (3.39)$$

with σ_{ij} being the sum of radii of particles i and j.

The hard-sphere potential is more like a toy model and clearly not suitable for real-system simulations due to its exclusively sharp behavior.

Lennard Jones 12-6 (LJ) potential is perhaps the most-widely-used form to model the van der Waals pairwise interaction and to mimic the EV effects in MD and coarse-grained simulations (Ashurst1975, Neogi1996):

$$U_{ij}^{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{3.40}$$

where σ_{ij} is the separation distance at which $U_{ij}^{LJ} = 0$, and ε_{ij} is the minimum energy, or well depth (Israelachvili2011, Kuriyan2012).

LJ potential can model the shortrange atomic attraction-repulsion: Once two neutral atoms become close enough to each other their fluctuating electron clouds rearrange, resulting in a transient dipole created (Israelachvili2011). This will produce an attractive force between atoms, which would be dominated by a much stronger repulsive force as soon as two atoms come closer to the extent that their electronic orbitals overlap (Kuriyan2012).

The LJ potential, however, needs a very small timestep in simulation as its behavior is very steep below σ (Knudsen1996, Cifre1999), leaving it computationally expensive (Andrews1998). Regardless of the computational cost, a short timestep should be avoided in Brownian dynamics due to physical considerations (see section "The choice of timestep in the Brownian Dynamics").

For computational simulation of polymeric systems, Milchev et. al. (1993) thus suggested another potential, termed Morse potential (reviewed in (Prabhakar2004b)), which allows larger timesteps than LJ does (Milchev1993):

$$U_{ij}^{M}(r_{ij}) = \varepsilon_{ij}\left(e^{\left[-2\alpha\left(r_{ij}-r_{\min}\right)\right]} - 2e^{\left[-\alpha\left(r_{ij}-r_{\min}\right)\right]}\right), \tag{3.41}$$

where $r_{\min} = \sqrt[6]{2}\,\sigma_{ij}$ and α is the potential parameter with the unit of inverse of the length.

Although the Morse potential is considered an alternative to LJ, they are similar and both behave steeply as two particles come close to each other below the characteristic length (Fig. 3.13). Like LJ, the Morse potential also has a sharp minimum at $r_{\min}$ for large values of α (Binder1995). The drawback for the WLC is that at short distances the repulsive force is extremely high and thus two beads experience large repulsion, which can lead to rashness in nonlinear springs like WLC.

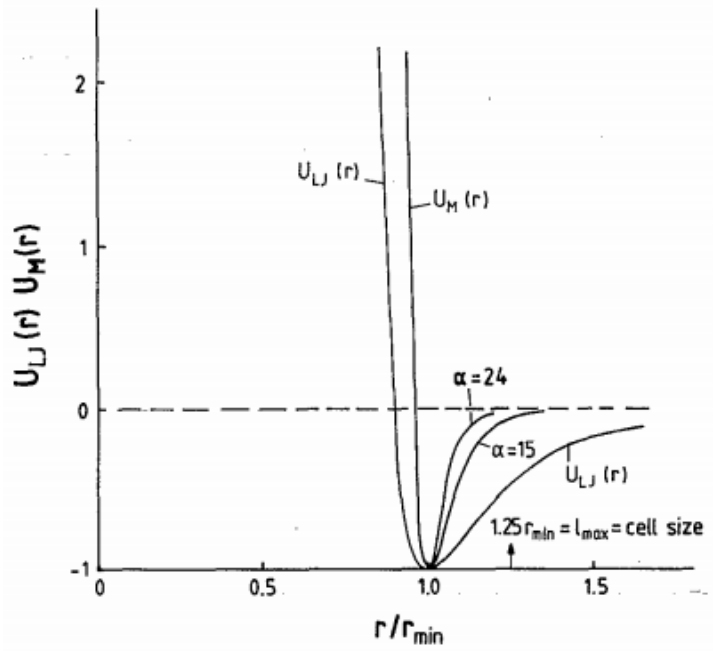


Figure 3.13. Comparison of LJ and Morse (for α=15 and 24) potentials. Both Morse and LJ have attraction and repulsion parts, and their behaviors are steep for distances below the characteristic length. Adopted from (Milchev1993)

After all, both LJ and the Morse potentials are purely phenomenological as they do not account for any quantum-chemical effects at atomic resolution (Binder1995). They have both attraction and repulsion energies. The attraction part arises solely from the quantum-chemical effect of an induced transient dipole, and it is not really necessary in a coarse-grained model.

All of these strongly suggest against the LJ and the Morse potentials in polymeric simulations under Brownian dynamics, although they have been widely used in such systems by many studies.

To overcome these barriers, an exponential, pure repulsive potential was suggested in Brownian dynamic simulations as (Rey1992):

$$U_{ij}^{rep}(r_{ij}) = \begin{cases} \varepsilon e^{-\frac{r_{ij}}{\sigma}} & r_{ij} \geq r_{cut} \\ 0 & r_{ij} < r_{cut} \end{cases} \qquad (3.42)$$

The repulsive force from the eq. 39 then will be:

$$F_{ij}^{rep}(r_{ij}) = -\frac{\partial U_{ij}^{rep}}{\partial r_{ij}} = \begin{cases} \frac{\varepsilon}{\sigma} e^{-\frac{r_{ij}}{\sigma}} & r_{ij} \geq r_{cut} \\ 0 & r_{ij} < r_{cut} \end{cases} \qquad (3.43)$$

Many studies have successfully employed this potential to account for particles repulsion and excluded volume effects (Cascales1991, Rey1992, Knudsen1996, Cifre1999).

Eq. 3.42 represents a very soft potential with no steep behavior, and is numerically stable, computationally efficient, and does not set any requirement for timestep; however, its resulting force (eq. 3.43) suffers from a discontinuity at $r_{ij} = r_{cut}$ (Fig. 3.14, red line). Therefore, I used the shifted version of the repulsion force equation for the modeling in this dissertation:

$$F_{ij}^{rep}(r_{ij}) = \frac{\varepsilon}{\sigma} e^{-\frac{r_{ij}}{\sigma}} \left(1 - e^{-\frac{\left(r_{cut} - r_{ij}\right)}{\sigma}} \right), \qquad (3.44)$$

which is now continuous over all r_{ij}'s as can be seen in Fig. 3.14 (black line).

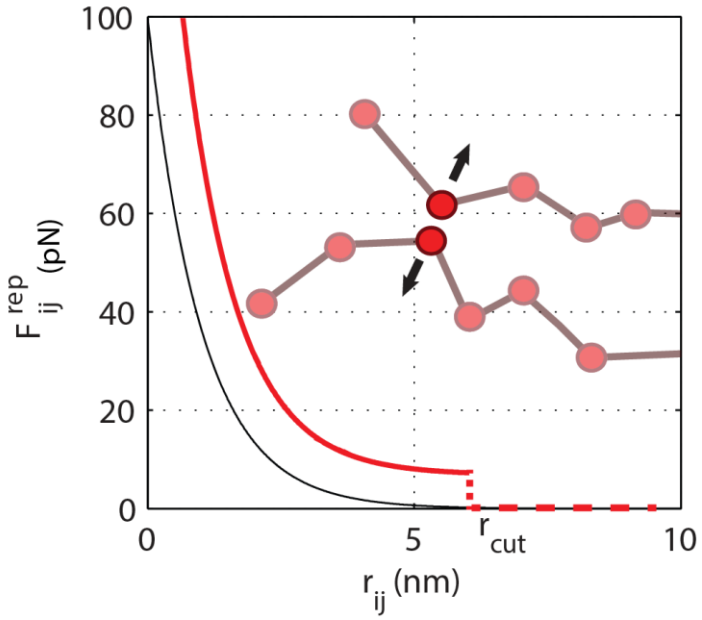


Figure 3.14. The original exponential force (red line; eq. 3.43) has a discontinuity at r_{cut}. The shifted force that is developed and used in the current dissertation (black line; eq. 3.44) is continuous over all r's.

3.3.3.2 Electrostatic force

The electrostatics is among the most fundamental interactions occur in the cell biology. The most basic form of an electrostatic interaction between two point-charges q_i and q_j is the famous long-range Coulombic potential:

$$U_{ij}^{Coul} = \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}}, \qquad (3.45)$$

in which the r_{ij} is the distance between two charges, ε_0 is the vacuum permittivity, and ε_r the relative permittivity of the medium. Relative permittivity, or dielectric constant, of a given medium filled by a specific material is a measure of polarizability of the material in response to an applied electric field (Davis1990). The higher the polarizability of the material, the stronger the dipoles

are induced in the medium, and consequently, the more the magnitude of the pairwise electrostatic energy U_{ij}^{Coul} is reduced. It is known that proteins are dielectrics (Neves-Petersen2003), due to presence of numerous permanent and induced molecular dipoles. More specifically, the relative permittivity of the cytoplasmic environment has measured to be 45.0 (Fikus1989), simply meaning the electrostatic interaction of two charges in such an environment is 45 times less than that in the vacuum. It is, however, important to note that the dielectric materials do not change the physics of interaction, i.e., the same long-range Coulombic relationship is kept.

The physics of electrostatic interaction dramatically changes when interacting charges are put in a solution, containing rapidly diffusing charged atoms, i.e. ions. It its most comprehensive form, the classical, as opposed to quantum mechanical, treatment of electrostatic interaction in such environment is based on the nonlinear Poisson-Boltzmann equation as (Honig1995):

$$\nabla.\left[\varepsilon(\mathbf{r})\nabla.\Phi(\mathbf{r})\right]-\varepsilon(\mathbf{r})\kappa^2\sinh\left[\Phi(\mathbf{r})\right]+\frac{4\pi\rho^f(\mathbf{r})}{k_BT}=0, \qquad (3.46)$$

in which $\Phi(\mathbf{r})$ is the dimensionless electrostatic potential, $\varepsilon(\mathbf{r})$ is the variable dielectric constant of the medium, ρ^f is the fixed charge density, and $\kappa=\lambda^{-1}$, where λ is the Debye screening length whose magnitude depends solely on the properties of the solution. In fact, the Debye length (the second term in eq. 3.46) accounts for the effects of mobile ions in the solutions, and would vanish when ions are not mobile ($\kappa=0$).

The magnitude of λ is equal to the thickness of the double-layer that is formed by diffusing counterions around the charge q is the solution. This layer is called **diffuse electric double-layer** and its thickness in a solution of different types of ions is given by (Israelachvili2011):

$$\lambda=\left(\frac{\varepsilon_0\varepsilon_r k_BT}{\sum_i \rho_i\left(ez_i\right)^2}\right)^{\frac{1}{2}}, \qquad (3.47)$$

where e is the fundamental charge, z_i and ρ_i are the valency and the ionic concentration of the ion i, respectively, ε_0 is the vacuum permittivity, and ε_r is the relative permittivity.

Under the condition of small electrostatic potential, $\sinh\left[\Phi(\mathbf{r})\right]$ can be approximated by $\Phi(\mathbf{r})$, yielding the linearized Poisson-Boltzmann equation (Honig1995, Fogolari2002). When there is no dielectric discontinuity (Honig1995) and in the absence of any fixed charge, the Deybe-Hückel approximation is retrieved and the equation 2.43 reduced to:

$$\nabla^2\Phi-\kappa^2\Phi=0 \qquad (3.48)$$

Now, for two particles of radii a_i and a_j with central net charges of q_i and q_j, the electrostatic energy potential can be obtained (Medina-Noyola1980, Ando2010, Dlugosz2011, Israelachvili2011) as:

$$U_{ij}^{elec}(r_{ij})=\frac{q_iq_j}{4\pi\varepsilon_0\varepsilon_r(1+\kappa a_i)(1+\kappa a_j)r_{ij}}e^{-\kappa\left(r_{ij}-a_i-a_j\right)}, \qquad (3.49)$$

which is usually referred to as the Deybe-Hückel or the DLVO (Derjaguin–Landau,Vervey–Overbeek) approximation for two net charges (q_i, q_j) in an ionic solution with a weak electrostatic potential. The relative permittivity, ε_r, for the cytoplasmic environment is 45.0 (Fikus1989), and the Debye screening length is $1.0\,nm$ for the cytoplasmic ionic strength (Nelson2013b).

In the current dissertation, the pairwise potential energy of electrostatic interaction between two charged beads, or a charged bead and the charged cargo surface is calculated based on the surface-to-surface distance as:

$$\begin{cases} U_{ij}^{elec}(r_{ij}) = \dfrac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r(1+\kappa a_i)(1+\kappa a_j)r_{ij}} e^{-\kappa\left(r_{ij}-a_i-a_j\right)} , & s_{ij} \geq \lambda \\ U_{ij}^{elec}(r_{ij}) = \dfrac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}} , & s_{ij} < \lambda \end{cases} \tag{3.50}$$

As indicated in Fig. 3.15, first the surface-to-surface distance, s_{ij}, between to charged beads is calculated and compared with the Debye screening length. The rationale is that if $s_{ij} < \lambda$, mobile ions cannot form a diffuse electric double-layer around the beads, and thus the charged beads electrostatically feel each other via a pure Coulombic interaction.

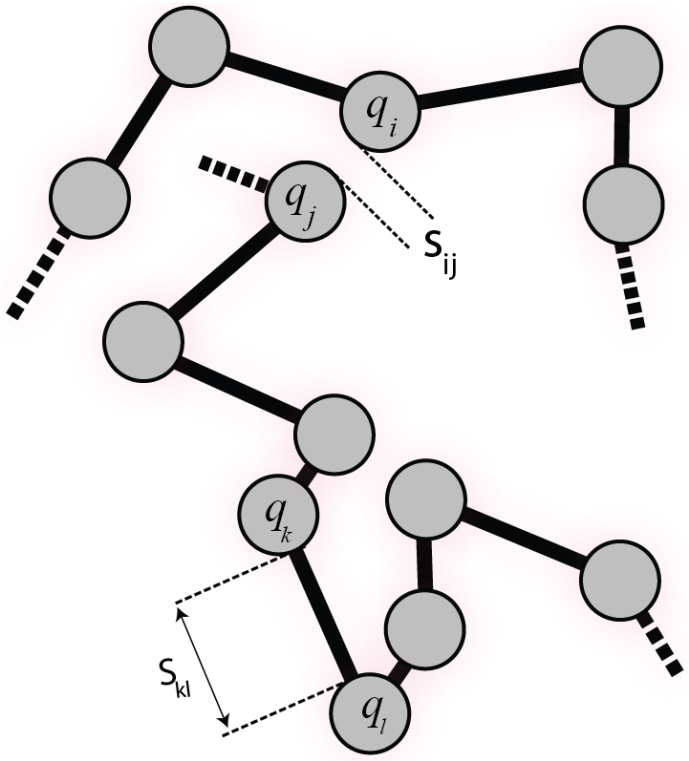


Figure 3.15. The physics of the electrostatic interaction depends on the surface-to-surface distance between two charged beads. $S_{ij} < \lambda$, thus the Coulombic interaction is applied between corresponding beads, while $S_{kl} > \lambda$, and thus, the screened Debye-Huckel interaction is applied between two corresponding beads.

3.3.3.3 Hydrophobic force

The Hydrophobic effects are relatively subtler to model since they are mainly entropic in nature (Tanford1980). As such, a hydrophobic interaction is long-range in nature (Claesson1988), because an exposed hydrophobic surface tends to decrease configurational entropies of surrounding water molecules (Kuriyan2012), which in turn would unfavorably affect the hydrogen bonds in the next layers of water molecules.

Although Trp has the highest hydrophobicity index and from this viewpoint, is naturally considered to be the most hydrophobic residue (Osada1997), but because of the nitrogen in its indole ring, Trp is significantly more polar than Phe. The same is true for the other aromatic residue, i.e. Tyr, which due to its hydroxyl group is way more polar than Phe (Nelson2013a). Moreover, in the context of FG-Nups, Phe is the most abundant hydrophobic residue within the disorder domains. These have made Phe a unique candidate to perform the hydrophobic bond in the FG-repeat domains.

Thus, in this work the hydrophobic interactions between all Phe residues are carefully taken into account. To do so, the area of interaction is chosen as the solvent-accessible surface area (SASA) of the nonpolar side chain of Phe in its extended conformation, i.e. $175\,\mathring{A}^2$ (Rose1985, Radzicka1988, Creighton1993) .

For the pairwise hydrophobic interaction, I tested a wide range of hydrophobic force fields (Claesson1988, Kellogg1991, Yoon1997, Hummer1999, Curtis2002, Donaldson2011). To determine the optimized force filed for FG-repeats five criteria were implemented: *i*) availability of the force field parameters for FG-Nups, *ii*) ability to represent the overall known physiological properties of FG-Nups, *iii*) minimal number of parameters involved, *iv*) numerical efficiency, and *v*) simplicity of implementation in the simulation.

Remarkably, the double-parameter exponential potential energy was found to be the most optimized force field for FG-Nups. This potential is proposed for purely hydrophobic interaction energy between two hydrophobic surfaces in a polar solution like water as (Israelachvili1992):

$$U_{ij}^{hphb} = -2\gamma \exp\left(-r_{ij}/\mu\right), \qquad (3.51)$$

with $\gamma = 0.4\ k_B T/nm^2$ representing the hydrophobic affinity per unit area and $\mu = 1.0\,nm$ being the characteristic length of interaction (Moussavi-Baygi2011a, Moussavi-Baygi2011b, Moussavi-Baygi2016).

The same hydrophobic potential is also applied between the cargo surface and FG-repeats. To represent NTRs bound to the cargo surface adequate number of hydrophobic binding spots are placed and restricted to only one side of the cargo, consistent with localization of hydrophobic binding patches on the convex surface of kap β (Peters2009). Indeed, the NTR-Phe hydrophobic affinity is set stronger than Phe-Phe affinity (Patel2007) to yield the nanomolar apparent dissociation constant between NTRs and FG-Nups (Peters2009) through:

$$k_d = \exp\left(U_{NTR-F}^{hphb}/k_B T\right). \qquad (3.52)$$

3.3.4 Discretization scheme of the FG-repeat domains into bead-nonlinear springs:

In the current dissertation, the overall 3D geometrical map of the NPC central channel was adopted from the Yang et. al. (Yang1998), while the upper and the lower half-spoke domains taken from Alber et. al. (Alber2007a, Alber2007b). The central channel is composed of eight identical spoke domains, composed of upper and lower half-spokes, and five FG-Nups per half-spoke (Alber2007a). I first made these two half-spokes, shown as light and dark green in Fig. 3.17A, with all of known FG-Nups tethered to them (Fig. 3.17). Subsequently, the entire spoke domain was rotationally repeated eight times around the central axis of the pore to form the central channel. This procedure is consistent with the proposed assembly of the NPC in vivo (Alber2007a, Alber2007b).

FG-motifs are scattered across the NPC structure and electron microscopy has localized FG-Nups to the central channel as well as the cytoplasmic and nuclear faces (Bayliss2000a). In agreement with this, in the 2D model, FG-repeats are considered all over the NPC (Fig. 3.16), while in 3D model due to computational considerations only central was considered. The anchor domains were taken from the work of Yamada et al. (Yamada2010), while positions of tethering points on the spoke domain (Fig. 3.17) were calculated from the work of Alber et al. (Alber2007a).

The discretization process for the intrinsically disordered FG-repeat domains is subtler than that for the NPC backbone for several reasons. First, the springs in the FG-Nups are nonlinear with mathematical singularity as opposed to Hookean springs in the backbone. Secondly, FG-repeat

domains are extremely dynamics as opposed to the backbone that behaves, for the most part, statically in our time scale of interest. Third, beads in the FG-repeat domain have a wider list of attributions compared to the backbone. They can be hydrophobic, negatively or positively charged, neutral, or any combination.

The process of FG-Nups discretization begins with sequencing the disordered domains of FG-Nups. As a prerequisite, however, we need the biophysical properties of all of the residues as listed in the Table 3.2. The sequence of the disordered domains of all of the 11 yFG-Nups (Yamada2010) were taken from UnitProt summarized in Table 3.3, as used in the current model. Incidentally, the majority of disordered FG-Nups have a larger number of Gly residue than Phe, indicating of the highly flexibility of disordered domains, which is consistent with their function and can describe why the small value of the l_p for FG-repeats (Table 3.1).

Individual FG-repeat domains are then discretized into WLC segments, with each bead representing five consecutive residues in the sequence, and the contour length of the spring being the N-terminal of the first residue to the C-terminal of the last one. The α-carbon-α-carbon distance between two adjacent atoms is 0.38 nm, but because of the tetrahedral geometry, the maximal extension per reside is about 0.34 nm (Yang2000, Ohashi2007). This is the effective length per residue in the current model for all disordered domains of FG-repeats.

Each bead encompasses the biophysical properties of its own constituting resides, including charge, hydrophobicity, and vdW radius as depicted in Fig. 3.16. Accordingly, there are six different type of beads in the model that arise from the constituting residues (Figs. 3.16 and 3.17):

1. Pure hydrophobic bead, representing a monomer that has Phe with no charged residue in the monomer.
2. Hydrophobic-negative, when there are negatively charged residues (Asp, Glu) in addition to the Phe in the monomer.
3. Hydrophobic-positive, when there are positively charged residues (Lys, Arg, His) in addition to Phe.
4. Pure negative, when the monomer is negative without any hydrophobic residue
5. Purely positive, opposite of the previous one
6. Neutral with zero net charge, when there are only neutral residues in the monomer.

The net charge of each bead is the sum of the charges of constituting residues. However, if one chooses to develop the model on a coarser basis, for example large portion of a Nup per simulation particle, one can estimate the effective charge of each monomer using Henderson-Hasselbalch equation under a desired pH (Moore1985). Among others, EMBOSS is a tool available online to calculate the effective charge of proteins as a function of pH based on their sequence. Initially, I used this tool when I started developing the 3D NPC model on a literally coarse basis.

The WLC springs play the role of the hydrophilic linkers between FG-motifs. Hydrophilic linkers are a sequence of 5-30 charged amino acids between FG-motifs (Jovanovic-Talisman2009b, Colwell2010a) that hypothetically drive the FG-repeat regions into disorder (Denning2003b). Based on their MD simulations on various peptides of disordered domains of FG-Nups, Dolker et. al. suggests that these linkers contribute to the interactions between FG-domains (Dolker2010). Consistently, our model predicts that the polymeric elasticity of the discrete WLC combined with the thermal fluctuations dictates randomness to the FG-repeat domains and increases the chance of FG-motifs to meet and interact with each other.

The WLC force-extension eq. (3.31) for a discretized segment of FG-repeat domains with length dL_c is then written as:

$$F_{ij}(r_{ij}) = \frac{k_B T}{l_p}\left[\frac{1}{4}\left(1-\frac{r_{ij}}{dL_{c,ij}}\right)^{-2} - \frac{1}{4} + \frac{r_{ij}}{dL_{c,ij}}\right] \qquad (3.53)$$

Importantly, in eq. (3.53) discretization of each FG-repeat should be such that an individual WLC segment represents a large number of persistence lengths to satisfy the condition of $dL_c \gg l_p$ (Marko1995). This ensures that the discretization criterion of the WLC model in which each spring must be sufficiently larger than the persistence length is satisfied (Underhill2004a), otherwise the WLC discretization scheme would not be physically accurate, and bead-stick model should be chosen instead. Fundamentally speaking, this requirement arises from the difference in the bead-stick vs bead-spring models (Fig. 3.2) for flexible polymers.

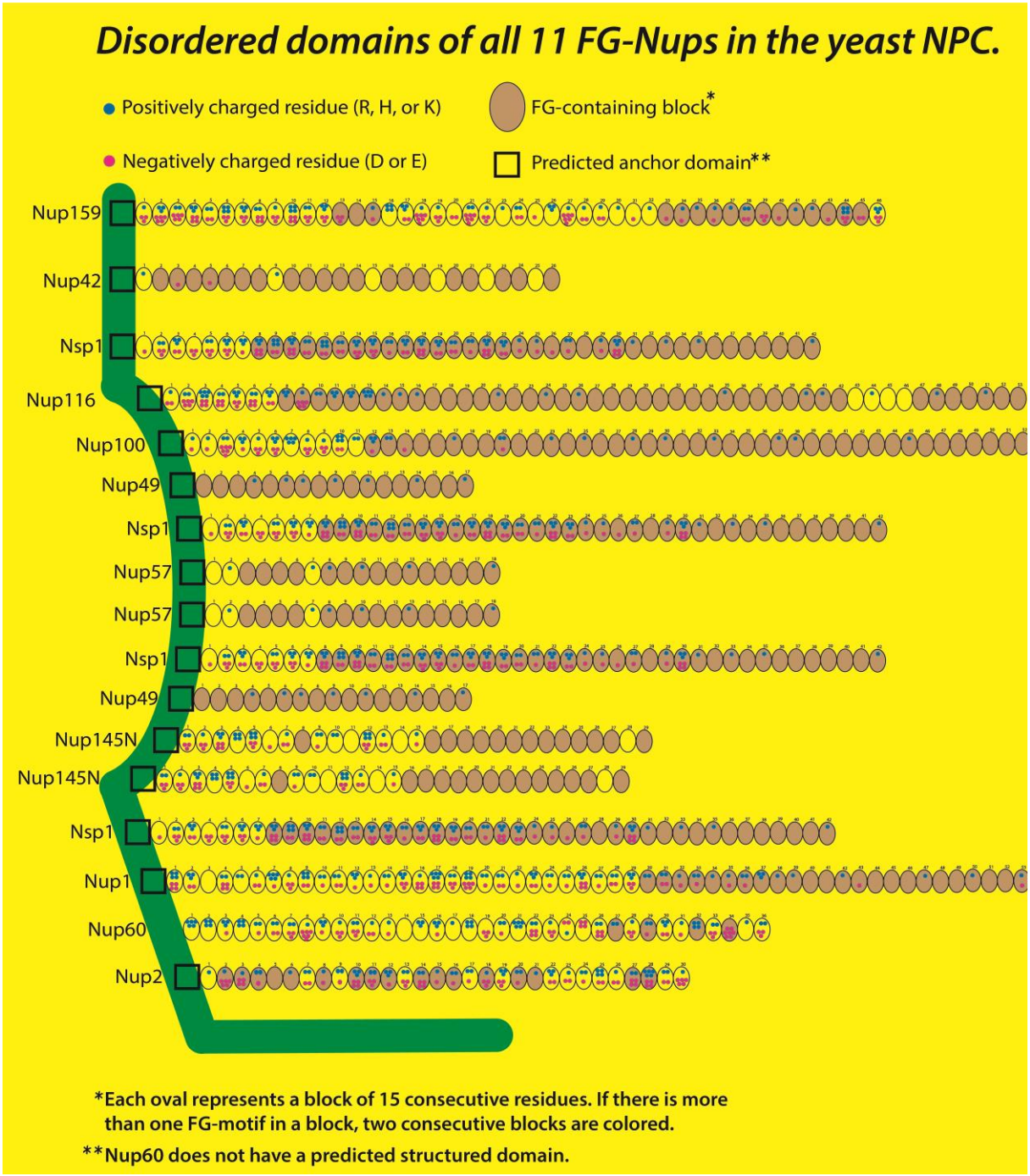


Figure 3.16. The 2D representation of discretization scheme of the 11 FG-repeat domains in the current model for an octagonal arm of the yeast NPC, composed of a cytoplasmic filament, a spoke domain, and a nuclear basket rod. The connecting WLC springs are not shown for clarity. Each bead represents a monomer, which includes the positive and negative charges as well as the hydrophobicity of constituting residues. Positive and negative charges of a monomer are represented by adequate number of blue and red dots inside the oval. Hydrophobic monomers are tanned. Adopted from (Moussavi-Baygi2016).

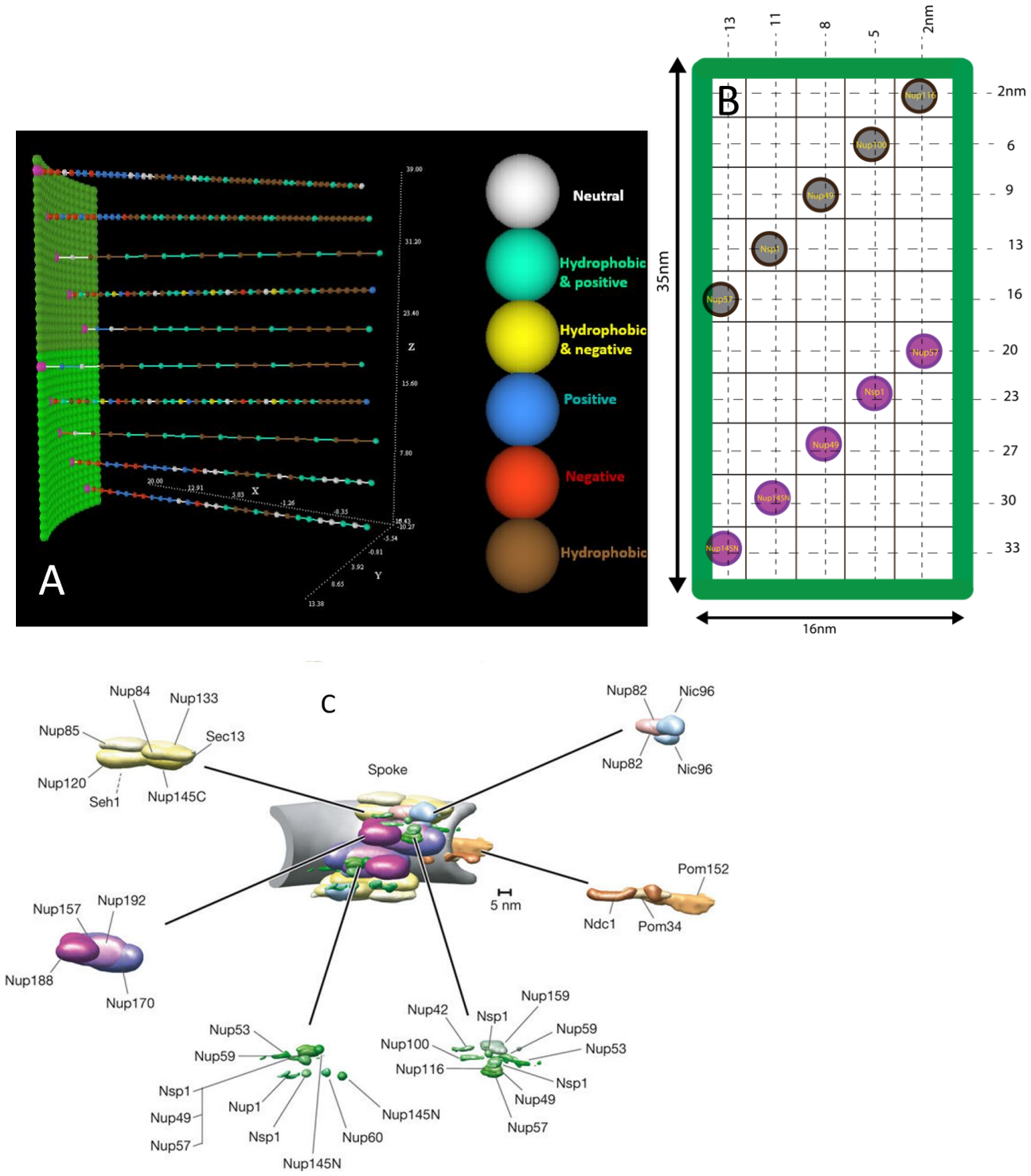


Figure 3.17. **A)** The single spoke with disordered FG-Nups being in their initial, cramped configurations. FG-Nups are modeled as bead-spring. Upper and lower spokes are colored in dark and light green, with the tethering point of each FG-Nup colored in pink. Each bead in the FG-Nup represents a monomer with the following coloring scheme: brown is hydrophobic (HB), yellow is hydrophobic-negative (HN), cyan is hydrophobic-positive (HP), red is purely negative (PN), blue is purely positive (PP), and white is neutral with zero net charge (N). Adopted from (Moussavi-Baygi2016). **B)** The exact location of FG-repeat anchor points on the spoke, projected onto a 2D surface. **C**) The location of each Nup in a spoke domain as calculated by a pioneering study of Alber et. al. (2007) Adopted from (Alber2007a).

Table 3.2. The biophysical properties of residues as used in the current model.

Residue	vdW radius (Å) (Creighton1993)	vdW volume ($Å^3$)	vdW surf ($Å^2$)	Mass (Da) (Creighton1993)	hydrophobicity index $kJ.mol^{-1}$ (Kuriyan2012)
Ala	2.5196	67	79.77509	71.09	2.1
Arg	3.2814	148	135.3086	156.19	7.5
Asn	2.8405	96	101.3907	114.11	3.8
Asp	2.7903	91	97.83894	115.09	15
Cys	2.7382	86	94.22144	103.15	0
Gln	3.008	114	113.6983	128.14	3.3
Glu	2.9633	109	110.349	129.12	15
Gly	2.2545	48	63.87216	57.052	4.5
His	3.0427	118	116.3426	137.1412	5
Ile	3.0935	124	120.2537	113.1595	-4.5
Leu	3.0935	124	120.2537	113.1595	-5
Lys	3.1824	135	127.2643	128.1742	11.7
Met	3.0935	124	120.2537	131.19	-2.9
Phe	3.1824	135	127.2643	147.18	-7.1
Pro	2.78	90	97.12085	97.12	0.4
Ser	2.5927	73	84.46939	87.0782	2.1
Thr	2.8106	93	99.26727	101.1051	0.8
Trp	3.3887	163	144.3032	186.2133	-8.8
Tyr	3.2288	141	131.0077	163.176	-2.9
Val	2.9266	105	107.6326	99.14	-2.1

Table 3.3. The summarized biophysical data of the FG-Nups' disordered domains used in the current model. Adopted from (Moussavi-Baygi2016)

Nup	Proteomic accession number	Approximate localization inside the NPC	Disordered domain (Yamada2010)	Number of F	Number of FG	Number of G	Length (aa)	Ratio of #G/#F	Total positive charge in disordered domain	Total negative charge in disordered domain	**Number of copies inside the channel**
Nsp1	P14907	Cytoplasmic, central, and nuclear sides	1-617	57	33	62	617	***1.09***	73	-65	32
Nup116	Q02630	Central channel	172-960	50	37	126	789	***2.52***	49	-36	8
Nup100	Q02629	Central channel	1-800	54	44	111	800	***2.06***	42	-26	8
Nup49	Q02199	Central channel	1-251	18	17	53	251	***2.94***	7	0	16
Nup57	P48837	Central channel	1-255	19	16	59	255	***3.11***	7	0	16
Nup2	P32499	Nuclear side	160-600	34	11	20	441	***0.59***	56	-63	8
Nup42	P49686	Cytoplasmic side	1-382	35	29	54	382	***1.54***	13	-2	8
Nup145	P49687	Central channel	1-433	24	11	43	433	***1.79***	32	-23	16
Nup1	P20676	Nuclear side	220-1076	64	17	76	857	***1.19***	96	-72	8
Nup60	P39705	Nuclear side	1-539	23	4	20	539	***0.87***	79	-60	8
Nup159	P40477	Cytoplasmic side	387-1071	45	25	55	685	***1.22***	72	-100	8

3.3.5 Sensitivity analysis on the discretization scheme

To ensure predicted results by the current model are not dependent on the spatial resolution of the bead-spring model (the level of discretization), I performed a sensitivity analysis. Specifically, a coarser discretization level, i.e. 15 AAs/monomer, was conducted and the time evolution of the averaged radii of gyration of FG-repeats, R_g, was examined over $400\mu s$. It appears that the time evolution as well as the average values of R_g exhibit similar patterns in both of the discretization levels (Fig. 3.18), and thus the model is not sensitive to the discretization level, suggesting that the spatial resolution does not bias the conformational behavior of FG-repeats. It must be noted that the discretization cannot go too fine as the model should be coarse enough to assure that the assumption of eliminating atomistic hydrogen bonds between solvent-solvent, solvent-reside, and residue-residue is valid. We thus adhere to a five-AAs/monomer coarse-graining level throughout the presented simulations, which is coarse enough to safely ignore the explicit hydrogen bonds, treating them with the stochastic Langevin treatment, and is fine enough to show the dynamics and distribution of hydrophobic and charged residues.

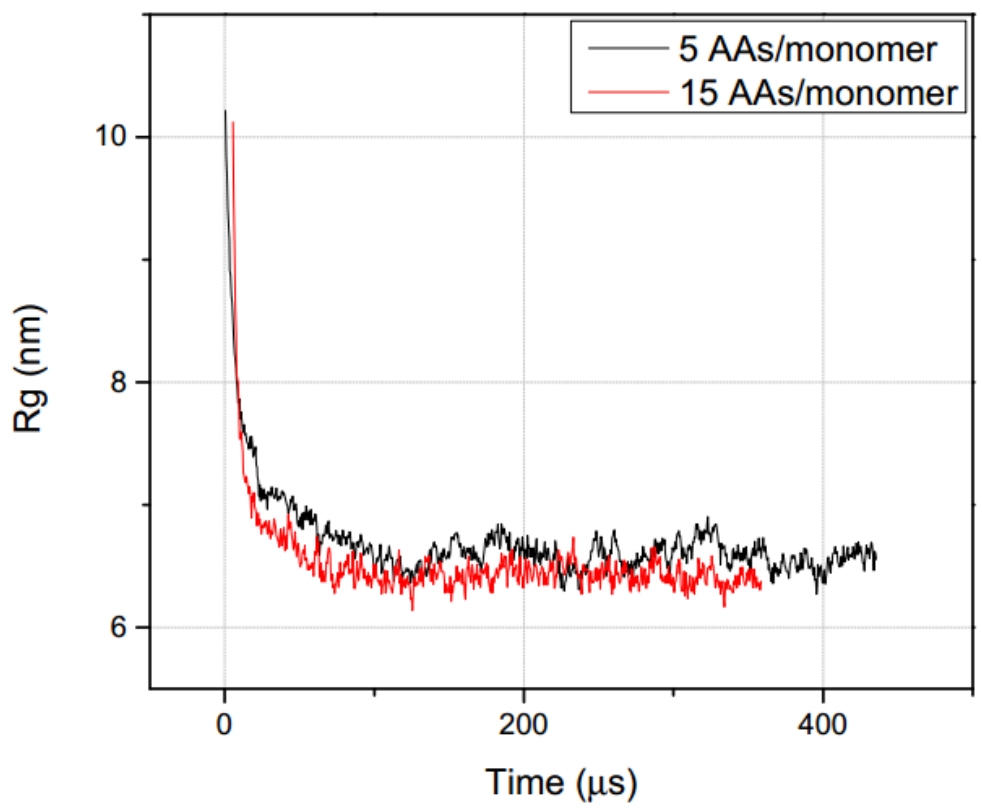


Figure 3.18. The time-evolution of the radius of gyration of FG-Nups in two different discretization schemes: five AAs per monomer, and 15 AAs per monomer (Moussavi-Baygi2016).

3.3.6 Granularity level in a computational model must be consistent with relevant scale-dependent physical phenomena

Although a finer grained model might seem to provide more detailed information of the system, the process of fine graining must be treated with utmost caution, as the physics of the model can be involuntarily challenged. This is because some physical phenomena are scale-dependent, meaning they become prominent at a specific scale while can be neglected at other scale. Particularly, while scaling down to finer levels of granularity, one should be aware of the specific physical properties that become significant at the relevant resolution, and should consider them accordingly in the model.

In case of the FG-repeat domains modeling, for instance, we can safely ignore the hydrogen bonds between residues within the FG-meshwork, so long as the granularity of beads permits. However, if we go to finer granularity of, say, one residue per bead (Ghavami2014, Ghavami2016), the inter-residue hydrogen bonds may not be arbitrarily dropped. At such level of granularity, ignoring hydrogen bonds on the basis that the FG-repeats are "flexible with no secondary structure", and thus, "no long-lasting hydrogen bonds are formed" can be challenged due to, first, the nature of hydrogen bond in sidechains, and second, the biophysical requirements for IDPs.

First, a hydrogen bond is transient and non-lasting by its very nature. At its basic definition, a hydrogen bond occurs as a result of difference in electronegativity between two covalent bonds, one bond having a donor and the other having an acceptor, like N and O atoms, respectively (Kuriyan2012). This does not have anything with the secondary or primary structures of the proteins, i.e. it can happen, as it does, anywhere as long as donors and acceptors, either in the same chain or different chains, are in appropriate proximity of each other. This situation is expected to be frequent in the super-crowded environment of polar and charged sidechains inside the NPC where there are thousands of dangling donors and acceptors. In a finely discretized model of one residue per bead, one cannot ignore the collective effect of such a huge number of transient and dynamic hydrogen bonds.

Secondly, the combined overall high net charge and low hydrophobicity is an important, and universal, property of IDPs (Fig. 3.19), to the extent that it has been suggested as a prerequisite for a given amino acid sequence to evaluate if it forms an IDP (Uversky2011). Indeed, the existence of high net charge, marked polar residues, and low overall hydrophobicity are all known within the FG-repeat domains (Denning2003a, Suntharalingam2003, Colwell2010b). These imply the importance of the hydrogen bonds in the crowded FG-meshwork once one zooms in on a single reside. Nevertheless, hydrogen bonds formed within inter- and intra-IDPs microenvironments are expected to be much more transient and short-lived compared to hydrogen bonds within the folded proteins where the bonds are shielded from the water molecules' attacks. This, however, does not justify ignoring the hydrogen bonds in a fine-grained computational model.

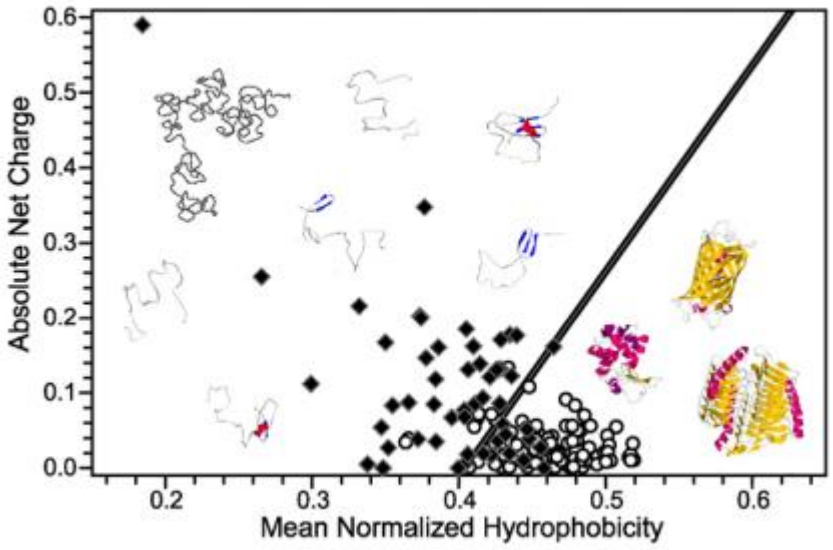


Figure 3. 19. The map of hydrophobicity vs charge for 275 folded proteins (open circles) and 91 IDPs (black diamonds). This map clearly shows the overall high net charge and low hydrophobicity content of the IDPs compared to folded proteins. Adopted from (Uversky2011)

3.3.7 Rejecting movement algorithm in WLC

The finite-extensibility of the WLC imposes a singularity on its force-extension law, such that as the chain approaches the contour length, the axial force goes to infinity. While the finite-extensibility is crucial to be considered in the (bio)polymer simulation, from numerical point of view it will be a drawback. This is because it can produce instability once the distance between two beads approaches the contour length, particularly in the Brownian dynamics as we talked in details in the last sections. This is very likely when explicit time marching scheme is used in the framework as represented in eq. 3.17 (Underhill2004a). To overcome this flaw in the simulation, I use the rejection algorithm proposed by Öttinger (Öttinger1996). In this method, after calculating the new position of each bead, the new end-to-end length is compared with the contour length of the relevant segment, $dL_{c,ij}$, before updating the WLC force of beads. The new force is accepted only if the new end-to-end length is less or equal than 95% of the $dL_{c,ij}$, otherwise the force is corrected. The flowchart is depicted in Fig. 3.20.

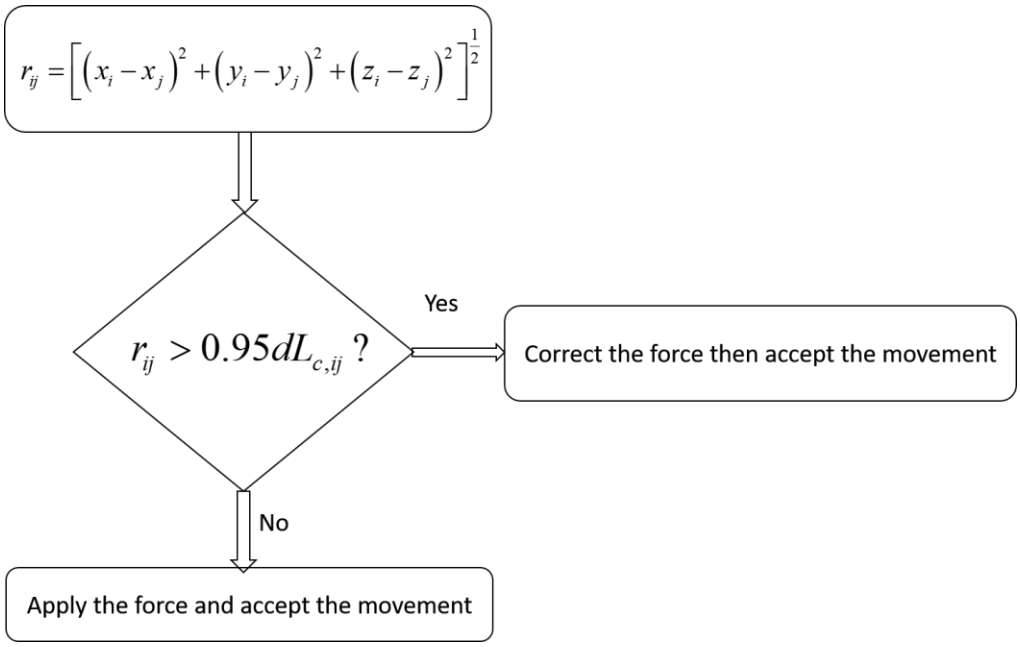


Figure 3.20 The rejection algorithm for discretized WLC spring in the current model.

3.3.8 Sensitivity analysis on the initial configuration of pre-exposed hydrophobic and charged monomers in FG-repeats inside the confined geometry of the central channel

The lengths of FG-repeats localized to the NPC central channel are in the range of 85 to 272 nm (Table 3.3) with an average length of 180 nm. Thus, the initial straight configurations of FG-repeats in our simulations are cramped so that they can be fit in the 40-nm wide central channel (Fig 3.21). This would naturally lead to hydrophobic as well as the charged monomers being pre-exposed.

A legit question is whether the predicted channel-filling conformational behavior of FG-repeats inside the central channel (to be studied in details in the next chapter) is potentially an artifact caused by the pre-exposed hydrophobic as well as charged residues from the same or different FG-Nups.

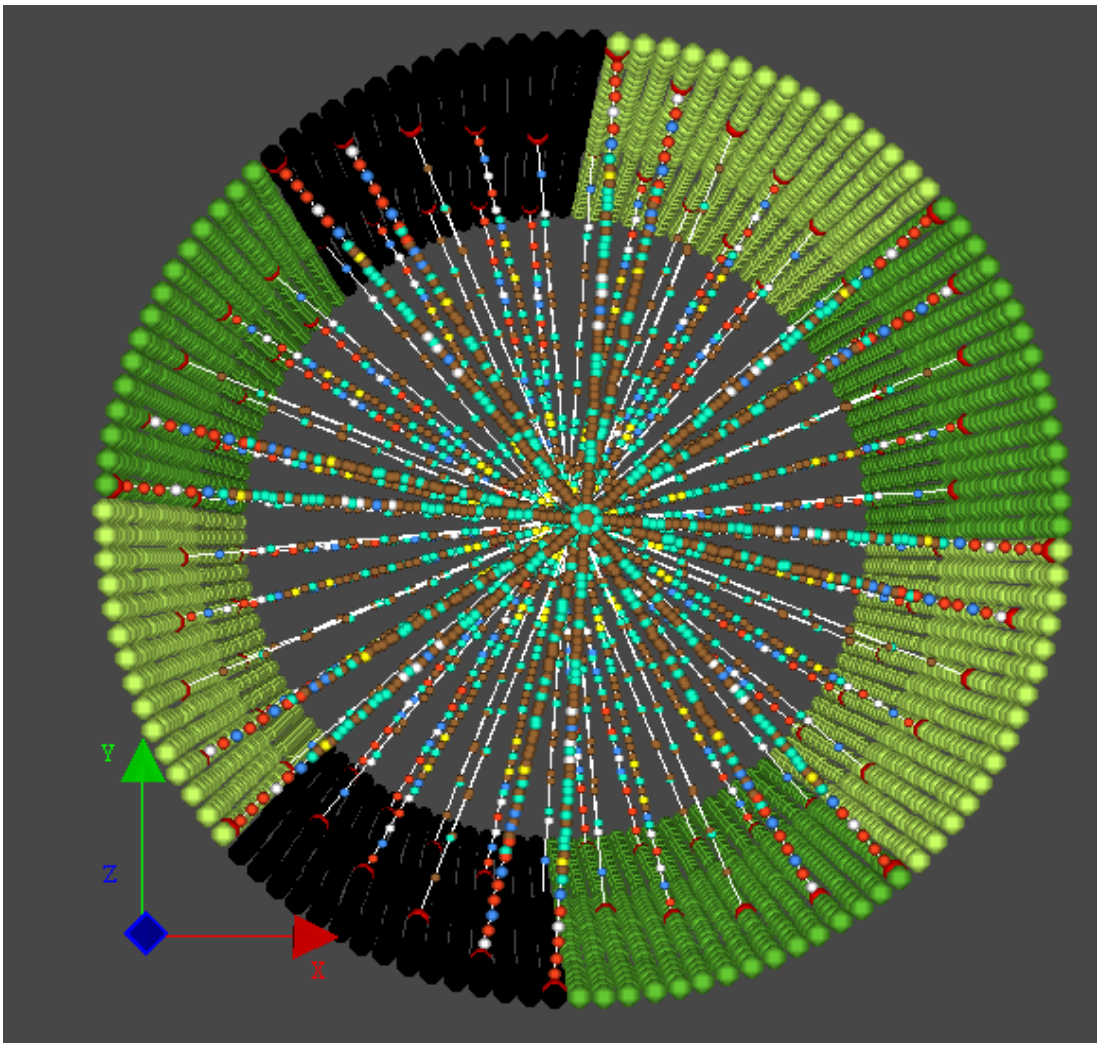


Figure 3.21. The initial configuration of FG-repeat domains it their compacted extended state.

To address this question, I removed the initial pre-exposing by artificially enlarging the channel diameter, and thus, relaxing the crampedness while keeping the positions of tethering points unaltered. The artificial channel had the same height as the central channel (40 *nm*) but its diameter is $550\,nm$ (Fig. 3.22), a bit more than twice the longest disordered FG-Nup, i.e. Nup 100 (Yamada2010). The initial configurations of disordered FG-Nups were set to their fully extended lengths, consistent with previous molecular dynamics studies performed in an open (unconfined) geometry (Miao2009b, Miao2010b), ensuring that hydrophobic as well as charged residues along the same or different FG-Nups were not initially pre-exposed (Fig. 3.22-A).

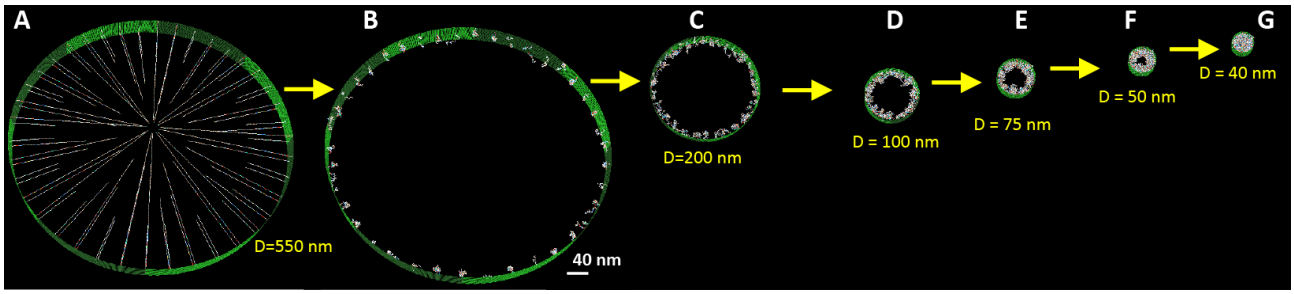


Figure 3.22. Starting from an artificially large channel, the channel diameter is decreased stepwise, and FG-Nups are equilibrated in each step. A) The initial configuration of the artificial central channel with unstructured domains of FG-Nups being in their fully extended length configuration. The diameter of the artificial channel is 550 nm, which guarantees monomers do not meet nor the unstructured domains cross each other. The height of the artificial channel is kept the same as that for the original yNPC central channel. B) The artificial channel after equilibration.
C-G) The diameter of the artificial channel is shrunk stepwise to the final value of 40 nm. In each step, the system is equilibrated for 5 μs before shrinking the diameter to the next step. In the final step, the channel-filling meshwork is reached, as in the original NPC channel.

FG-Nups within the artificial channel were first equilibrated for 5 μs (Fig. 3.22-B), during which they rapidly shrunk and deposited on the channel wall. After that, the channel diameter was shrunk in a stepwise manner to reach its native dimension, i.e. $40\ nm$. In each step, FG-Nups inside the new artificial channel were equilibrated for another 5 μs (Fig. 3.22-C to G). Finally, after the diameter was downsized to $40\ nm$ the FG-meshwork formed and occupied the channel with similar average density and conformations as those in the native channel. This proved that formation of the FG-meshwork inside the native central channel in our model is not an artifact of the pre-exposed hydrophobic or charged residues.

3.3.9 The hydrodynamics interactions can be neglected in the FG-meshwork

A legitimate question in simulating the selectivity barrier inside the NPC and its main components, i.e. FG-repeat domains is about the role of hydrodynamic interactions (HI) inside the pore, particularly given that the HI has a long-range nature (Freed1974). HI is a complex, but classical (De Gennes1976), way of force communication between monomers through a fluidic medium. The hypothesis is that a when a monomer moves in a solvent under the force $\mathbf{f}$, it causes some local disturbance in the form of backflow velocity (De Gennes1976) which varies inversely with distance (Freed1974). This, in turn, is transmitted to the neighboring monomers through the solvent molecules, applying a type of frictional force on them. The transmitted force to the neighboring body is the hydrodynamic force and this process is called hydrodynamic interaction (Chandran2010).

When it comes to the FG-repeat dense polymeric meshwork inside the NPC, however, HI should not be an important and/or determining factor in dynamics and conformational behavior of the system. First, in two pioneering studies, one on ideal chains (Freed1974) and the other in the presence of repulsive force and excluded volume effects (De Gennes1976), authors showed that HI are screened in dense polymer solutions. Their works revealed that the medium effectively

screens out the perturbed fluid field around a monomer in a concentration-dependent manner through decaying exponential factor $e^{-\kappa r}$. Here, κ^{-1} is the concentration-dependent screening length, mathematically resembling Debye screening length in ionic concentration. Thus, in the highly concentrated polymeric meshwork inside the NPC, 1.2 M (Moussavi-Baygi2016), HI are totally screened out.

Secondly, it has been shown more recently that for high aspect ratio polymers, i.e. long and thin rods, the HI are of minor importance and can be neglected in the dynamics of the system (Dhont2003). This notion has been indeed implemented in successful Brownian simulations (Kim2009b, Kim2014). Considering the natively disordered FG-Nups that are long and linear sequences of residues with the average length of 150 nm (in the yeast NPC), the HI can be reasonably ignore.

Thirdly, to elucidate the extent of the HI effects in dynamics and rheology of rod polymers suspension, a comparative study conducted over a wide range of concentrations from dilute to condensed systems (Pryamitsyn2008). It was found that the HI is of secondary importance compared to steric repulsions which accounts for excluded volume effects, and that HI only marginally lead to small corrections in such systems (Pryamitsyn2008).

All in all, we can safely neglect the HI in Brownian simulations of the polymeric meshwork of FG-repeat domains inside the NPC.

3.3.10 The choice of timestep in the Brownian Dynamics

Importantly, in designing a Brownian dynamic simulation, the timestep Δt must satisfy the condition $\Delta t \gg \tau$, or the physics of the system is not correct. This is a consequence of being in the ballistic vs diffusive regimes, as discussed earlier. This subtle point becomes crucially important when one is analyzing the RMSD of time-evolved trajectory at maximum resolution, i.e. Δt. For example, calculation of the frequency-dependent viscosity of a complex fluid (like a FG-meshwork inside the NPC) by particle tracking microrheology is solely based on the RMSD of the tracked particle. At very high frequencies f_{high}, however, it requires inertial considerations (Mason1995) because of the ballistic nature of the motion (see microrheology derivation in the next chapter). Theoretically, the highest frequency one can reach in a simulation is $(\Delta t)^{-1}$, which must be smaller than the f_{high} for the microrheological calculations to be valid. This is because in Brownian dynamics we assume that inertial term is negligible, i.e. we are far from ballistic regime. The choice of $\Delta t \gg \tau$ guarantees that physics of calculations hold true at all time windows accessible to the simulation.

Chapter 4. Results and predictions of the coarse-grained model

In this chapter, I will present the results of the model, predicting different aspects of conformational behavior of FG-repeat domains obtained from the model. The computational microrheological analysis of FG-repeats will be discussed in more details.

4.1 FG-repeats form a nonuniform, porous, and highly dynamic meshwork.

Throughout the simulation, an extremely dynamic, cohesive meshwork of unstructured FG-repeats formed inside the channel, reminiscent of a bowl of 'oily spaghetti' (Macara2001b) that *seemingly* fills the channel. I repeated the simulation 10 times with different initial conditions and verified that the apparent meshwork was repeated, and thus the formation of the meshwork did not depend on the initial conditions.

Next, I explored the extremely dynamic characteristics of this FG-meshwork, arising from the rapid Brownian motion of individual FG-repeats. The total concentration of all AAs within the disordered domains of FG-Nups localized to the NPC channel is about 1.2 M. However, due to the highly dynamic nature of unstructured domains, the concentration is significantly fluctuating (Fig. 4.1). The time-dependent concentration fluctuations in radial direction shows that the fluctuations are largest in the middle and smallest near the wall (Fig. 4.1). This is because of accumulation of free ends along the central axis and more stationary situation near the wall (Giannelis1999).

The time-averaged 2D map of all AAs' radial concentration (Fig. 4.1) exhibits a 6-fold increase in concentration from near the wall (0.4 M) to the middle (2.4 M). This is a signature of formation of a nonuniform meshwork with a dense zone in the middle that weakens toward a low-density zone near the wall. While the low-density zone resembles a sparse polymeric shell in the wall vicinity, the dense zone features a rod-like high density region that is about 22-*nm* thick and extends axially along the channel (Figs. 4.1 and 4.13).

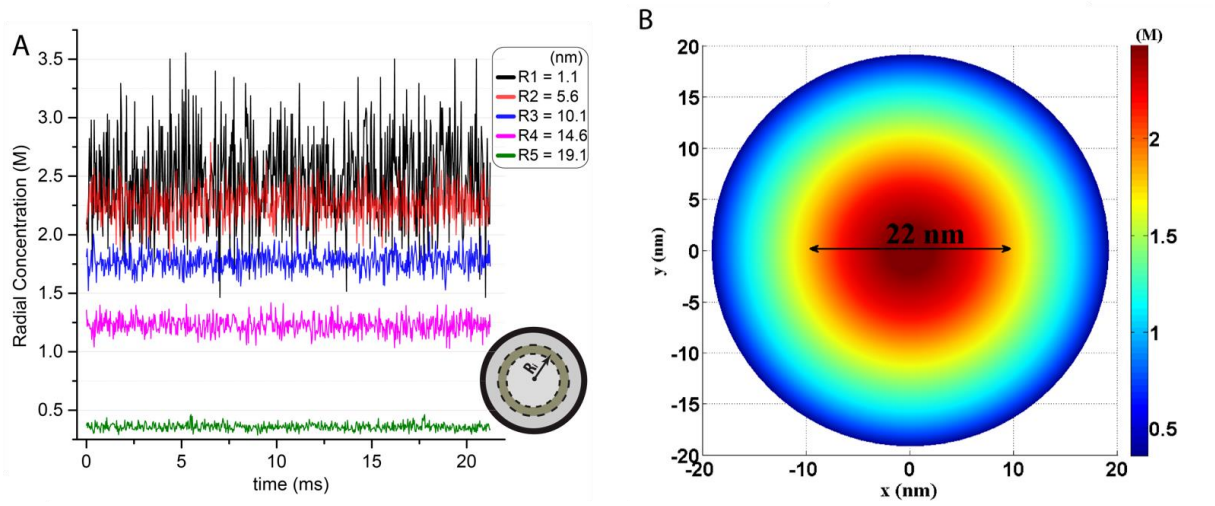


Figure 4.1. Radial concentration of FG-repeats inside the NPC channel. A) Time-dependent concentration fluctuations in radial direction at five different radii inside the channel. At each R_i, the concentration is calculated within a cylindrical shell (the ring inside the yellow circle) with a width of the monomer's vdW diameter. B) Time-averaged 2D map of radial concentration of all AAs within the FG-repeat domains. The concentrations are averaged along the channel main axis over 20 *ms*. The diameter of the rod-like high-density zone is about 22 nm.

Notably, the average density of the central zone decreases as the diameter of the channel increases, in agreement with the behavior of single-type FG-Nups grafted inside the solid-state nanopores (Kowalczyk2011). At higher diameters of the pore, the central zone disappears, and instead, the radial concentration spreads more evenly within the channel (see Fig. 4.2). This suggests that the channel diameter affects the formation of different zones within the FG-meshwork confined to the channel interior. However, for the rest of this study we consider the channel diameter of the yeast NPC that is 40 nm (Aitchison2012a).

To further analyze the rod-like high density zone I determined its constituent residues and the type of dominant interactions therein. The time-averaged 2D map of the single-type AA radial concentration shows a relatively large number of Phe—the most hydrophobic residue (Bayliss2002a)—in the rod-like zone with a concentration of about 1 M (Fig. 4.3). This suggests that the most hydrophobic region of the FG-meshwork is along the rod-like zone.

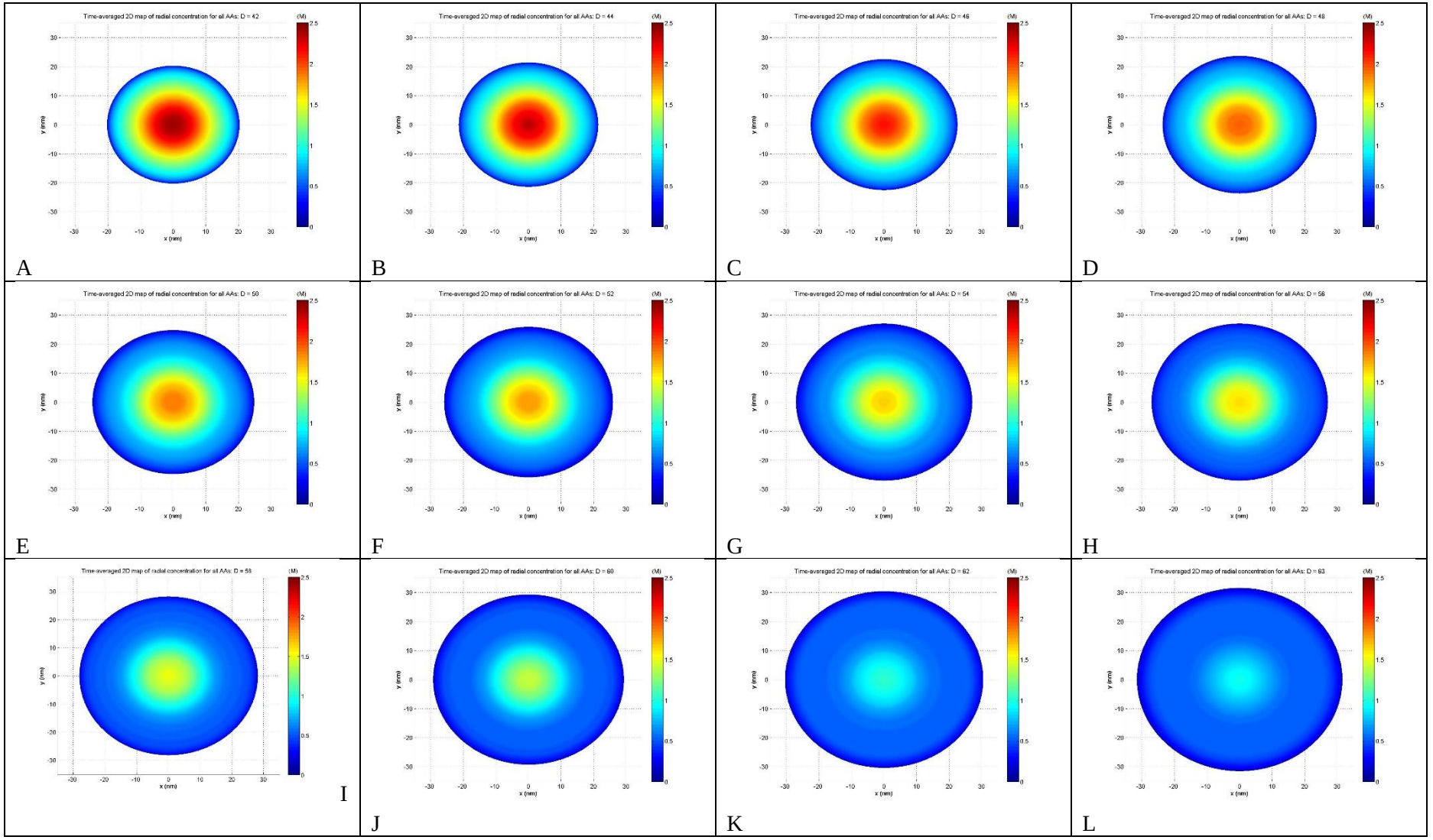


Figure 4.2. The dependence of the intra-FG-meshwork zone formation on the diameter of the channel. The channel diameter is increased from 42 nm in panel A to 64 nm in panel L. From each panel to the next, the increment in the channel diameter is 2 nm. Color bars in all panels have the same limits (0 to 2.5 M), and x and y axes' limits are -30 to 30 nm

Nevertheless, since the concentration map is a 2D radial projection, it does not convey information about dynamics and the accurate locations of hydrophobic interactions. Thus, we dissected the time-evolution of hydrophobic interactions within the FG-meshwork through the time-dependent 3D spatial density map.

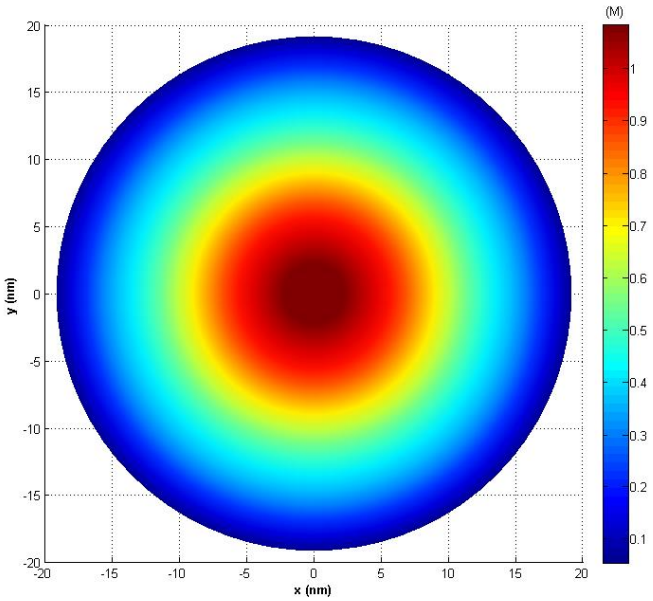


Figure 4.3. Time-averaged 2D map of radial concentration of all Phe residues within the FG-repeat domains. The concentrations are averaged along the channel main axis.

The 3D map of hydrophobic interactions affirmed that the majority of those interactions occur in the central part, both radially and axially, forming a dynamic hydrophobic core (Fig. 4.4)

Next, I examined the porosity of the FG-meshwork using the monomer-monomer pair distribution function (PDF) (Allen1989), with the pair of monomers belonging to two distinct FG-Nups. The location of the first peak in the PDF plot represents the most probable pairwise distance between monomers (Allen1989), and thus is a measure of the meshwork porosity. It was found that the dominant size of passive pores within the FG-meshwork is about 6 nm (Fig. 4.5-A), setting an upper limit to the free diffusion of small particles across the NPC. Yet, because the meshwork is spatially heterogeneous, we performed the zonal PDF within the high- and low-density zones separately, and determined that the peak location shifts to 4 nm in the high-density zone (Fig. 4.5-B), implying the porosity inside the rod-like zone is tighter compared to the entire meshwork. The PDF plot for the low-density zone, however, is very similar to the entire FG-meshwork (Fig. 4.5-C), suggesting that the porosity of the entire FG-meshwork is mainly dictated by this zone.

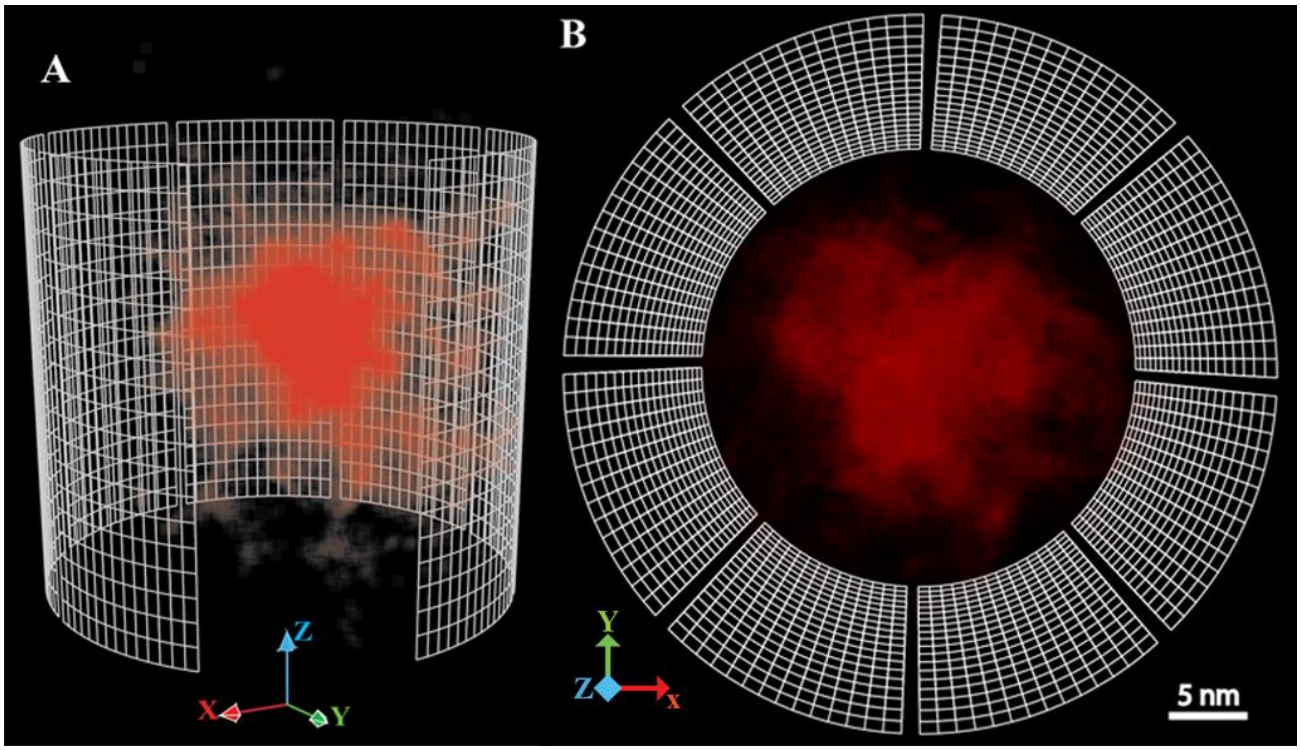
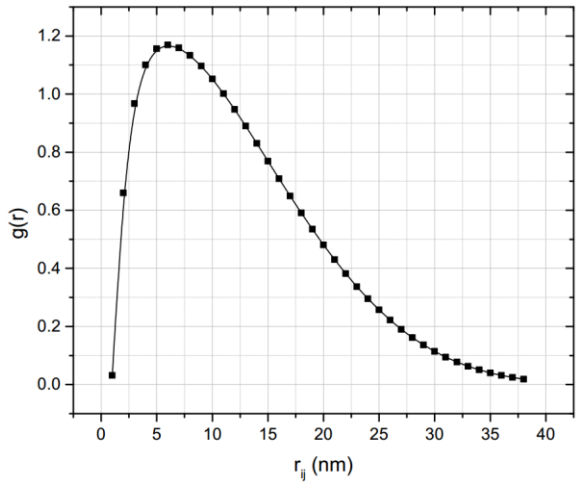
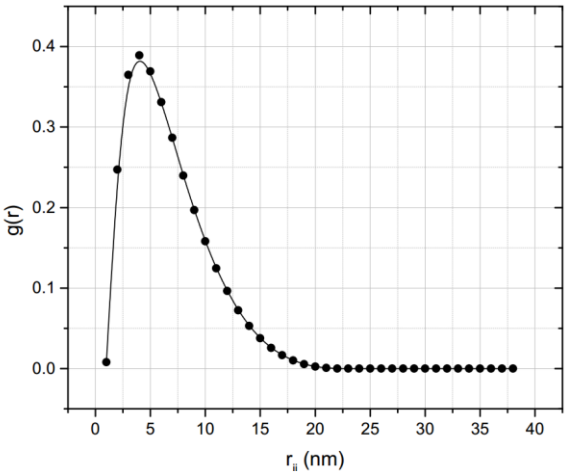


Figure 4.4 The 3D spatial density map of inter- and intra-FG-repeats hydrophobic interactions inside the channel. We pinpointed every single hydrophobic interaction and recorded its coordinates along with the strength of interaction throughout the simulation. Each dim point shows a single hydrophobic interaction, and thus, the color intensity is proportional to the number of hydrophobic interactions. The channel is represented as a wired frame for clarity. A) side view and B) top view snapshots.





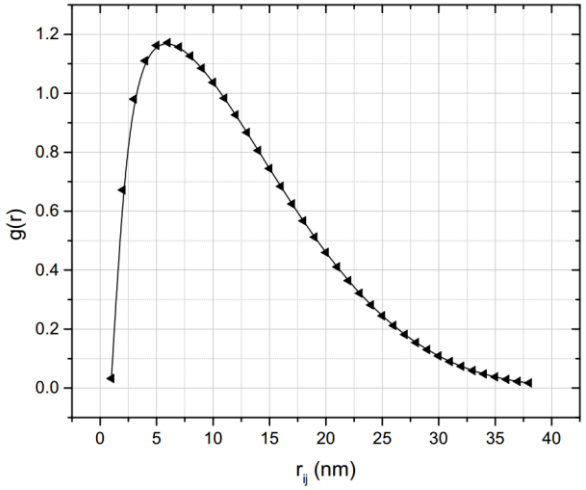


Figure 4.5. The pair distribution function (PDF) is used to calculate the frequencies and distances of particles' pairwise separations inside a fluid with a given structure and density relative to a fluid with a completely random structure at the same density. Thus, it can provide profound insights into the microstructure of the FG-meshwork in the current study. The monomer-monomer PDF plot of the FG-meshwork for A) the entire meshwork (rectangles), B) high-density zone in the middle (circles), and C) low-density zone near the wall (triangles).

4.2 Dynamicity underlies concentration fluctuations, facilitating the NTR-bearing macromolecule transport.

The model predicts that the end-tethered FG-repeats formed a porous and extremely dynamic meshwork that constantly jiggles and wiggles, showing a high degree of concentration fluctuations inside the NPC channel (Fig. 4.1). Indeed, highly dynamic nature of free IDPs and their erratic motions on time scales of sub-nanoseconds is reported elsewhere (Uversky2008). Here, the model suggests that IDPs retain their highly dynamic characteristics, even when they live in a dense cohesive meshwork and are tethered to, and confined in, a compact geometry.

The rapid Brownian motion, and thus the concentration fluctuations within the FG-meshwork, are conceivably among the major underlying factors in efficiently conducting transport. This is achieved by augmenting the back-and-forth motions of the NTR-bearing macromolecule that is engaged in the FG-meshwork through hydrophobic, and to a lesser extent, electrostatic interactions (see following). Analogously, but in a different context, it has been suggested that concentration

fluctuations of the actin filaments drive lamellipodia protrusion and retraction, which in turn provokes the cell crawling mechanism (Ryan2012).

The concentration fluctuations within the FG-meshwork are reminiscent of the cyclic motions of self-oscillating gels induced by an oscillatory chemical reaction called the Belousov-Zhabotinsky (BZ) reaction (Maeda2008). In those systems, periodic chemical energy of the BZ reaction is converted to mechanical oscillations within the polymeric meshwork. Indeed, fluctuations in self-oscillating hydrogels has been harnessed for mass transport and cargo delivery purposes (Murase2008, Shinohara2008). However, there is a fundamental difference between those systems and the NPC in that no chemical reaction occurs inside the FG-meshwork, nor is there any external source of energy to wriggle the FG-meshwork. Instead, we propose that the thermal noise, spreading through the geometrically confined NPC channel that hosts numerous transient, individually weak hydrophobic and electrostatic bonds, along with the delicate structures of *tethered* FG-repeats, *en masse* produce incessant rapid Brownian motion, leading to continuous concentration fluctuations in the NPC channel.

4.3 FG-repeat domains percolate inside the native NPC

Percolation is a mathematical concept, describing the behavior of connected clusters in a random graph. On a lattice-based model, there is a probability attached to each site being occupied to make a bond with the neighboring site. As the bond formation probability increases, at some critical probability, a path will emerge, connecting two opposite sides of the box. Fig. 4.6 shows the progressive increase in the bond formation probability and the emergence of percolation as a connecting path at the critical probability of 0.59 (Christensen2002).

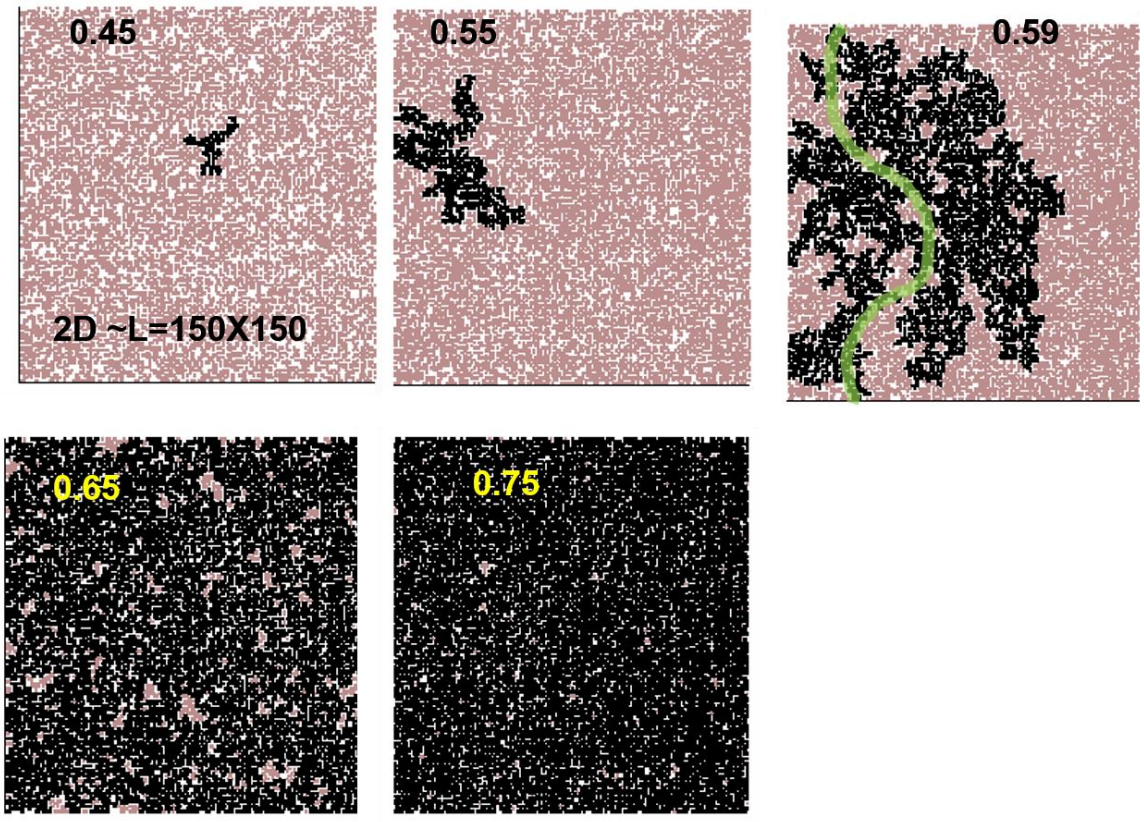


Figure 4.6. A 150 by 150 lattice model start percolating at the critical probability of 0.59 (Christensen2002). The bond formation probability is indicated on each panel.

This concept was introduced in polymer physics by de Gennes (De Gennes1976, de Gennes1979a) and Stauffer (Stauffer1976) as a geometrical criterion for gelation. Thus, percolation is often referred to as the geometrical phase transition. As indicated in Fig. 4.7, before gelation, there are sporadic clusters in the lattice, while at a critical point, those clusters connect together to form a percolating meshwork. From geometrical viewpoint, this is the gel point (Hess1988). The network continues to grow and fill the lattice during the post gel. The definition of the critical point of percolation in the gelation process is somewhat arbitrary. Depending on the system, it can be a specific concertation, temperature, pH, molecular weight of polymers, length of polymers, or any combination at which the percolation starts.

Nevertheless, later studies on the gelation of colloidal solution suggested that percolation is not equivalent to gelation in physically associated polymers (Grant1993, Bergenholtz1999). It was argued that if the pairing energy between associating monomers along polymers is much higher than the thermal noise, $\varepsilon_{ss} \gg k_B T$, the clusters last effectively for a very long time, and thus the network shows a solid-like behavior, i.e. a gel is formed and maintained. In these systems, the percolation is equivalent to gelation. However, for polymeric systems with weak physical

associations where $\varepsilon_{ss} \sim k_B T$, the associating monomers pair and unpair rapidly. Here, the timescale of the cluster formation is the same of the polymers' diffusional displacement. Thus, the percolation does not guarantee a long-lasting solid-like behavior (Kumar2001). On this basis, the geometrical percolation might be considered a necessary, but not sufficient, condition for gelation (Guo2005).

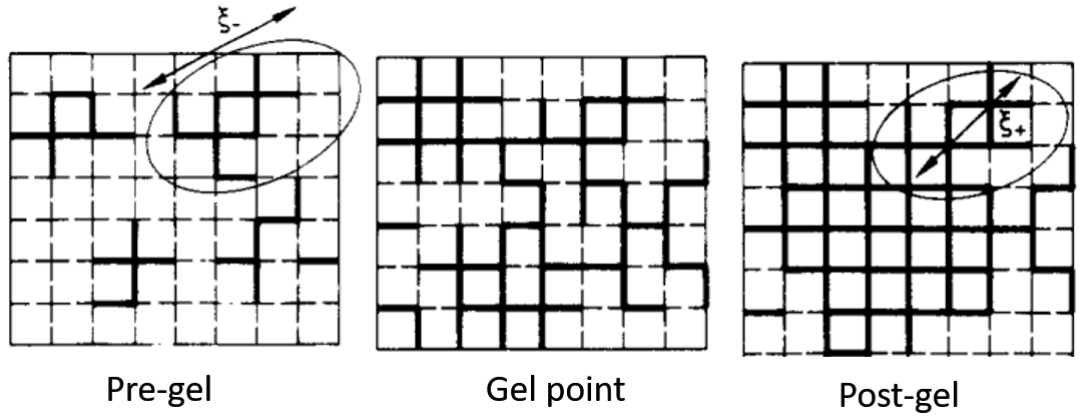


Fig. 4.7. The pre-gel, gel, and post-gel points from geometrical point of view. Adopted from (Hess1988).

To determine whether FG-repeats satisfy the geometrical conditions of gelation, I analyzed the connectivity profile of the meshwork, using the geometrical percolation theory (Larson1999) as adopted in polymer physics (De Gennes1979b). It was found that the FG-meshwork percolates inside the channel both in radial and axial directions (Fig. 4.8), implying that the FG-meshwork exhibits a geometrical connectivity, similar to a gel. This further implies the physiological concentration of FG-repeats is sufficient to form a connecting meshwork inside the NPC, but does not necessarily denote that the channel is *uniformly* filled up as implied by the hydrogel model (Frey2007).

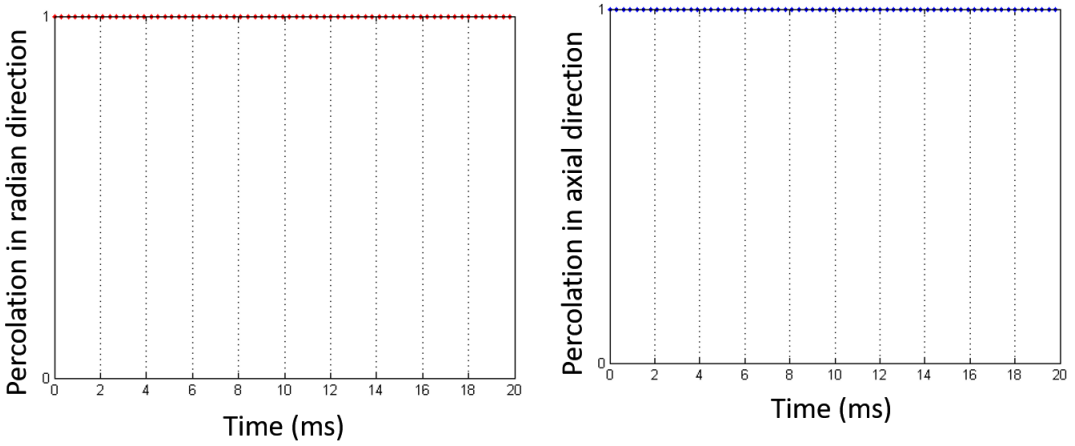


Figure 4.8 Percolation of the FG-meshwork in the radial and the axial directions within the central channel during the simulation. Vertical axis shows whether percolation occurs (=1) or not (=0) for a particular configuration corresponding to a certain simulation time. Adding up those numbers, we can calculate the probability of percolation for the total simulation time of 20 ms. The probability of percolation is 100% in radial as well as the axial directions, meaning the FG-meshwork always forms a connected pathway in both directions. A) Radial direction, and B) axial direction.

4.4 The FG-meshwork locally collapses upon interacting with the NTR-bearing macromolecule, but autonomously reconstructs itself between sub-microsecond and $\leq 7\mu s$, depending on the macromolecular shape and size.

For efficient nucleocytoplasmic transport cycles, the NPC channel must constantly accommodate macromolecules of different sizes. If the channel is sealed by a dense FG-meshwork, then the challenging questions would be: Once a NTR-bearing macromolecule perforates and passes through, how fast does the FG-meshwork reconstruct afterward? What is the timescale of reconstruction process in comparison with the milliseconds-long transport time? Does not the constant flow of macromolecules and/or NTR leave the channel in an unsealed state and destroy the permeability barrier?

To explore these questions, I simulated the globular and elongated NTR-bearing macromolecules, i.e. active cargo, by placing the macromolecule at the NPC cytoplasmic entry, with the NTR hydrophobic binding spots closely faced against the FG-repeats. The macromolecule was then nudged toward the channel by a small force, on the order of magnitude of the thermal noise (see Table 4.1). The nudging force was applied only to save computational time by guiding the macromolecule to the right direction and gracefully steering it across the FG-meshwork, instead of wandering around and after a long time finding the correct position. Thus, the force does not have a direct physiological equivalence. Yet, it is too small that does not obscure the native transport of the macromolecule.

Once the NTR-bearing macromolecule reached the mid-channel, it was kept fixed there for 10 μs so that the polymeric FG-repeats around the cargo surface were equilibrated. Then, the cargo was abruptly removed, leaving a cavity inside the meshwork. Subsequently, we measured the time it

took for the FG-meshwork to refill the cavity and adopt its pre-cavity conformation—the reconstruction process.

It appears that the reconstruction process follows a saturating pattern (Fig. 4.9), for which we suggest the following equation:

$$\rho^*(t) = 1 - e^{-\frac{t}{\tau}} , \qquad (4.1)$$

where $\rho^*(t)$ is the time-dependent dimensionless density inside the cavity and equals $\frac{\rho(t)}{\rho_0}$, with $\rho(t)$ being the time-dependent local density and ρ_0 the pre-cavity value of the local density. More importantly, τ, the characteristic time of reconstruction, is the time for the cavity to regain about 2/3 of its initial density, ρ_0.

Significantly, the reconstruction pattern is biphasic for both the globular and the elongated macromolecules; initially it proceeds very rapidly and then gradually reaches a saturating phase (Fig. 4.9). The characteristic time τ is the determining factor for the speed of reconstruction. To find out how τ varies with the cavity size, we studied different sizes of NTR-bearing macromolecules (Fig. 4.10 and Table 4.1). Furthermore, to draw conclusion about the shape effects, volumes and surface characteristics of globular and elongated macromolecules were kept the same in a pairwise manner, so that for every globular macromolecule, an elongated counterpart of equal volume and the same surface characteristics was studied (Table 4.1). For each macromolecule with a particular size and shape, we performed 30 independent simulations, from which the mean of τ was computed.

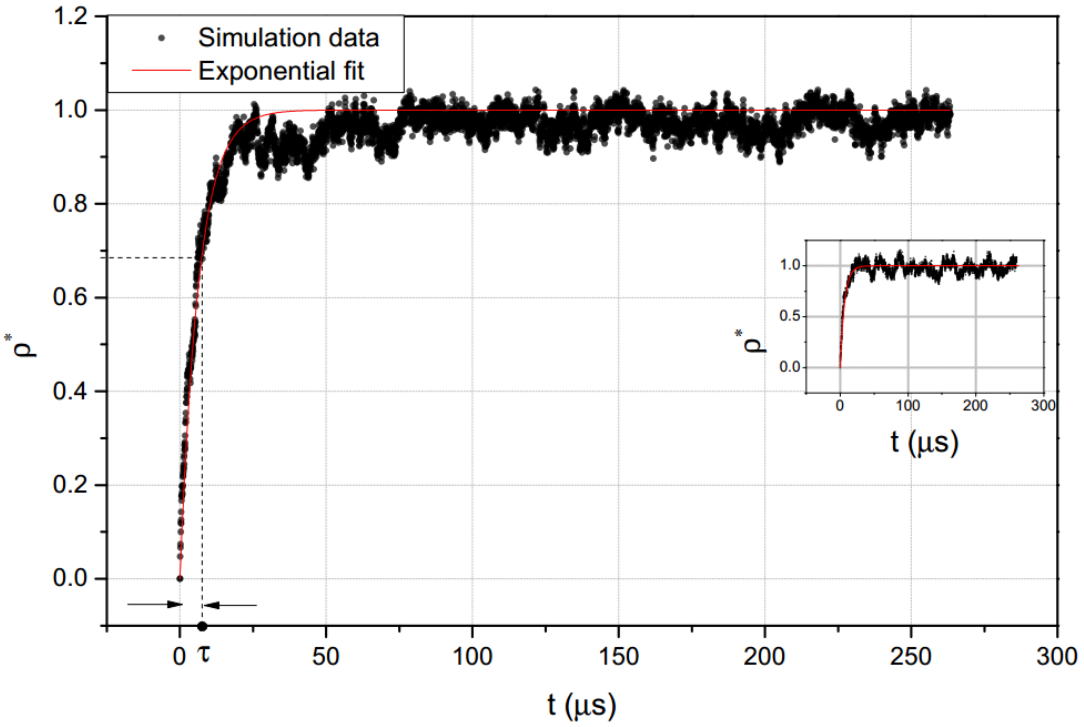


Figure 4.9 Reconstruction pattern of the FG-meshwork after a 20-nm globular macromolecule passes through. Black dots are simulation data and red line shows the fitted eq. (4.1). Inset: the reconstruction pattern of the FG-meshwork for an equivalent elongated macromolecule (with the diameter of 11.6 nm and the length of 40 nm).

Generally, τ increases with the cavity size (Fig. 4.10), indicating there is a direct relation between them. For different sizes of macromolecules, both globular and elongated, we determined that τ varies between half a microsecond and about $7\mu s$ (Fig. 4.10 and Table 4.1). Particularly, for globular cargoes beyond 20 *nm*, τ increases more rapidly, i.e., relatively it takes longer for the FG-meshwork to reconstruct. This is conceivably due to extensive conformational changes in the FG-meshwork as a result of the large cavity size ≥ 20 *nm*.

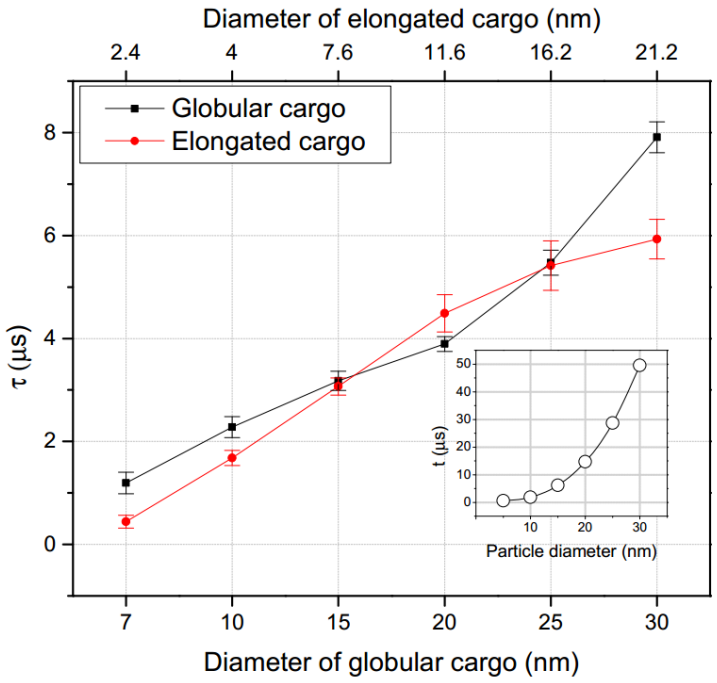


Figure 4.10 Distribution of the reconstruction characteristic time for different the cavity size of globular and elongated NTR-bearing macromolecules. Inset: The time of local diffusional motion versus the particle diameter for a globular particle diffusing in the cytoplasmic viscosity.

Notably, the reconstruction process here resembles the 'self-healing' feature proposed in other studies that has been attributed either to inter-FG hydrophobic bonds, or to continuous presence of the NTR population (Frey2009, Schoch2012). We thus sought to identify the potential role of hydrophobic interactions in reconstruction by mutating all of the Phe residues to polar residues, having the same vdW dimensions. Subsequently, after equilibration of the mutated meshwork, the macromolecule was forced into and kept fixed in the mid-channel for 10 μs, followed by an abrupt removal. It is found that the reconstruction process repeated in the mutated meshwork, following the same exponential pattern (Fig. 4.12), but with a smaller τ, i.e., reconstruction occurs faster than in the native FG-meshwork. For instance, for a globular cavity of 20-nm in the mutated meshwork, $\tau = 1.38 \pm 0.25 \mu s$; this is 65% faster than that in the native FG-meshwork (Fig 4.5, and Table 4.1).

I then looked into how individual FG-repeats conformationally behave upon interacting with the NTR-bearing macromolecule (Lim2007b). To this end, a ring of 15-*nm* width in the middle of the channel was chosen and monitored the average height of the brush created by the FG-repeats tethered to the ring. As the NTR-bearing macromolecule approaches the ring, FG-repeats show a continuous decrease in their heights, driving them to adopt a collapsed conformation (Fig. 4.11). As soon as the macromolecule disappears, FG-repeats start regaining their heights, indicating the collapse is promptly reversible.

Table 4.1. The characteristic times along with the properties of different NTR-bearing globular and elongated macromolecules used in the model.

Globular macromolecule				**Elongated macromolecule**				
Diameter (nm)	**τ (μs) ± SE**	**Number of hydrophobic binding spots**	**Nudging force (pN)**	**Diameter (nm)**	**Aspect ratio (h/D)**	**τ (μs) ± SE**	**Number of hydrophobic binding spots**	**Nudging force (pN)**
7	1.20 ± 0.21	10	0.5	2.4	20.83	0.44 ± 0.12	7	0.25
10	2.28 ± 0.20	22	1.0	4.0	12.5	1.68 ± 0.15	15	0.5
15	3.18 ± 0.19	51	2.25	7.6	6.58	3.07 ± 0.17	34	1.25
20	3.89 ± 0.14	91	4.0	11.6	4.31	4.49 ± 0.36	61	2.0
25	5.47 ± 0.24	120	6.25	16.2	3.09	5.42 ± 0.48	80	3.25
30	7.91 ± 0.30	252	9.0	21.2	2.36	5.93 ± 0.38	140	4.5

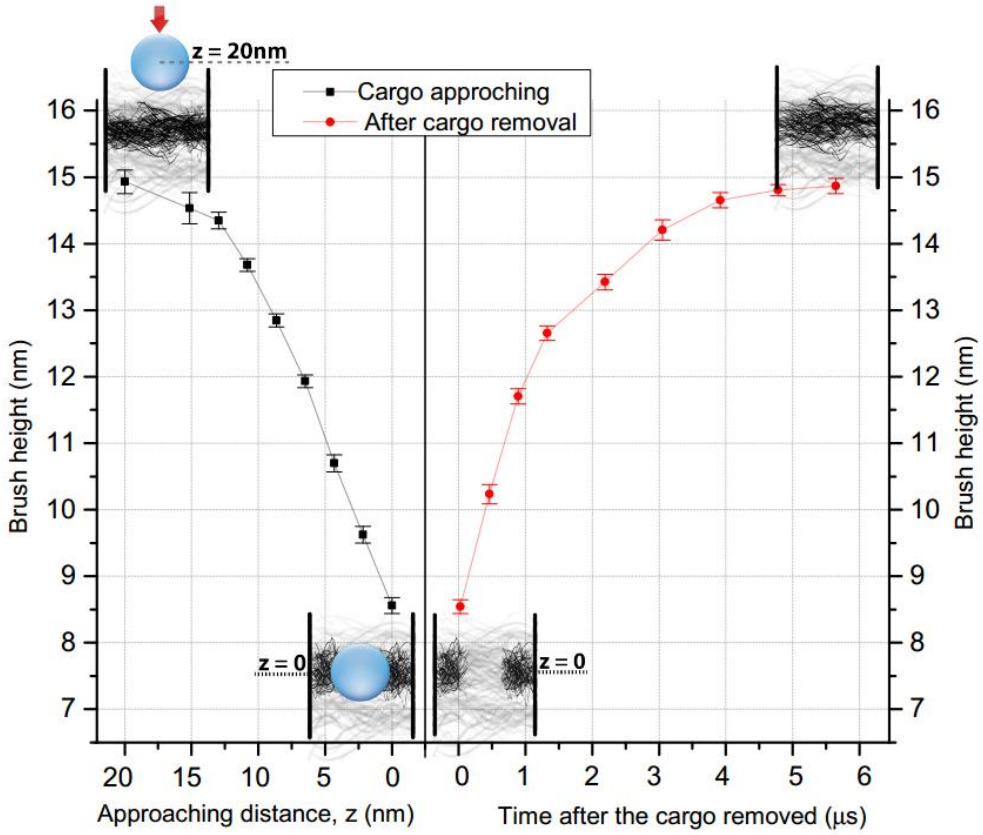


Figure 4.11 The rapid and reversible collapse of individual FG-repeats in the mid-channel as the NTR-bearing macromolecule approaches from the top. Left graph shows the brush height as a function of approaching distance. Right graph represents after the macromolecule is removed how fast FG-repeats regain their heights as a function of time.

4.5 Ultrafast reconstruction of the FG-meshwork arises solely from rapid Brownian motion of FG-repeats, keeping the NPC channel in a perpetually-sealed state.

In this work, I quantified the reconstruction pattern and time for the FG-meshwork in detail and proposed a time-dependent relation for this process (Eq. 4.1), which is markedly biphasic with a rapid and a slow phase (Fig. 4.9). The reconstruction occurs mainly during the first phase, which is entropically driven; a cavity within a dense meshwork is entropically highly unfavorable, and thus, once the macromolecule passes through, the meshwork rearranges itself rapidly. Once the configurational entropic cost is compensated, the density fluctuations within the cavity, $\rho(t)$, continues more slowly with lots of 'small-amplitude' peaks and valleys during the saturating phase (Fig. 4.9). Imaginably, this phase is primarily due to favorable inter- and intra-FG-meshwork hydrophobic as well as electrostatic interactions.

Our results revealed that the characteristic time of reconstruction, τ, lies anywhere between $0.44 \pm 0.12 \mu s$ and $7.91 \pm 0.30 \mu s$, depending on the macromolecule size and shape (Fig. 4.10 and Table 4.1). Given the transport time of a single macromolecule is anywhere between 3 *ms* to several seconds (Yang2004, Kubitscheck2005, Yang2006, Dange2008a, Mor2010, Moussavi-Baygi2011b, Moussavi-Baygi2011a, Siebrasse2012, Ma2013, Goryaynov2014), the reconstruction occurs three to seven orders of magnitude faster than the actual transport. This indicates that the reconstruction process is ultrafast and 'instantaneous' compared to the timescale of the entire transport process.

More precisely, τ is even shorter than the time of *local* diffusional motion of the macromolecule. Here, by *local* we mean the time it takes for a particle to diffuse a distance equal to its characteristic length, *L*. For example, for a globular cargo of 20 *nm* in diameter it takes about $\langle t \rangle = \langle L^2 \rangle / 6D \cong 15.0 \mu s$ to diffuse its diameter in cytoplasmic viscosity (Gorlich2003). Since the cargo diffusion inside the NPC channel is slower than that within the cytoplasm (Ma2013), $15.0 \mu s$ is the lower bound on the *local* diffusional time of a 20-*nm* cargo inside the NPC. Yet, τ for the cavity created by the same cargo is $3.89 \pm 0.14 \mu s$ ($\pm SEM$) (Table 4.1), meaning that the reconstruction process is at least fourfold faster than the *local* diffusional motion. This implies that the cargo does not leave any void behind itself, suggesting the NPC channel is perpetually sealed so that the traffic of macromolecules does not lead to breaking the permeability barrier or 'leaking'.

The key factor here is the existence of different timescales involved in the game: the FG-meshwork reconstruction occurs three to seven orders of magnitude faster than the transport process. Moreover, the reconstruction characteristic time is at most 1/4 of the cargo *local* diffusional time, i.e. the reconstruction is accomplished while the NTR-bearing macromolecule is still around. Consistently, based on the difference in influxes of nonspecific (inert) versus NTR-bearing cargoes through an *in vitro* hydrogel obtained from saturated FG/FxFG repeats of the Nsp1, it was shown that the 'lesions' created in the hydrogel were short-lived, and thus, the hydrogel 'immediately' resealed behind particle translocations (Frey2009). Nonetheless, to the best of our knowledge, the mechanistic details of reconstruction have never been investigated thoroughly, regardless of which conformational behavior FG-repeats adopt.

The current model also shows this ultrafast reconstruction, reminiscent of 'self-healing', is merely due to the highly dynamic characteristics of the FG-meshwork, and not due to the inter- or intra-FG-Nup hydrophobic interactions, nor the presence of NTRs. In a study by Görlich and co-worker, based on a single-type FG-Nup (Nup49p or Nup57p) synthetic hydrogel at macroscopic scale *in vitro*, it was suggested that the 'self-healing' stems from inter-FG-repeats hydrophobic cross-linking (Frey2009). More recently, however, from the basis of another single-type FG-Nup (hNup62) grafted on a plane (unconfined) geometry, Lim and co-workers argued that the 'self-healing' was due to the continuous presence of kap β population and was triggered by the kap β-FG-Nup binding avidity (Schoch2012).

Obviously, in the current study neither NTR nor NTR-bearing macromolecules are continuously present inside the NPC channel. Particularly, after the NTR-bearing macromolecule is abruptly removed, it is solely the FG-meshwork on its own that autonomously fills the cavity and reconstructs itself with no aid from the NTR or NTR-bearing cargo. Therefore, our model does not support the notion that the 'self-healing' is conferred by the NTR or any other external agent as

proposed by Schoch et al. (Schoch2012). Likewise, the 'self-healing' character does not stem from inter- or intra-FG-repeats hydrophobic interactions, as the mutated meshwork, lacking hydrophobic interactions, reconstructs itself even faster than the native FG-meshwork (Fig. 4.12). For example, for a 20-*nm* globular cavity in the mutated meshwork, τ is $1.38 \pm 0.25 \mu s$ $(\pm SEM)$ (Fig. 4.12); about three times smaller than that in the native FG-meshwork (Table 4.1). This is because the mutated meshwork is non-cohesive and lacks the 'stickiness', and thus, entropic repulsions are not counterbalanced by favorable hydrophobic interactions. Incidentally, this also explains why the amplitudes of peaks and valleys during the second phase of reconstruction in the mutated meshwork are larger compared to the native FG-meshwork (compare Fig. 4.9 and Fig. 4.12). Therefore, the molecular basis of the reconstruction process is not due to the inter- or intra-FG-repeats hydrophobic interactions. Instead, it solely arises from the rapid Brownian motion of the confined and end-tethered FG-repeats, manifested in the dynamicity of the FG-meshwork.

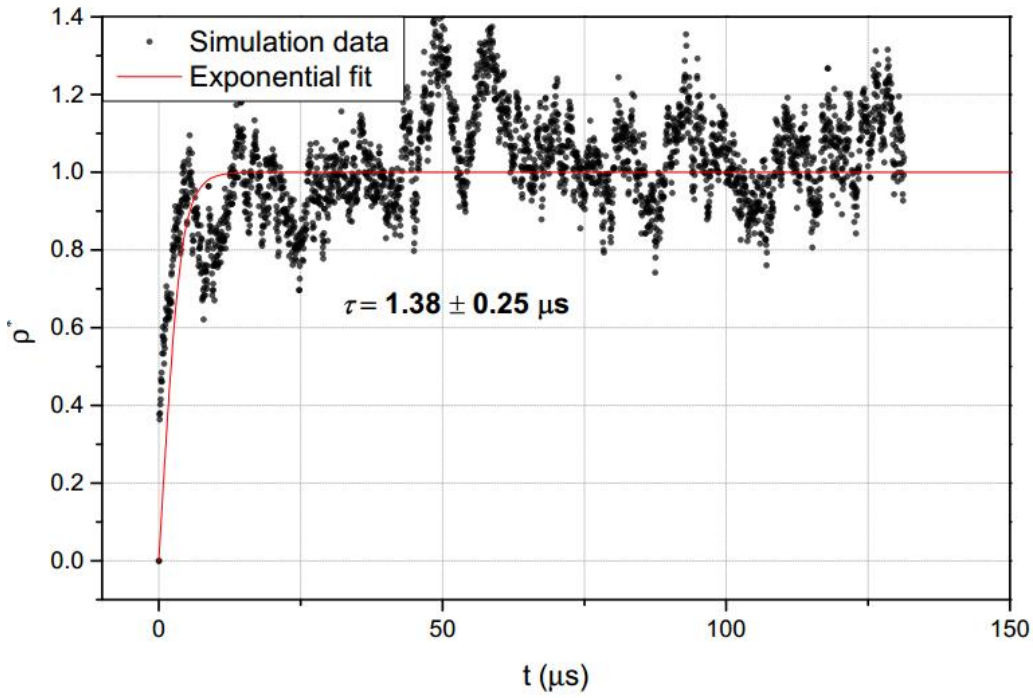


Figure 4.12. The reconstruction patter of the non-FG-meshwork, where all of the Phe residues are replaced by polar residues with the same vdW dimension. The cargo in this case is a 20-nm NTR-bearing globular macromolecule.

Although we do not capture mechanical characteristics of the FG-meshwork in the current study, it has been shown that the Phe→Ser mutation in the Nsp1 synthetic hydrogel changes mechanical properties and converts the semi-solid gel into a more fluidic aggregation (Frey2006b). However, a generalization of this observation to the interior of the NPC channel is questionable, as four

major contributing factors in determining the FG-meshwork conformation were absent in that study. These factors are heterogeneity, end-tethering, confinement, and nano-scale size effects of the channel.

Our model also suggests that *individual* FG-repeats undergo *rapid* collapse upon interacting with NTR-bearing macromolecules (Fig. 4.11). The collapse is rapidly reversible, i.e., the moment the NTR-bearing macromolecule dissociated from the FG-Nup, disordered domains stretch away from the wall rapidly due to the entropic restoring force. Indeed, plane-grafted disordered domains of Nup153 was shown to undergo a reversible and rapid transition from polymer brush to collapsed state when they interact with kap β (Lim2007b). Here, however, we extend that result to the interior of the NPC, with the heterogeneous mixture of all FG-Nups confined to and grafted in the channel. This agrees with a recent theoretical and computational study where the self-consistent field theory as portrayed by the strong-stretching theory was employed to investigate a polymer-brush-based nanovalve (Coalson2015). Our model suggests that the *rapid* collapse of individual FG-domains is truly an essential step for nucleocytoplasmic transport, because in its absence the NTR-bearing macromolecule cannot quickly and efficiently penetrate into the FG-meshwork (see following).

4.6 The cargo shape effects

FG-meshwork generally exhibits a mildly faster reconstruction for an elongated cavity than a globular cavity having the same volume (Fig. 4.10 and Table 4.1), suggesting elongated cargoes tend to maintain the sealing state of the FG-meshwork more efficiently than their globular counterparts. Notably, compared to a globular-shaped macromolecule, an elongated counterpart also requires about 35% fewer number of hydrophobic binding spots, and about 50% smaller nudging force to penetrate into the FG-meshwork (Table 4.1), implying that the cargo shape matters in nucleocytoplasmic transport. To further investigate the cargo shape effects, we also examined nonspecific (inert) globular and elongated macromolecules and applied different values of constant nudging forces to find the minimum force required to break the FG-meshwork and overcome the permeability barrier. For the globular-shaped inert cargo of 20 *nm*, the minimum force was approximately 70 *pN*, while for the elongated shape of equal volume, it was 40 *pN*; i.e. about 42% smaller force. Therefore, both for NTR-bearing and inert cargoes, elongated shapes penetrate more readily into the FG-meshwork and thus, cross the pore more efficiently than a globular cargo of equal volume.

This underlines the significance of the cargo shape effects in nucleocytoplasmic transport; particularly, in transport processes where pressure is utilized to send large nonspecific cargoes into the nucleus. In those processes, an elongated shape costs less, and therefore are more favorable, which might be, among others, why the nature has conserved the pressure-driven DNA ejection. Indeed, it has been recently shown that upon docking to the NPC cytoplasmic entry, herpes simplex virus type 1 (HSV-1) ejects its genome to the nucleus using the internal pressure within its capsid that is about 18 *atm* (Bauer2013). Intriguingly, this pressure amounts to a force of about 54 *pN* on the HSV-1 highly packed genome, having a nearly cylindrical shape with a diameter of about 20 nm (Bauer2013).

While the effects of size have been extensively studied in inter- and intracellular transport in a broad sense (Petros2010), and also more specifically in nucleocytoplasmic transport (Moussavi-Baygi2011a, Moussavi-Baygi2011b, Popken2015a), the shape effects just more recently proved

important in cellular uptake and target specificity (Geng2007, Petros2010, Kolhar2013). However, effects of the shape in nucleocytoplasmic transport have been mainly overlooked. Here, our findings imply that shape effects are important in nucleocytoplasmic transport; however, this awaits further in-depth elucidation.

4.7 The hydrophobic affinity difference acts like a stimulus that rapidly 'opens' the FG-meshwork and paves the way for successful transport of NTR-bearing macromolecules.

A NTR-bound macromolecule can easily perforate the FG-meshwork with a minimal nudging force of 4.0 *pN* (see Table 4.1), i.e. in the same order of magnitude as thermal kicks. The quick perforation of the FG-meshwork can be attributed to higher hydrophobic affinity between NTR and FG-repeats than between two FG-repeats (see SI) (Patel2007). Having higher affinity for the NTR, individual FG-repeats would favor to interact with hydrophobic patches on the NTR surface over interacting with themselves once a NTR is around. This would prime the FG-repeats to locally collapse and thus, the cargo can easily penetrate into the FG-meshwork. This is also the principal factor contributing to the rapid collapse of individual FG-repeats upon interacting with a NTR (Fig. 4.11). In this picture, the 'affinity difference' functions as a stimulus that triggers the open state of the FG-meshwork. Indeed, our model suggests that in the absence of such a stimulus there would be a much higher entropic barrier for the NTR-bearing macromolecule to enter the channel. In addition, we found that for a successful transport, the number of sufficient hydrophobic binding spots on the macromolecule surface is proportional to its size (see SI), implying larger cargoes need a proportionally larger number of NTRs to overcome permeability barrier (Table 4.1) in agreement with previous reports (Ribbeck2002).

Curiously, even after adding net negative charges to the nonspecific cargo surface, the cargo is not capable of perforating the positively charged FG-meshwork. To break the FG-meshwork and penetrate into, the negatively-charged nonspecific cargo requires the same amount of nudging force that is needed in the absence of surface charge. This highlights the dominating role of hydrophobicity versus charge in crossing the NPC channel, in agreement with a recent study (Goryaynov2014). Therefore, our model corroborates the observation that energetically favorable electrostatic interactions between NTRs and FG-repeats only marginally facilitate the transport process (Goryaynov2014).

4.8 Within the yeast NPC channel, the FG-meshwork is composed of two porous zones: a central high-density rod-like zone, and a peripheral low-density shell-like zone.

The results markedly show that there are two distinguishable regions inside the channel: a rod-like high-density central zone and a polymeric shell-like low-density peripheral zone (Fig. 4.13). The dense zone is largely hydrophobic with a diameter of about 22 *nm*, having the highest concentration of Phe residues (Figs. 4.1 and 4.3) that extends along the channel axis and possesses a dynamic hydrophobic core (Fig. 4.1). This hydrophobic core is reversibly dissolved upon interaction with NTR-bearing macromolecules (Fig. 4.11). The low-density zone is about 9 *nm* thick next to the channel wall and resembles a polymeric shell (Figs. 4.1).

In the study of Yamada et al., based on the stoichiometry, topology, hydrodynamic dimensions, and *speculated* interactivity of the yeast FG-Nups, the existence of two zones of traffic in the channel was proposed (Yamada2010). The low-density zone we observed here closely resembles what they termed ‘zone 2’ in the vicinity of the channel wall. Their proposed ‘zone 1’ located in the middle of the channel has a hydrophobic environment, consistent with the hydrophobic rod-like high density zone in our model. However, the zone 1 in Yamada et al.’s study features hollowness, as opposed to the rod-like zone observed here, which is densely packed with hydrophobic residues. This is believably because Yamada et al.’s study did not consider the dynamics, geometrical restrictions, and the end-tethering of FG-repeats—three important elements that literally underpin every aspect of the conformational behavior.

Remarkably, the presence of a cylindrical 'central plug' with low electron density, or a 'central transporter' with a weak smeared density in the channel was first proposed more than two decades ago by several independent studies in the *Xenopus* and yeast NPCs (Reichelt1990, Yang1998). Based on these studies, the central plug was hypothesized to adopt two distinct states, "open" and "closed" (Reichelt1990). It had also been realized that the central plug was basically the same as transporter (Reichelt1990). However, later on, it was presumed that the central plug was, at least in part, the remnants of the collapsed nuclear basket (Stoffler1999), or the cargo caught in transit during preparation (Stoffler2003). Importantly, some later studies hinted that the central plug might have been the result of aggregation and interaction of FG-repeats inside the channel (Yamada2010). Consistently, the spatial arrangement of Phe within FG-Nups in the absence of any dynamics, suggests that Phe residues tend to aggregate mainly toward the central axis of the channel (Fig. 3.16), implying the static picture supports the formation of a highly hydrophobic region within the middle of the FG-meshwork, in agreement with dynamical observations of our model.

Both zones in our model show porosity with the primary pore size of 4 *nm* in the rod-like dense zone and 6 *nm* in the low-density zone (Fig. 4.5). The pores can be viewed as passive holes to conduct free diffusions and thus set an upper limit on the size of small particles freely crossing the NPC. It should be noted, however, that these passive holes are not static; rather they constantly form and re-form within different locations of the highly dynamic FG-meshwork.

The passive pore sizes observed in our model closely match those from other qualitative studies reporting the largest size of freely diffusing particles across the NPC (Patel2007). In a detailed study by Mohr and coworkers (Mohr2009), authors found the apparent widths of the passive holes were 5.2 *nm*. Indeed, in that study they did not have control over specifying a region inside the NPC to differentiate between zones. Our results benefitted from a more spatial accuracy since we can discriminate between different regions within the NPC and quantify the porosity in each zone. Thus, 5.2 *nm* passive pores (Mohr2009) can be interpreted as an average value of the passive holes in two different zones.

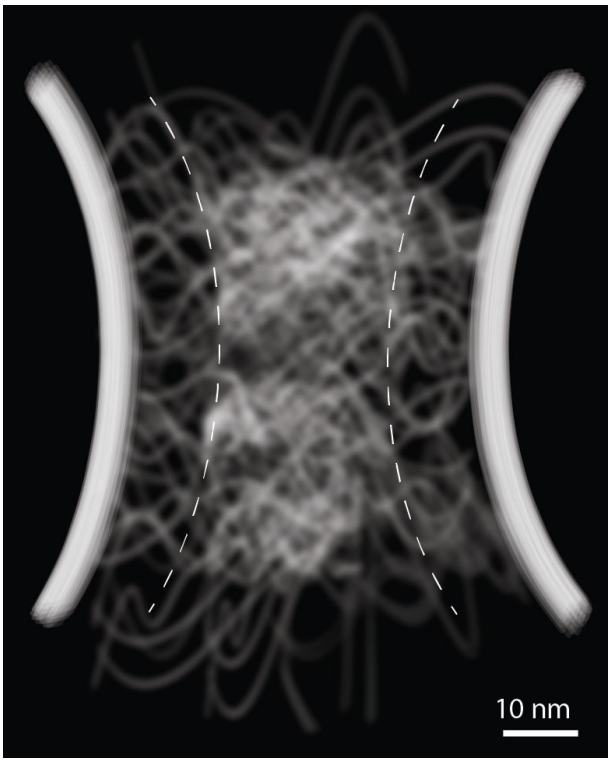


Figure 4.13. The side-view of the yeast NPC central channel based on the prediction of the current study, representing two zones inside the FG-meshwork (Fig. 4.1). Dashed lines are to guide the eye to distinguish between two zones.

4.9 The average configurational entropic cost sets the theoretical minimum number of hydrophobic binding spots: larger macromolecules need more NTR to overcome permeability barrier.

When a large active cargo enters the central channel, the disordered FG-repeats in its vicinity experience a reduction in their configurational entropy to locally accommodate the cargo. The decrease in configurational entropy of a single monomer, Δs_{mon}, can be estimated by calculating the reduction in the available space to the monomer (Teraoka2002a) as:

$$\Delta s_{mon} = k_B \ln\left(\frac{V_0 - V_c}{V_0}\right), \qquad (4.2)$$

where V_c is the volume of the active cargo, and V_0 the available space to the monomer before the cargo enters the channel, taking into account the excluded volume effects of other monomers: $V_0 = V_{chn} - (N-1)v_{mon}$. Here, V_{chn} is the bare channel volume, N the total number of monomers, and v_{mon} the vdW volume of a single monomer (Creighton1993).

If ρ is the average number density of the monomers inside the channel, for a single active cargo, on average, $n_{mon} = \rho V_c$ monomers must be dislocated to locally accommodate the cargo entirely inside the FG-meshwork, leading to an entropic cost of $\Delta S = n_{mon}\Delta s_{mon}$. This is, of course, a

minimal estimation for the entopic cost because we only consider dislocation of monomers immediate to the cargo surface and ignore farther layers.

The configurational entropic cost is compensated by *entropically* and *energetically* favorable hydrophobic, and to a lesser extent, *energetically* favorable electrostatic interactions between NTR and FG-repeats. Thus, ΔS for a given cargo can be used to determine the minimal number of hydrophobic patches on the active cargo surface required to overcome the lower bound of the configurational entropic barrier. For example, for a globular cargo of 15 *nm* in diameter, $\Delta s_{mon} \cong -0.04 k_B$ and $n_{mon} \cong 300$, and thus the lower bound of the entropic cost is $\Delta S \cong -12 k_B$. Therefore, at the cell temperature the free energy gain required to overcome this entropic barrier must be greater than $12 k_B T$. Considering the NTR-FG-repeats hydrophobic interaction, at least six hydrophobic patches would be required on the surface of a 15-*nm* cargo for the energy gain of $12 k_B T$. However, in reality this number of hydrophobic patches is not sufficient for this cargo to overcome the entropic barrier and perforate the FG-meshwork, at least during the simulation times in the current study. Yet, this calculation can set a theoretical lower bound to estimate the minimal number of hydrophobic binding spots required for a cargo with a certain size to overcome the entropic barrier of the NPC. It should be remarked, however, that this is a rough estimation; for simplicity we used an average value of the density for the monomers inside the channel and also, we ignored the end-tethering of the FG-repeats, which otherwise would not be straightforward to find the entropic cost. Moreover, the solvent translation entropy was ignored in our implicit solvent model. Nevertheless, this theoretical approach shows that larger cargos need more number of hydrophobic binding spots, and thus more NTRs, to overcome the permeability barrier, as was observed experimentally (Ribbeck2002).

4.10 Computational microrheology of the NPC

4.10.1 Introduction: rheology

> *"Everything flows and nothing abides; everything gives way and nothing stays fixed." Greek philosopher Heraclitus of Ephesus (536–470 BCE)*

The word *rheology* literally means "study of flow" from Greek words *rheo* ("flow") and *ology* ("study of"). Rheology is the study of the material deformation and flow (Morrison2001) through mathematical descriptions. In a gross sense, rheology detentions how 'soft' or 'hard' a material is; more specifically, it indicates the degree to which a material is 'fluid-like' or 'solid-like' (Larson1999). While in principle it is not limited to a particular type of fluid, the field is dominated by the viscoelastic study of complex fluids, like polymeric solutions, melts, and gels.

In a rheometer, two main sought-after parameters are stress and strain. The elastic modulus, a measure for the solidity of a material, is calculated from the ratio of stress to strain, while the viscosity as a measure of the fluidity of material is calculated from the strain rate (Chen2010). Viscoelasticity of a material manifests to extent to which the material is viscous and elastic. Newtonian fluids and Hookean elasticity are the most basic forms of viscosity and elasticity, respectively, as both are independent from the frequency of the stress and strain rate.

The viscous behavior of complex fluids, however, deviates from Newtonian fluids, while their elastic behavior deviates from Hookean materials (Morrison2001). The deviation arises naturally from the complicated intra- and inter-molecular interactions as well as molecule-solvent interactions, and is perhaps the most common feature in almost all polymeric materials with *long* molecular chains.

In addition to the natural viscoelastic deviation, several physical conditions can certainly complicate the behavior and the rheology of long molecular chains. Among these conditions are i) the co-existence of charged and hydrophobic groups along the polymers' backbones, ii) the spatial confinement of long polymers inside geometries with characteristic length much smaller than the chain's length, and iii) the reduction in polymers' d.o.f by end-tethering. Surprisingly, the long biopolymers of FG-repeat domains encompass all of these physical conditions, warranting their detailed rheological analysis. Nevertheless, to date there is no information about the FG-repeat domains rheology, predominantly due to lack of access to their functional state *in vivo* as well as *in vitro* at physiologically relevant scales. Here is where a computational biophysical model can provide invaluable insights into the microdynamics of the FG-repeat domains. In this section I will delve into the rheology of FG-repeat domains at microscopic scale solely from a computational viewpoint, which forms a part of my efforts in the current model.

4.10.2 Microrheology

For decades rheology has been dealing with viscoelasticity analysis of bulk materials at macroscopic scale in mechanical rheometer with solid moving plates (Macosko1994, MacKintosh1999). However, since 90s more novel techniques appeared toward rheological analysis at microscopic scales, particularly for soft materials like biological systems. The technique later called 'microrheology'. Based on the interpretation of thermal fluctuations of suspended micrometer-size probes in the complex microenvironment (Valentine2001), microrheology can provide detailed insight into microdynamics (Waigh2005) and local micro-viscoelastic properties of the medium (MacKintosh1999) (Fig. 4.13).

The key component here is an appropriate accompanying technique to measure the MSD of the probe. While the laser-deflection tracking (Gittes1997, Mason1997) and the diffusing wave spectroscopy (Gisler1999, Palmer1999) were the first techniques used to measure the MSD of the probe, later on more techniques were employed. Perhaps the most widely-used techniques are image-based tracking (Crocker2000, Chen2003, Valentine2004), AFM (Benmouna2004), optical tweezers (Starrs2002, Addas2004), and dynamics light scattering (Dasgupta2002, Popescu2002). 'Computational microrheology based on many-particle tracking method without external probe', is a novel technique that is developed in this dissertation. In essence this method is a natural extension of the conventional microrheology in computational domain; the strong prerequisite here is a physically and biologically validated model (see next section).

There are generally two modes of microrheology: passive and active. In the passive mode, the probe diffuses non-invasively in the complex fluid merely through thermal fluctuations, but the active mode involves some type of external force on the probe (Vázquez-Quesada2012), and thus might involve some alterations in the medium. The passive microrheology lies within the linear

viscoelasticity regime, while the active mode extends the analysis into the nonequilibrium and nonlinear regime (Squires2010). Due to its non-invasiveness, the passive microrheology is preferred over active mode in biological systems.

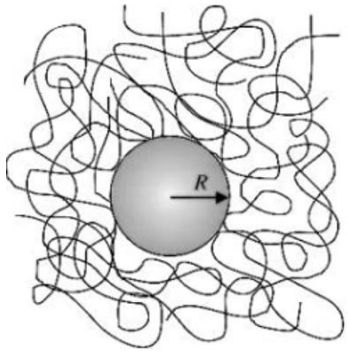


Figure 4.13. The thermal fluctuations of the spherical probe bead is monitored in the polymeric system. From the bead's MSD, the rheological properties of the polymeric system are extracted.

Microrheology in general, and in particular its passive mode, is motivated by several advantages, especially in biological investigations (MacKintosh1999). First, compared to traditional rheological approaches, there is need to only minuscule sample in microrheology (MacKintosh1999). This is notably important in the case of cellular and subcellular studies. Additionally, the frequency range in conventional rheology is limited by the mechanical inertia of the apparatus. This limitation is lifted in microrheology, and thus the local viscoelasticity can be probed at much higher frequencies in a non-invasive manner (Vázquez-Quesada2012). More importantly, the non-invasive nature of passive microrheology is absolutely appealing in biological investigations. Furthermore, heterogeneity of (sub)cellular medium can be straightforwardly probed through microrheology (MacKintosh1999). This is an important factor in dealing with intracellular studies. Moreover, the properties of many biological fluids, including cytoplasm, change while they are being measured in vitro conditions, demanding the in-vivo measurements (Vázquez-Quesada2012). Finally, the ability of probing viscoelasticity at microscopic scale has led to new detailed insights into the molecular csbasis of viscoelasticity (Zaner1989, Gardel2003).

The original idea of converting the information embedded in thermal fluctuations into microscopic viscoelasticity comes from a study by Mason and Weitz (Mason1995). The idea was later developed and two similar, but independent, formalisms emerged by Gittes et. al. (Gittes1997, Schnurr1997) and Mason et. al. (Mason1995, Mason1997), relating the thermal fluctuations to the microscopic energy storage and energy dissipation, i.e. micro-viscoelasticity.

In the former, Gittes et. al. (1997) reported a novel technique using laser interferometer microscope to measure the microscopic viscoelasticity from thermal fluctuations of micrometer probe (Gittes1997, Schnurr1997). Their study proved the approach worked well in predicting the viscoelasticity, loss (G'') and storage (G') moduli, of biological polymers—F-actin solutions—as well as industrial polymers—polyacrylamide gels (Gittes1997).

Using the linear response theory for a spherical probe with radius a thermally fluctuating in the fluid, authors derived the **generalized Stokes equation** under certain approximations (Schnurr1997):

$$\alpha^*(\omega) = \frac{1}{6\pi a G^*(\omega)}, \tag{4.3}$$

where $\alpha^*(\omega) = \alpha'(\omega) + i\alpha''(\omega)$ is the complex compliance coefficient. $\alpha^*(\omega)$ is the complex response function for the probe displacement. $G^*(\omega) = G'(\omega) + iG''(\omega)$ is the complex shear modulus in frequency domain. In their study they showed that from eq. 4.3, one can get the loss and storage moduli of the probe in terms of the compliance coefficient as (Schnurr1997):

$$\begin{aligned} G'(\omega) &= \frac{1}{6\pi a}\frac{\alpha'(\omega)}{\alpha'(\omega)^2 + \alpha''(\omega)^2} \\ G''(\omega) &= \frac{1}{6\pi a}\frac{-\alpha''(\omega)}{\alpha'(\omega)^2 + \alpha''(\omega)^2} \end{aligned} \tag{4.4}$$

As an alternative formalism, Mason et. al. (Mason1995, Mason1997) used a phenomenological approach (Mason2000) to describe the unforced motion of a probe particle in a given polymeric network. The investigation starts with the generalized Langevin equation for a probe particle with mass m and radius a:

$$m\dot{v}(t) = R(t) - \int_0^t \mu(t-\tau)v(\tau)d\tau, \tag{4.5}$$

where $R(t)$ is the random force on the particle and $\mu(t)$ is the time-dependent viscoelastic memory function.

Because there is energy stored in the medium due to elasticity, the functional form of FDT differs from what was introduced earlier in the previous chapter (eq. 3.8) for a purely viscous environment. The FDT for this viscoelastic microenvironment is written as:

$$\langle R(0)R(t)\rangle = k_B T \mu(t), \tag{4.6}$$

which does not include the delta function anymore.

If one takes one-sided Laplace transform from both sides of the eq. 4.5, a relation between the viscoelastic memory function and the velocity autocorrelation function, and thus the MSD (mean-squared displacement), will emerge in the frequency domain. The key assumption here is that at microscopic scale, the viscoelastic memory function is proportional to the frequency-dependent viscosity of the complex fluid as:

$$\tilde{\mu}(s) = 6\pi a\tilde{\eta}(s), \tag{4.7}$$

with s being the frequency in the Laplace domain. It then follows that:

$$\tilde{G}(s) = s\tilde{\eta}(s) = \frac{k_B T}{\pi a s \left\langle \Delta\tilde{r}^2(s) \right\rangle} - \frac{ms^2}{6\pi a}, \tag{4.8}$$

where the Laplace-transformed modulus $\tilde{G}(s)$ is the viscoelastic spectrum of the fluid and is the unilateral (one-sided) Laplace transform of the stress relaxation modulus $G_r(t)$. The stress relaxation modulus is the time-dependent shear stress to the corresponding strain (Ferry1980). It is a common material property used to discuss the linear viscoelasticity of a given polymeric system (Dealy2007). The discrete form of $G_r(t)$ is written as (Dealy2007):

$$G_r(t) = \sum_{i=1}^{N} G_i e^{-\frac{t}{\tau_i}} \tag{4.9}$$

in which the set of $\{\tau_i, G_i\}$ are the empirical parameters, constituting the discrete relaxation spectrum. To develop the theoretical foundation of microrheology, Mason et. al. (Mason1997, Mason2000) used the unilateral Laplace transform of the $G_r(t)$ as:

$$\tilde{G}(s) = s\mathrm{L}\{G_r(t)\} = s\int_0^\infty G_r(\hat{t}) e^{-s\hat{t}} d\hat{t} \tag{4.10}$$

The first term in the RHS of eq. 4.8 is due to the thermal fluctuation, and the second term is due to inertial effects (Mason1995). At low-frequency domains, the eq. 4.8 is simplified for a particle freely diffusing in a purely viscous liquid to $\left\langle \Delta\tilde{r}^2(s) \right\rangle = \frac{6D}{s^2}$ with the frequency-independent viscosity of $\eta = \frac{k_B T}{6\pi a D}$ (Waigh2005). Similar to Brownian analysis of a free particle, inertial effects in microrheological analysis can also be neglected in low *Re* numbers regime (Mason1997, Schnurr1997, Squires2010), and thus:

$$\tilde{G}(s) = \frac{k_B T}{\pi a s \left\langle \Delta\tilde{r}^2(s) \right\rangle} \tag{4.11}$$

This is the **generalized Stokes-Einstein relation** (GSER) in the Laplace domain originally developed by Mason et. al. (Mason1995, Mason1997), laying the foundation of the modern microrheology in the linear regime.

The GSER was developed based on an empirical approach (Vázquez-Quesada2012). Nevertheless, Levine et. al. (Levine2000) put it on a firm theoretical foundation (Waigh2005), and a decade later, Xu et. al. (Xu2007) used a detailed fluid mechanical analysis to show that GSER is valid whenever the usual Stokes-Einstein relation holds, proving the solid physical foundation for the GSER.

All in all, although both microrheological formalisms briefly introduced here are very similar, the GSER approach enjoys a more rigorous physical basis. Thus, in the current work, I have adopted this approach as the basis of microrheological analysis of the NPC. The only key assumption in

the GSER is that the fluid around the spherical probe is isotropic and incompressible, which is indeed applicable to a vast majority of physical and biological systems, including FG-repeat domains inside the NPC.

As for the frequency-domain representation of the linear viscoelasticity, an alternative to the Laplace transform is the Fourier transform of the stress relaxation modulus, yielding the complex shear modulus as: $G^*(\omega) = i\omega F_u\left\{G_r(t)\right\} = i\omega\int_0^\infty G_r(\hat{t})e^{-i\omega\hat{t}}d\hat{t}$. Then the GSER in Fourier domain will be (Mason2000):

$$G^*(\omega) = \frac{k_B T}{\pi a i\omega F_u\left\{\left\langle \Delta r^2(t)\right\rangle\right\}} \qquad (4.12)$$

Eq. 4.12 has the same information about the viscoelasticity of the system as eq. 4.11 does (Mason2000). The formalism in eq. 4.12 has, however, the advantage that it directly relates the MSD to the complex shear modulus. The importance of the complex shear modulus lies within the main task of (micro)rheology: to determine how energy is stored and dissipated in the material. This task is precisely expressed by the complex shear modulus (Crocker2000) straightforwardly through its real and imaginary components. Therefore, in the current work, I am using Fourier representation of the microrheology, instead of Laplace representation: *just by knowing thermal fluctuations of the beads in FG-repeats, we can obtain the complex shear modulus inside the NPC*.

Using an estimation method for the Fourier transform of the MSD as $i\omega F_u\left\{\left\langle \Delta r^2(t)\right\rangle\right\} \approx \left\langle \Delta r^2\left(\frac{1}{\omega}\right)\right\rangle \Gamma\left[1+\varsigma(\omega)\right] i^{-\varsigma(\omega)}$, Mason (2000) derived the following equation for calculating the complex shear modulus solely from thermal fluctuations of the probe (Mason2000):

$$G^*(\omega) = \frac{k_B T}{\pi a\left\langle \Delta r^2\right\rangle \Gamma\left[1+\varsigma(\omega)\right]}\left(\cos\frac{\pi\varsigma(\omega)}{2} + i\sin\frac{\pi\varsigma(\omega)}{2}\right), \quad (4.13)$$

where the gamma function can be estimated by:

$$\Gamma\left[1+\varsigma\right] \approx 0.457\left(1+\varsigma\right)^2 - 1.36\left(1+\varsigma\right) + 1.90\,, \qquad (4.14)$$

in which:

$$\varsigma\left(\omega = \frac{1}{t}\right) = \frac{d\ln\left\langle \Delta r^2(t)\right\rangle}{d\ln t}\,, \qquad (4.15)$$

is the power-law exponent, describing the logarithmic slope of $\Delta r^2(t)$ at t. Equations 4.13-15 are the basis of calculating the viscoelastic response within Brownian dynamic simulations in the current dissertation.

4.10.2.1 Two notes on GSER: breakdown of GSER; the inertial effects

4.10.2.1.a Breakdown of GSER

It should be noted that although the eq. 4.11 has a solid physical foundation as it is the based on the ensemble-averaged of the velocity autocorrelation function, it is not a universal law and there are certain cases that GSER is not applicable. First, as the GSER follows from equilibrium statistical mechanical analysis, it fails in the analysis of a nonequilibrium system that is significantly far away from equilibrium. In addition, GSER cannot provide accurate measurement for active microrheological setups in which the probe is strongly forced into the medium such that the material is derived out of the equilibrium state. The third case where the GSER is not generally applicable is an active material that consumes internal chemical energy to produce mechanical stresses (Squires2010). There are alternative methods to conduct microrheological analysis in such systems (Squires2010), which are beyond the scope of my work. For example, an alternative formalism termed 'two-point microrheology' has been developed, in which a cross-correlated pair of probes is employed to analyze the viscoelasticity of the medium under the assumption of extreme heterogeneity (Crocker2000, Lau2003, Crocker2007). The predictions of this formalism deviate from microrheology and is not considered in the current work.

While in general none of the aforementioned conditions happen in biological applications, ATP-dependent energy-producing processes can potentially confound the Brownian motion of the probe (Crocker2007). This is applicable to the microrheological analysis of intracellular substances, and thus to get reliable results, the cell needs to be depleted of ATP. For FG-repeat domains, however, GSER is perfectly applicable, and is the best available method, as there is no source of energy producing inside the NPC.

4.10.2.1.b Inertial effects

The characteristic time $\tau = \frac{m}{\xi}$ was introduced in the previous chapter as a criterion to determine whether the observation time scale lies within the ballistic or diffusive regimes. Analogously, the characteristic inertia-friction frequency for the probe bead is defined as $\gamma_c = \frac{6\pi a\eta}{m}$ in the frequency domain (Mason2000). For frequencies below γ_c the inertial effects of the probe bead can be neglected and GSER holds (Mason1997). Beyond this regime, the inertial effects in microrheology can indeed be pursued analytically. For example, Liverpool et. al. (Liverpool2005) and Atakhorrami et. al. (Atakhorrami2005) developed an exact solution of the linear response theory at very short times and high frequencies in the presence of inertial effects. As such, GSER (eq. 4.11) is not applicable to an MD simulation. This regime represents ballistic motions and is not considered in Brownian dynamics. For example, in the current dissertation a five-residue long monomer with a mass of about 500 Da would lead to a diameter of about 1.2 nm. In cytoplasmic viscosity, the characteristic frequency would be about $\gamma_c \approx 10^{14}\,Hz$. The maximum frequency in the simulation within the current dissertation is two orders of magnitude lower than γ_c, justifying the application of microrheological analysis under the eq. 4.12.

4.10.3 Mechanical spectrum of FG-repeat domains: computational microrheology based on many-particle tracking method without external probe

In the current dissertation, I extended the passive microrheology of 'single-particle tracking with external probe' to the 'many-particle tracking without external probe'. The many-particle method based on computational biophysical modeling produces a significantly richer pool of samples compared to the single-particle method, resulting in statistically more reliable values for the rheological properties of the system.

More importantly, it directly measures the native, unperturbed thermal fluctuations of the polymer network rather than dealing with an external probe (Fig. 4.13). This would nullify any possibility of the polymeric network being perturbed as a result of the existence of an external probe therein (Levine2000). In this approach, in place of an external probe, we monitor *all* of the constituent particles of the medium. Thus, instead of thermal fluctuations of the probe, the native thermal fluctuations of all monomers are *directly* analyzed. In addition, computational microrheology can extend the analysis to very high frequencies beyond the capability of the available experimental techniques (Fig. 4.14).

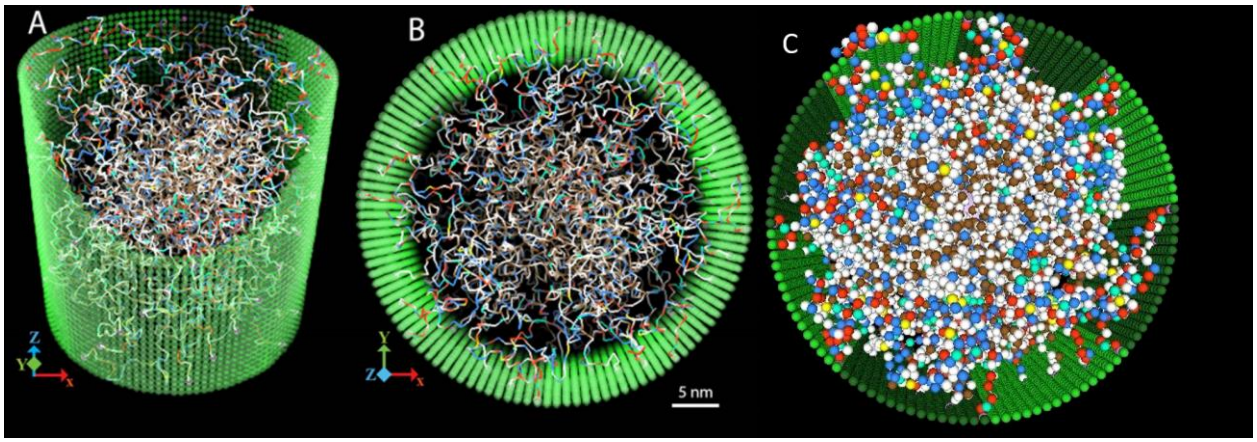


Figure 4.13. The representation of microrheological measurement of FG-repeat domains through a many-particle tracking method. Every single monomer in FG-repeat domain is treated as a bead probe whose thermal fluctuation is monitored and measured to obtain the shear moduli inside the channel. A and B) Side and top view of the simulated NPC channel with FG-repeats. C) The bead representation of the FG-repeat domains. Each bead represents the physical a monomer (connecting springs are not shown for clarity), playing the role of a native probe for microrheological analysis. See Fig. 3.17 for the meaning of different colors.

It should also be mentioned that a rheological analysis in the Fourier domain (eq. 4.12) mathematically requires frequencies from 0 to ∞ (Crocker2007). Technically, it requires evaluation of MSD over a wide range of frequencies. In practical measurements there is not a

single instrument or technique being able of measuring shear moduli over all frequencies (Chhabra2011), and thus, the results of several experiments are superimposed onto a master curve (Dealy2006).

Similarly, for a computational microrheology results to be inclusive, the MSD should be evaluated over a wide range of frequencies, which is impeded by the limited simulation time. One way to overcome this limitation is to conduct multiscale simulations and superimpose the results to obtain a master curve, covering a wide range of frequencies. There are, however, problems in relating the shear moduli obtained from different scales, arising from different levels of granularity at distinct scales. Alternatively, one can run a single-scale simulation for a very long time, with the penalty of dealing with huge trajectory files and overdue results!

Here, I chose to run the single-scale simulation for a long time to sufficiently cover the frequency domain. The simulation was running for 7,700 hours of wall-clock time over the course of 40 months from August 2014 to December 2017 (this included run time, queue time, analysis time, and overhead time) on the Stampede supercomputer at TACC (Texas Advanced Computing Center). This resulted in a huge trajectory encompassing 100 ms of simulation time. The trajectory is large enough to cover a wide mechanical spectrum. It also enjoys rich statistical number of sampling points, so that it does not suffer from statistical unreliability at large lag times. For the current model of FG-repeat domains with over 7,000 five-residue long monomers, the number of samples of squared displacements varies from 10^9 for the shortest lag time to 10^8 for the longest lag time. This can be seen in the stability in the tail of the MSD graph in Fig. 4.15.

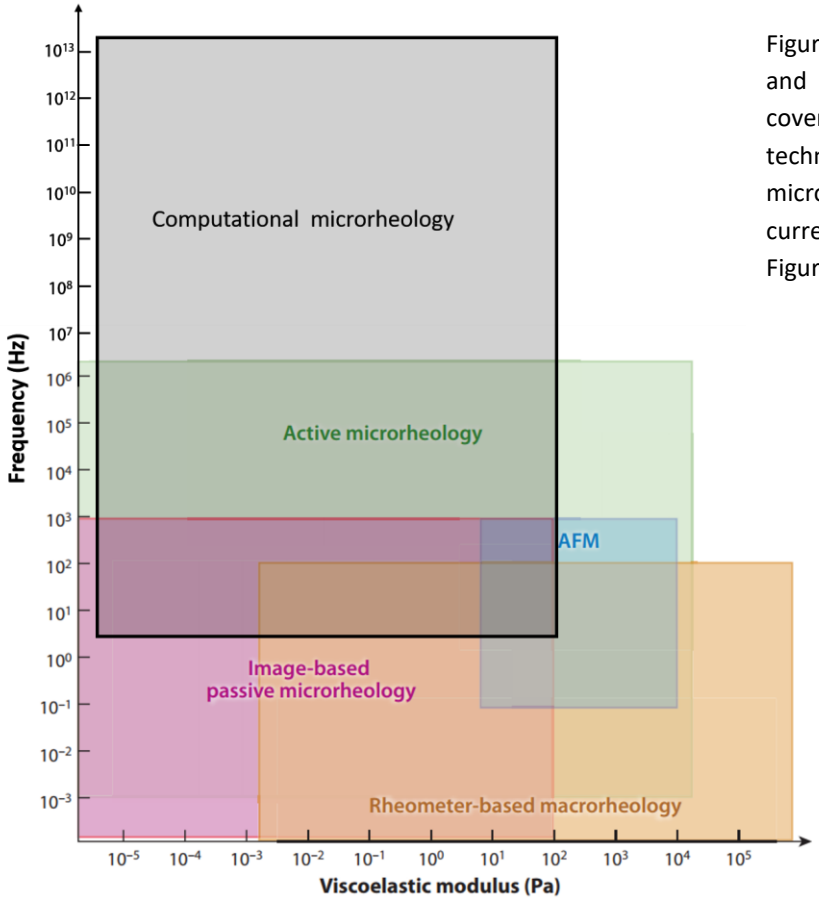


Figure 4.14. The range of frequencies and viscoelastic moduli that can be covered in different (micro)rheological techniques. The range of computational microrheology is drawn based on the current dissertation. The rest of the Figure is adopted from (Chen2010).

4.10.3.1 Mean-squared displacement of FG-Nups sheds light on the microdynamics inside the NPC: a layered-based model.

As elegantly portrayed in microrheology, the time-dependent MSD is the most fundamental quantity in shaping the microdynamics of a polymeric system. In principle, having the MSD of a system is equivalent to having the detailed picture of the system's microrheology.

Thus, here I delve into the MSD of FG-repeats inside the NPC and try to understand how different factors affect the MSD. For a single monomer we have:

$$MSD_i(t) = \left\langle \left| \mathbf{r}_i(t+T) - \mathbf{r}_i(t) \right|^2 \right\rangle, \qquad (4.16)$$

in which T is the lag time. The MSD is then averaged over all monomers, providing a measure of the collective self-diffusion of FG-repeat domains inside the NPC.

In its most general form, the MSD varies with time as $MSD \propto t^{\alpha}$. Normally, for a freely diffusing particle with random walk statistics, the exponent is one and $MSD \propto t$, where eq. 3.4 is recovered

in 3D. As the diffusion drifts from normality to anomalous diffusion, the exponent deviates from unity. Superdiffusion ($\alpha > 1$) occurs when the MSD grows faster than normal diffusion due to convection in the medium (Rogers2008), Levy walks (Klafter2005, Klafter2011), or any other kind of directed motion accelerating the diffusing particle in some direction (Saxton1997, Caspi2002). The vast majority of anomalous diffusions, however, lie within the class of subdiffusive motion $(0 < \alpha < 1)$, where the MSD increases more slowly than the normal diffusion, particularly at long times.

There are several factors that are known to turn the free diffusion of a particle into a subdiffusive motion, rendering MSD nonlinear in time. The major factors are:

i) crowdedness (Weiss2004),
ii) confinement (Saxton1997),
iii) inter- and intraparticle interactions that are known to suppress the Brownian motions of freely diffusing particles (Bressloff2013).

Clearly, the diffusion analysis becomes more complicated for the aggregation of long molecular chains, as microdynamics of the system dramatically shifts compared to an aggregation of globular particles. Here, in addition to three aforementioned factors, iv) end-tethering and v) the length of polymer chains also contributes in anomalous diffusion.

All of these factors are elegantly present inside the NPC, posing tremendous challenge to investigation of the self-diffusivity of the selectivity barrier. In what follows, my goal is to elucidate this aspect of the FG-repeats.

I start with native NPC channel (Fig. 4.13). Expectedly, the MSD at short times increases linearly in log-log scale, while at very long times plateaus, confirming the anomalous self-subdiffusion motion of the FG-repeats (Weiss2004). This is precisely do to the constrained motion of the long, end-tethered FG-repeats confined to the inner channel.

Importantly, from the log-log plot it can be seen that before the onset of the plateau regime at $t = 10^{-4} s$, there are two distinguishable diffusion phases in the MSD, i.e., $\alpha = 0.83$ and $\alpha = 0.49$. Initially, the MSD increases as $MSD \propto t^{0.83}$, but then the exponent further drops so that $MSD \propto t^{0.49}$. Therefore, at very short times as small as $t = 10^{-11} s$, the MSD deviates from normal diffusion by $\alpha - 1 = -0.17$, while at longer times the deviation is stronger $\alpha - 1 = -0.51$. At very long times of $t = 10^{-2} s$, however, the anomality is significantly augmented and the MSD plateaus.

Admittedly, the first phase in the MSD arises from the fact that at very short times, monomers diffuse so little that they barely feel the effects of the main subdiffusion-driver factors, such as composition, confinement, and end-tethering. As longer times, which amounts to the second phase and begins at $t = 10^{-8} s$, chains diffuse around and presumably entangle with one another in the crowded environment inside the channel (see the section under 'the effects of end-tethering'). This would result in a significant decrease in their diffusional motion. In fact, a decrease in the exponent from 0.83 in the first phase to the 0.49 in the second is a big change. Yet, other hypothesized subdiffusion-driver factors, i.e. composition, confinement, and end-tethering, have not been felt by the chains in this phase (this argument will be more clarified in the next three sections). At very long times, i.e. the last phase, while FG-repeats presumably disentangle, the subdiffusion-driver factors completely dominate the diffusive motion of monomers, and MSD plateaus.

From this argument, it believably becomes clear that the underlying physical reasons for the anomalous diffusion are different in each of these three phases. Nevertheless, it is important to note that over all times accessible to the current model the self-diffusion of FG-repeat domains deviates from normality and is subdiffusive, as the exponent is always less than unity. The amplitude of deviation, however, is strongly dependent on the time scale.

Another interesting observation from the log-log plot is that the MSD is z-direction is about 30% higher that than in x- and y-directions (37 nm^2 vs 29 nm^2). This is reasonably due to the open ends of the channel along the z-direction, where the nearby FG-repeats can diffuse, subject to their constraints, into the cytoplasmic and nuclear spaces. Thus, FG-repeats in general, and those near the entries in particular, have higher mobility in the z-direction. The physiological importance of the higher agility of FG-repeats in the axial direction (z-direction in the current model) perhaps lies within the efficiency of nucleocytoplasmic transport which is conducted in the same direction. The higher axial mobility of FG-repeats presumably enables them to establish more rapid interactions with NTR-bearing cargos, overall resulting in a more efficient transport.

The height of the plateau represents the maximum area swept by FG-repeats. Every single FG-repeat, on average, cannot discover beyond this area, because its diffusion is constrained. The square root of the plateau height can be thought of as the average characteristic length of the region that an FG-repeat diffuses in. The value of the squared root of plateau height is consistent with the gyration radii of FG-repeat domains (Fig. 4.16). The averaged radius of gyration of all FG-repeat domains is about 7 nm, while the individual Rg for each FG-Nup varies from 5 to 10 nm (Fig. 4.16). By definition, the variation of Rg is correlated with the polymer length; and thus, the longest Nups, namely Nup116 and Nup100 have the largest Rg (See Table 3.3), as presumably they take up largest space. Similarly, shortest Nups (Nup57 and Nup49) exhibit smallest values of Rg.

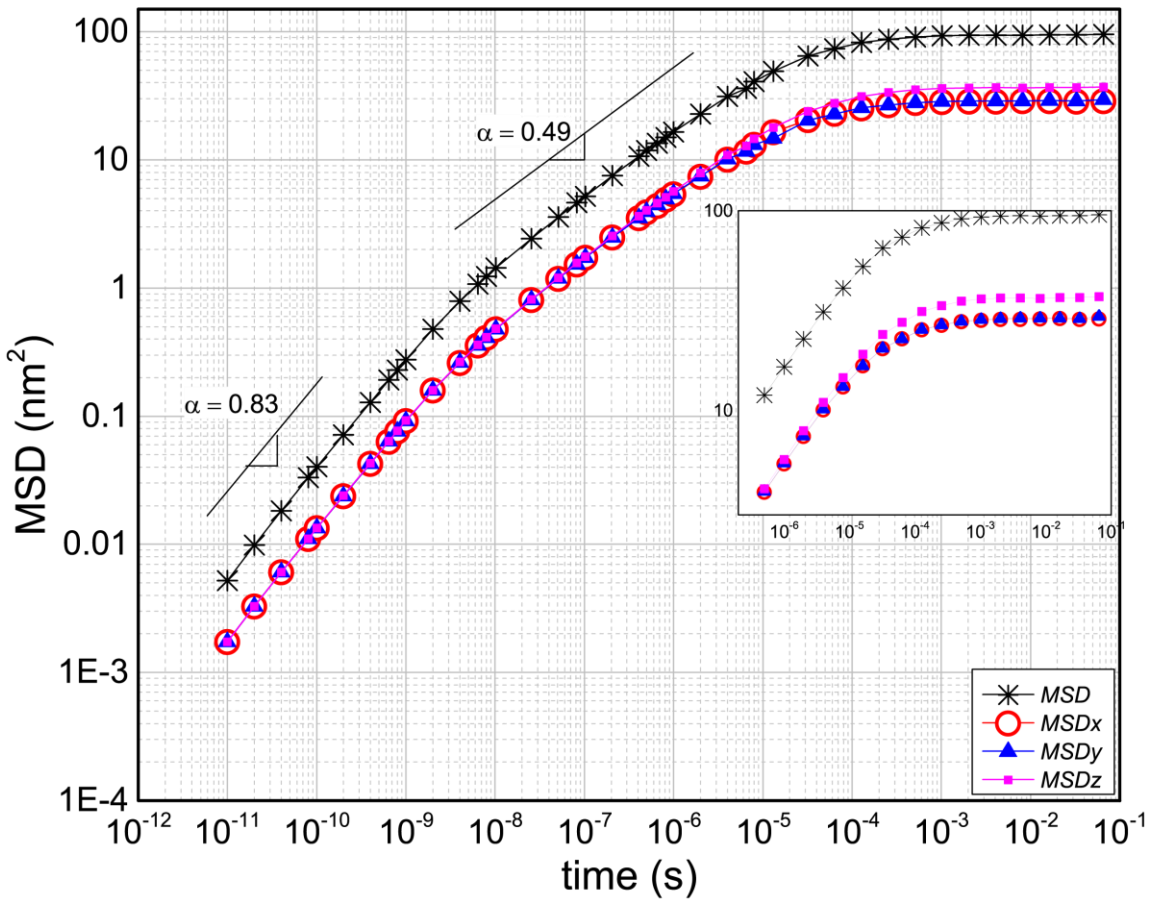


Figure 4.15. The mean-squared displacement of FG-repeat domains inside the NPC shown in x-, y-, and z-direction. The MSD in x and y directions are identical, because of the rotational symmetry. The MSD in z-direction shows higher magnitudes, particularly at large time lags, stemming from two open ends. Inset: zoom-in on the long-times region to more visibly see the differences between MSDs in orthogonal directions.

The radius of gyration roughly represents the volume of a sphere that the polymer chain on average takes in the space (Fig. 4.16 Inset). The Rg is determined not only by the linear size of the polymer, but also by every factor that affects the dynamical structure of the polymer. Particular to FG-repeats is that the Rg is affected by the grafting point of the FG-Nups inside the NPC. Nup116 is at the cytoplasmic entry and enjoys access to the cytoplasmic free space (See Fig. 3.16). Therefore, Nup116 can exhibit more extension than other FG-Nups and consequently more fluctuations, as can be seen from large variations in its Rg. Interestingly, the other longest FG-Nup, i.e. Nup100, is right below the Nup116 (Fig 3.16), showing the next highest degree of fluctuation in Rg.

Curiously, two longest FG-Nups are next to each other at the cytoplasmic entry. Although the exact locations of FG-Nups cannot be determined unambiguously, the present arrangement of FG-Nups in the current model inspired by (Alber2007a) is consistent with the NPC function. Mechanistically speaking, the presence of two longest FG-Nups at the

cytoplasmic entry would contribute in formation of an efficient rejecting barrier against nonspecific cargos. This is similar to the role of flexible cytoplasmic filaments that was predicted by our 2D model to effectively reject nonspecific cargos (Moussavi-Baygi2011a) as we discussed earlier.

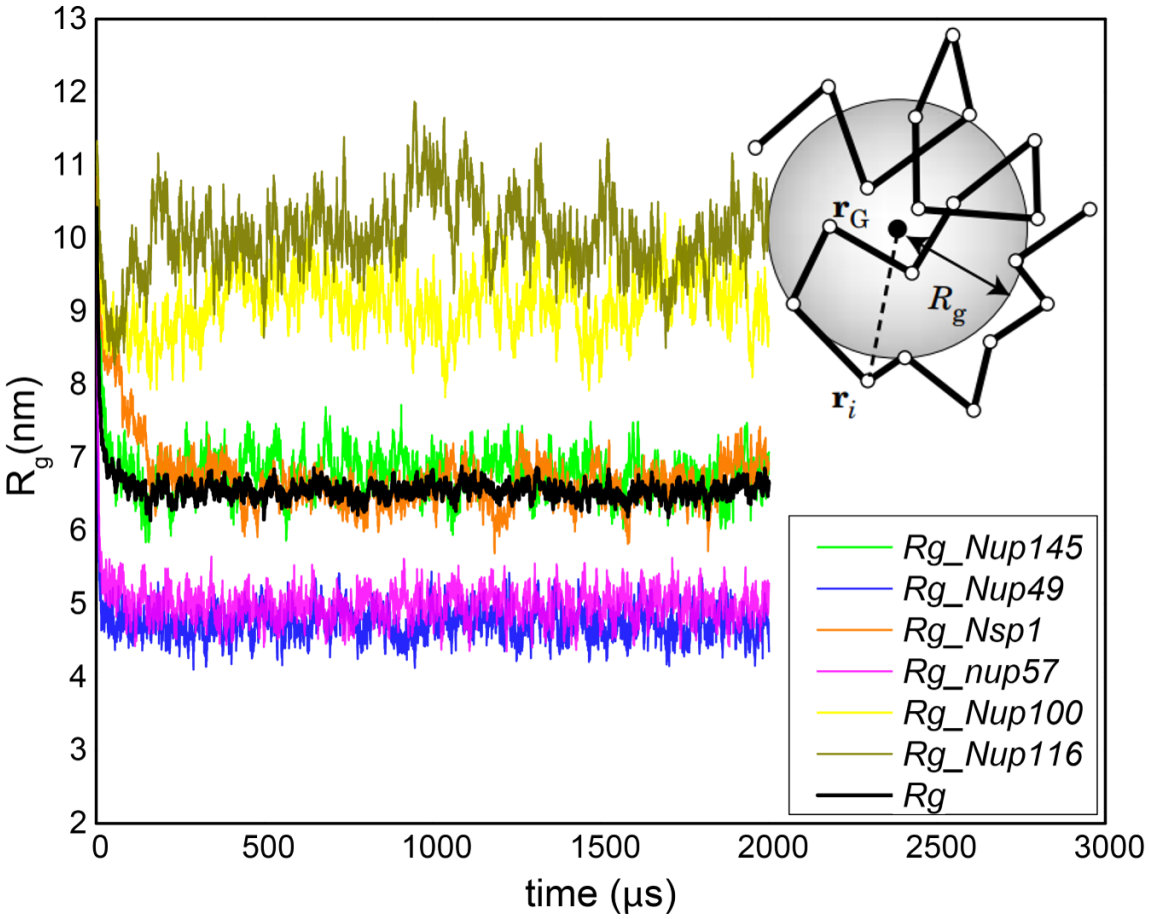


Figure 4.16. The time-evolution of gyration radius of FG-Nups in a native NPC. The black line indicates the Rg averaged over all of the FG-repeats. The assorted colored line show Rg for individual FG-Nups. Inset: graphical representation of the physical meaning of Rg (adopted from (Teraoka2002b)).

This simple observation is a signature of an erratic local microdynamics within the NPC, warranting a more in-depth analysis. To this end, I analyzed the MSD of individual residues along the backbone of each FG-Nup. The results are summarized in Fig. 4.17. All MSD heatmaps are shown on the scale of 0 to 200 nm^2 to facilitate the direct comparison between different FG-Nups.

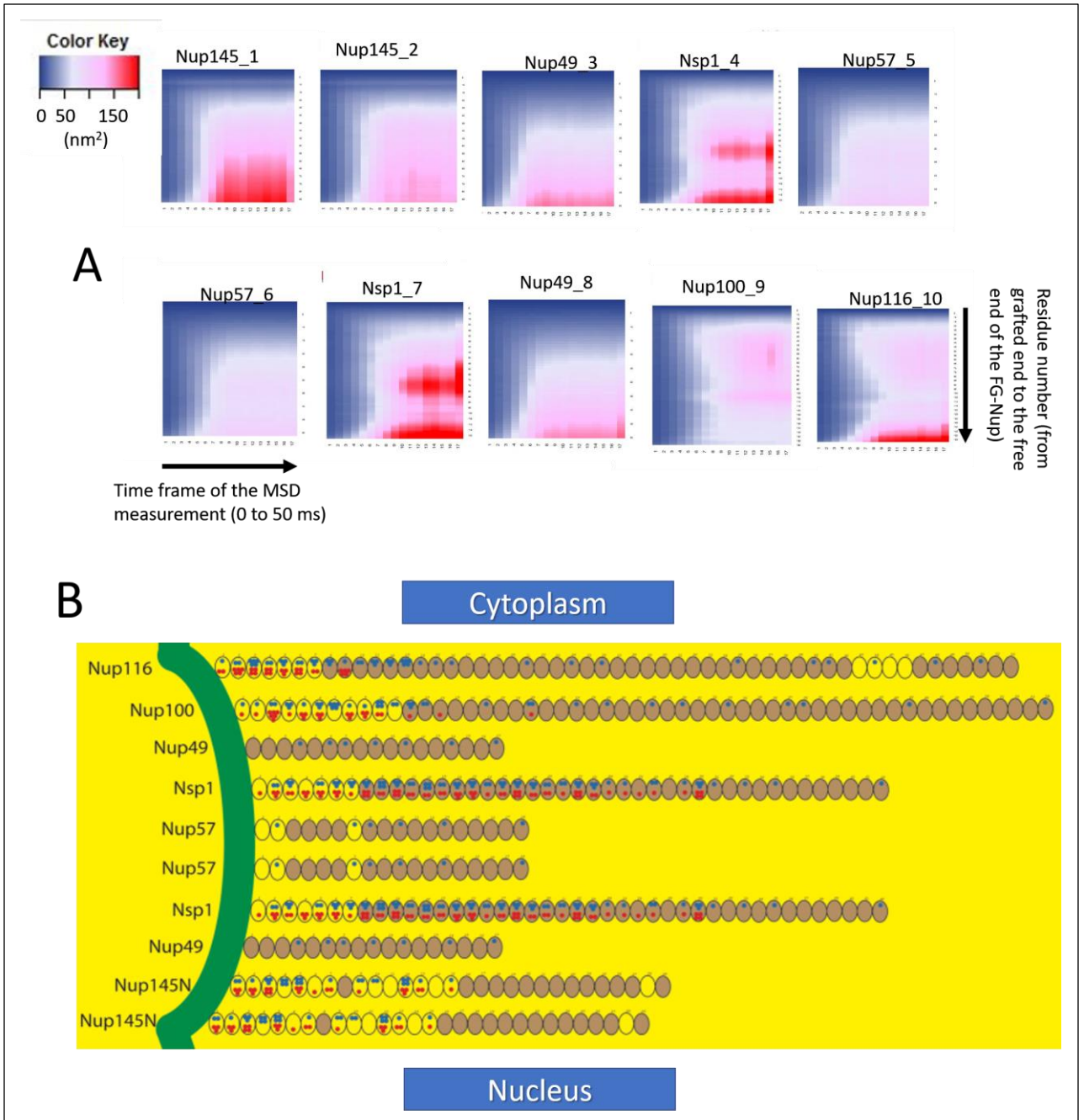


Figure 4.17. A) Heatmaps show the MSD per residue in each FG-Nups on the scale of 0 to 200 nm^2. The name of the FG-Nup is indicated at the top of the corresponding heatmap. The vertical axis is the residue number along the FG-Nup backbone, numbered from the tethered end to the free end. The horizontal axis is the lag time over which the MSD is calculated. FG-Nups are numbered from 1 at the nuclear entry to 10 at the cytoplasmic entry. Therefore, Nup145_1 on the heatmap indicates the first FG-Nup in the NPC channel on the nuclear entry. There are two copies of some FG-Nups, like Nup145 and Nsp1. B) A 2D representation of the FG-Nups in the central channel as an eye-guide to compare the location of FG-Nup with its MSD.

Expectedly, within each individual FG-Nup, residues toward the free end of the Nup have generally the highest MSD, with an exception of Nup100 whose free-end residues have smaller MSD than its mid-residues. This has presumably a lot with the composition of the Nup100. The Nup100 has the biggest number of FG-motifs, and is also the longest FG-Nup in the channel (Table 3.3). It is thus very likely for the eight copies of Nup100 to reach each other from opposite spoke domains and engage in adhesive hydrophobic interactions. As a result, on average, the mobility of the residues on the backbone of Nup100 is reduced compared to those on the backbone of other FG-Nups.

Another stood-out Nup inside the channel is the Nsp1. There are two copies of Nsp1 within the NPC central channel, Nsp1_4 and Nsp1_7. From the MSD heatmaps it appears that in addition to residues located at the free end of Nsp1, the mid-residues along this Nup also have large MSDs. No other FG-Nup inside the channel show such a pattern in the of MSD its residues. Again, one way to explain this observation is based on the composition of the Nsp1; it has the highest density of charge within the hydrophobic region (Fig. 4.17. B). The presence of charged residues derives the polymer backbone into an extended conformation, and lots of screened electrostatic interactions. This would lead the mid-residues to have higher mobilities compared to residues in the middle of other FG-Nups. Thus, while the middle large MSD zone along the Nsp1 is due to the high charge density, the large MSD toward the free end of the Nsp1 is similar to other FG-Nups.

Also visible from MSD heatmaps is that FG-Nups located near the cytoplasmic or nuclear entries show larger values of MSD compared to FG-Nups in the middle of the channel. This is anticipated, as there is a larger crowdedness in the middle of the channel, reducing the movability of FG-Nups therein. For example, there are two copies of Nup145 per spoke domain. The Nup145_1 that nearest to the nuclear entry has the higher MSD than Nup145_2. These two Nups are identical, except for their grafting point on the wall, suggesting that Nups near the entry are more dynamic. Similarly, on cytoplasmic entry, residues of Nup116 reveal larger mobility than neighbor residues on Nup100.

All in all, it is tempting to hypothesize a layered-based barrier inside the NPC. The first layer is an entropic barrier formed at the cytoplasmic entry of the channel by eight copies of Nup116. The second layer is a physical barrier formed by eight copies of Nup100. A similar two-layered arrangement is formed on the nuclear entry of the central channel by Nup145_1 and Nup145_2. In the mid channel, there are combination of highly mobile and sticky residues.

4.10.3.1.a The effects of FG-repeats' composition on the MSD: hydrophobicity and charge

Hydrophobic and charged residues are the main compositional characteristics of the selectivity barrier. It is now well established that the hydrophobicity is among, if not the only, major factors in the formation of selectivity barrier. Here, I am trying to understand how hydrophobicity affects the microrheological dynamics of the FG-repeat domains. To this end, I followed the same protocol as the Gorlich group in which all of the Phe residues are mutated to S (Frey2007), eliminating all of the hydrophobic content of the FG-repeat domains. The averaged MSD over all monomers is then calculated after a several long runs of the simulation.

It appears that the F>S mutation significantly augments the amplitude of the MSD compared to the native FG-repeat domains (Fig. 4.18). The plateau occurs at about 230 nm^2, more than twice in the native NPC. This is undoubtedly due to the lack of inter- and intra-chain hydrophobic adhesiveness in the mutate FG-repeats. In this scenario, a purely repulsive entropic meshwork is formed inside the channel, so that chains erratically scape from the inner channel to the outer space (Fig. 4.18). This behavior is also visible from the MSD curve in the z-direction, which has significantly larger plateau height than MSD in x- and y-directions.

It is also seen from the log-log plot that similar to the native NPC, there are two diffusion phases before the MSD plateaus; $MSD \propto t^{0.87}$ and $MSD \propto t^{0.52}$. In both phases the exponents are a bit higher than the corresponding exponents in the native NPC (0.87 vs 0.83 and 0.53 vs 0.49). Understandably, this is due to the lack of adhesive hydrophobic interactions among monomers, leading to higher mobility. It is also worthwhile to note that the onset of the second and the third (plateau) phases are the same as in the native NPC ($t = 10^{-8} s$ and $t = 10^{-4} s$, respectively)

The electrostatic interactions within the NPC was suggested to contribute in increasing the transport rate (Colwell2010b), and the charged-like regions inside the NPC were found to regulate the dynamics of the FG-repeat domains (Peyro2015b). Another study found that the electrostatic interactions between the NTR-bearing cargo and FG-repeats only enhance the binding probability, but do not play a role in transport efficiency (Goryaynov2014). The results of our 3D model also suggest that the NTR-FG electrostatic interactions only marginally participate in the transport efficiency and selectivity (Moussavi-Baygi2016). Yet, the effects of FG-repeat domains' charged content on the microrheology is not known.

To see the contribution of electrostatic interaction in the microrheology of mutated meshwork, here I performed a second round of mutation by changing all of the charged residues to Ser. Therefore, the meshwork is depleted of hydrophobic residues as well as charged residues (ΔF_ΔCharge). Interestingly, the MSDs for ΔF and ΔF_ΔCharge are identical, not only in pattern, but also in values (Fig. 4.18). This suggests that the electrostatic interaction does not play any role in the microrheology of FG-repeat.

The results here suggest that from FG-repeats composition viewpoint, the hydrophobicity, but not electrostatic interaction, is a determining factor in the MSD, and thus in microrheology of FG-repeat domains. This is because the electrostatic interactions are highly screened inside the NPC, but hydrophobic interactions actively associate FG-motifs.

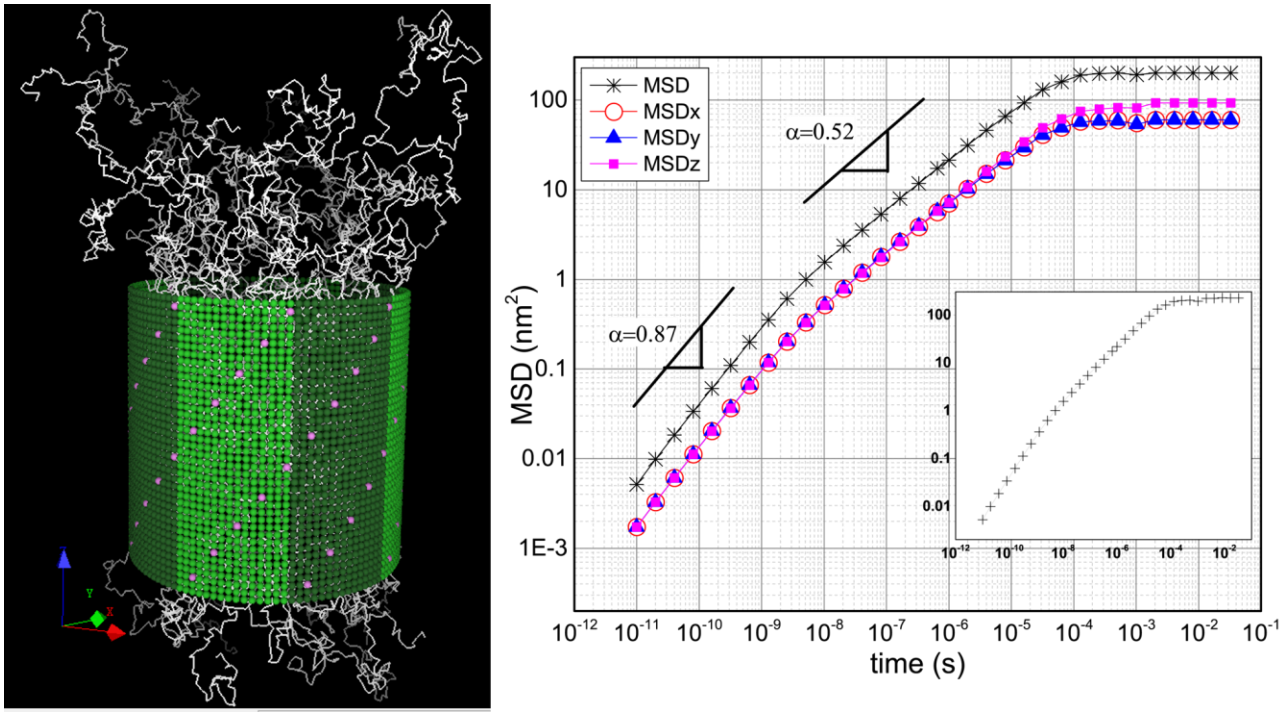


Figure 4.18. Left: The mutated NPC (ΔF) after equilibration. Because there is no hydrophobic affinity between the mutated monomers, the pure repulsions derive chains out of the NPC channel. Right: The MSD of mutated chains (ΔF) inside the NPC. Inset: The MSD of mutated chains (ΔF_ΔCharge). The patterns and values of MSD are identical in both cases.

4.10.3.1.b The effects of confinement on the MSD of FG-repeats inside the NPC

Next, I focused on the geometrical confinement factor. The confinement stems from the wall-particle repulsion. Depending on the wall geometry then, the degree of confinement can be quantified. For a cylindrical geometry, one may crudely quantify the degree of confinement as the ratio of the wall height to the radius of curvature, $\frac{h}{R}$. It the follows that degree of confinement for this geometry varies between two extremes of infinite (for $R=0$) and zero (for $R=\infty$).

In the case of the NPC channel, there is a physical limit on the minimum diameter of the channel in that all FG-repeat must be physically fit therein. Thus, at one extreme, the channel diameter has the minimum value, for which all of the FG-repeats can be fit and undergo thermal fluctuations without getting stuck. I determined the minimum permissible value of channel diameter by systematically reducing the diameter and monitoring the time-evolution of the FG-repeats. It was found that $D_{\min}=41nm$. Below this diameter, monomer-monomer penetration occurs (because of extremely limited space), leading to errors in the simulation.

I designed different artificial NPCs with diameters changing from 41 nm to 100 nm, and the same height as the native NPC (Fig. 4.19). Since the purpose of this section is to analyze the effect of

degree of confinement on the end-tethered FG-repeats, the only parameter to be changed is the degree of confinement. Thus, in each artificial NPC the FG-repeat compositions and numbers remain the same as those in the native NPC.

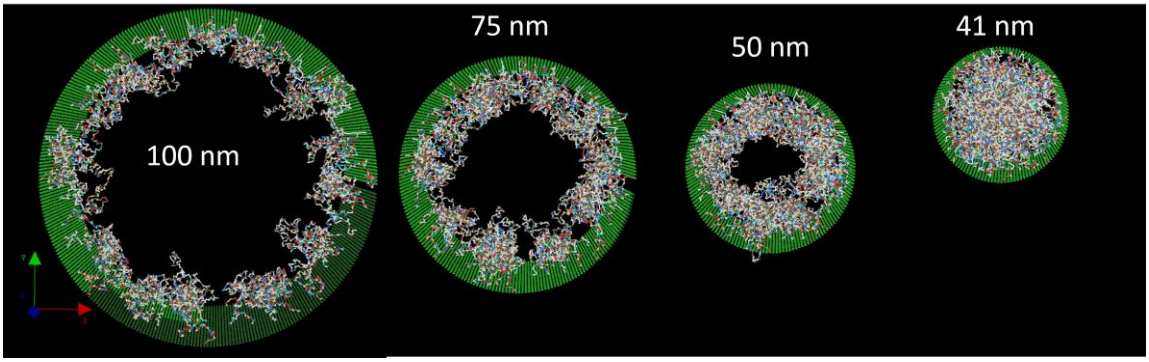


Figure 4.19. Artificial NPCs are made by increasing the diameter of the pore. Sample representation of four different diameters after equilibration. The diameter of each channel is indicated above it. Different colors have the same meaning as before (see Fig. 3.17).

The averaged MSD over all FG-Nups was calculated in each artificial NPC (Fig. 4.20). The patterns of the time-dependent MSD inside artificial NPCs are almost identical to that in a native NPC with two distinct phases before plateau (Fig. 4.15). However, as the degree of confinement decreases (the channel diameter enlarges), the values of the MSD increase and the plateau shifts up, confirming the effect of geometrical confinement on diffusion. Interestingly, it appears that beyond the channel diameter of D = 80 nm, the rate of the rise in the MSD plateau slows down. This implies that within the NPC, the reduction of the degree of confinement intensifies the amplitude of diffusion in a saturating manner. The more we deplete the confinement, the less the rate of the increase of the plateau height would be. Thus, the height of the plateau does not increase linearly with decreasing confinement. This is important, because it explains to what extent the channel confinement plays a role in shaping the permeability barrier inside the native NPC.

In particular, dissecting the role of confinement is significant in regard to NPCs of different species. It is known that the NPC channel diameter is strongly species-dependent such that it changes from 38-40 nm in yeast (Yang1998, Alber2007a) to about 90 nm in Xenopus oocyte (Eibauer2015), i.e. a variation of more than 100%. Therefore, the degree of confinement posed to the FG-repeats strongly depends on the species. If the degree of confinement plays a significant role in shaping the conformational behavior of FG-repeats, then the formation of selectivity barrier might be unique to each species.

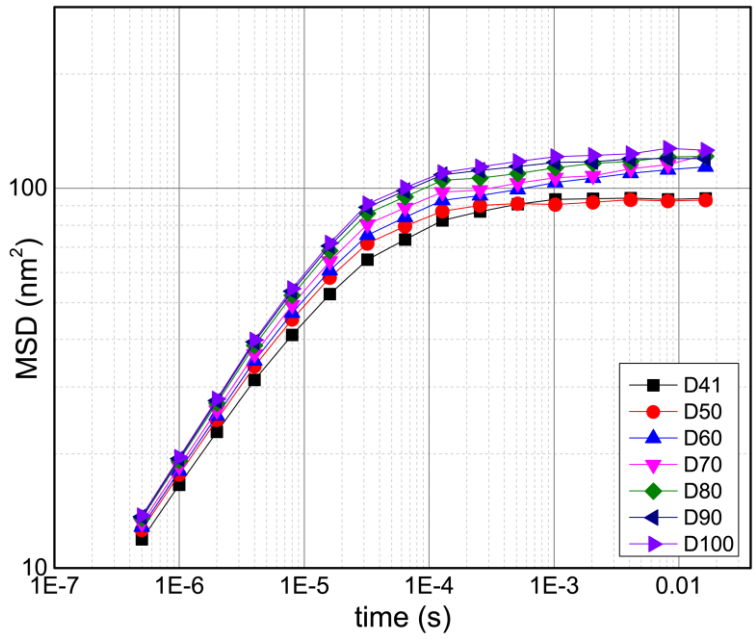


Figure 4.20. Changing of MSD vs diameter of the channel as a measure of the degree of confinement. The legend shows the diameter of the artificial NPC channel.

To find the extent of the confinement effect observed in Fig. 4.20, I designed an extreme, artificial NPC with no confinement. At this extreme, the channel has an infinitely large diameter, i.e. it is a flat wall. Moreover, to remove all aspects of the wall effects, the flat surface was made completely permeable to FG-repeats. Thus, while FG-repeats are end-tethered to the flat surface, they do not feel any wall repulsion and can freely cross the wall with no enthalpic or entropic penalty (Fig. 4.21). Nevertheless, because of the adhesive hydrophobic interactions inter- and intra-FG-repeats, the monomers generally aggregate on one side of the surface and barely cross the permeable wall (Fig. 4.21).

Similar to the extremely confined FG-repeats (D = 41 nm in Fig. 4.20), the MSD of fully non-confined FG-repeats varies nonlinearly in time and reaches a plateau (Fig. 4.21). Here, the plateau height is about 164 nm^2, showing an increase of more than 60% compared to the extreme confinement.

These results suggest that in the context of the NPC, the presence or the absence of geometrical confinement does not change the pattern of the FG-repeats' self-diffusion. Instead, the confinement only affects the amplitude of the self-diffusion, so that the degree of confinement clearly alters the values of MSD and the plateau height (Fig. 4.21).

Notably, both the composition and the confinement affect the MSD values, and thus the amplitude of FG-repeats' self-diffusion (Fig. 4.18 and 4.21). It is, however, astonishing that the composition has more prominent effects than the confinement. A total depletion of the confinement in an extreme case of a channel with an infinitely large diameter and permeable wall would result in only 60% increase in the MSD of FG-repeats. However, the mutation of F>S inside the native yeast NPC channel would introduce more than 200% increase in the

MSD. It is therefore reasonable to state that the variation in the NPC dimensions across different species would not give rise to dramatical change in the microdynamics of FG-repeats. Instead, the microdynamics of FG-repeats is exceedingly more sensitive to the hydrophobic composition of FG-repeats.

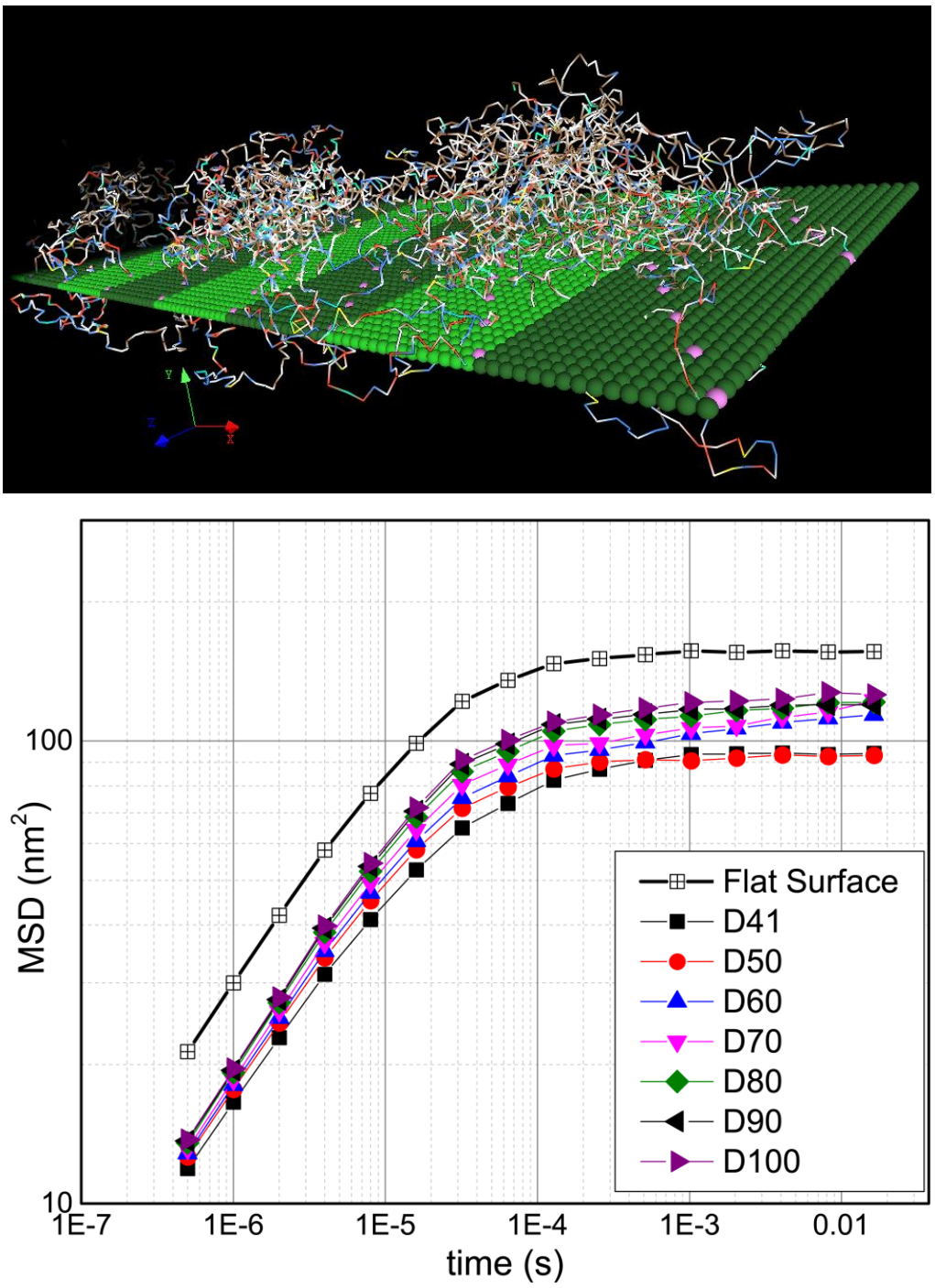


Figure 4.21. The FG-repeat domains are end-tethered on a perfectly permeable flat surface. There is no geometrical confinement felt by FG-repeats in this system. FG-repeat can also pass through the wall without any enthalpic or entropic cost. Top: The representation of the simulation after equilibration. Different colors have the same meaning as before (see Fig. 3.17). Bottom: The MSD vs time for flat surface. MSD of different diameters are repeated here for reference.

4.10.3.1.c The effects of end-tethering on the MSD of FG-repeats inside the NPC

The last factor that is reasonably playing role in determining the microrheology of FG-repeat domains is the end-tethering. Following the NPC attributes in terms of FG-repeat composition and concertation, here I designed two different experiments to scrutinize the effects of end-tethering. In the first experiment, the grafting points of the end-tethered FG-repeats are released, so that FG-repeats can freely diffuse inside the channel. As the channel is not closed in the z-direction, the now free FG-repeats can diffuse through cytoplasmic and nuclear entries out of the channel. To ensure that the concentration of FG-repeats remains the same as inside the native NPC, periodic boundary conditions under the minimum image convention (Allen1987) are applied in the axial direction at two ends of the pore. Therefore, if a chain leaves the pore from one entry, an identical chain would enter from the opposite entry. Effectively, this represents an infinitely long channel, having the same concertation of FG-repeats as in the native NPC.

As the results are depicted in Fig. 4.22 show, MSDs in x- and y-direction reach plateaus, while the MSD in z-direction increases so rapidly at long times. The value of the overall MSD in the longest lag time is two orders of magnitude bigger than that in the native NPC (10^4 nm^2 vs 10^2 nm^2).

Similar to the native NPC, diffusive motions of free FG-repeats undergo three different phases in time. The first two phases are entirely comparable to those in the native NPC with almost the same values of exponents. Curiously, the onset of second and the third phase remains the same as in the native NPC ($t = 10^{-8} s$ and $t = 10^{-4} s$).

The conspicuous departure from the native NPC emerges at very long times (the third phase), but only in the z-direction. The diffusional motion of the nontethered chains inside the long channel approaches the free diffusion in the z-direction with an exponent of 0.96. This is because after sufficiently long times chains start to disentangle and freely diffuse in the z-direction.

Notably, the FG-repeats' motion in the x- and y-directions is still confined, and thus the diffusion in those directions is anomalous as can be seen from the corresponding MSD plateaus. Therefore, the anomality in x- and y-direction is purely due to confinement, as the only disparity between these two direction and z-direction is the confinement. Noteworthily, the overall diffusive motion of the polymers in long times mimics free diffusion, meaning the diffusive motion in the z-direction dominates in long times. This makes sense because the diffusion in 3D is a scalar sum of three 1D diffusions in thee orthogonal directions.

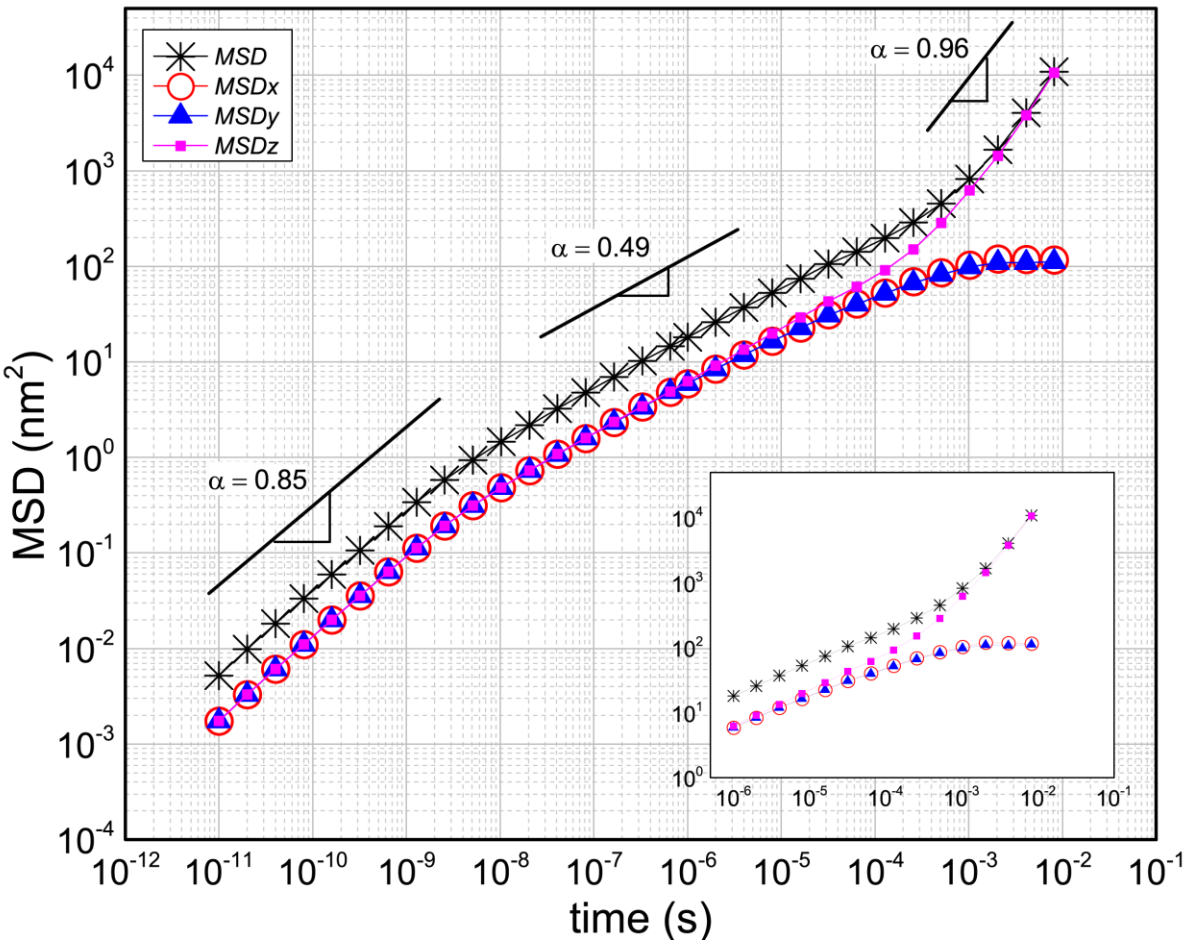


Figure 4.22. The MSD for free FG-repeats inside the NPC. The periodic boundary conditions in the axial direction guarantee that the FG-repeats concentration will be kept constant with the same value of that in the wild-type NPC. Inset: The zoom-in on the long-times domain of the MSD. The difference between MSDs of orthogonal can be seen more visibly.

This result shows that even if chains are confined in two directions, their MSD will eventually follow free diffusion, as long as their motion in the third direction is not constrained. In this regard, it is important to bear in mind that any disorder, i.e. deviation from the conditions of classical free diffusion, would affect the normal diffusive motion of the particle. The effects of the disorder can be classified under two main categories (Bouchaud1990): i) Only the value of MSD and transport coefficients (such as velocity, diffusion coefficient, shear viscosity, and thermal and electrical conductivity) is changed. ii) The underlying laws of Brownian motion is changed, e.g. the MSD varies nonlinearly in time.

Both confinement and end-tethering are examples of the former disorder, as they drive the time-dependent MSD into the nonlinear regime; they change the Brownian law of free diffusion. However, the extent of their effects is not the same. The end-tethering has markedly more influence

on the MSD than confinement. The combined results of this experiment (Fig. 4.22) and the grafted FG-repeats on the permeable flat wall (Fig. 4.21) clearly demonstrate the more powerful influence of end-tethering on the diffusional motion. In the flat wall experiment, there is absolutely no confinement in any direction, yet the motion of chains is subdiffusive. In the infinitely-long-channel experiment, there is no end-tethering, but there is confinement in both x- and y-direction. The overall motion of the chains approaches free diffusion.

Therefore, the end-tethering plays exceedingly a more active role in shaping the physics of diffusive motion than any other factors investigated in previous sections. In the context of the NPC, it implies that the microrheology of the selectivity barrier is mostly influenced by the end-tethering factor than by the composition of FG-repeats or the size of the channel. It is known that there are disparities in both composition of FG-repeats and the size of the NPC channel across different species. The results of our model, however, suggest that the microrheology of FG-repeats inside the NPC should be the same across different species.

In the second experiment to investigate the effects of the end-tethering factor, nontethered FG-repeat domains are constructed in a free 3D space with the same concentration as inside the native NPC. Here, I made a cubic simulation box with periodic boundary conditions in all directions under the minimum image convention (Fig. 4.23), similar to the common simulation box used in molecular dynamics.

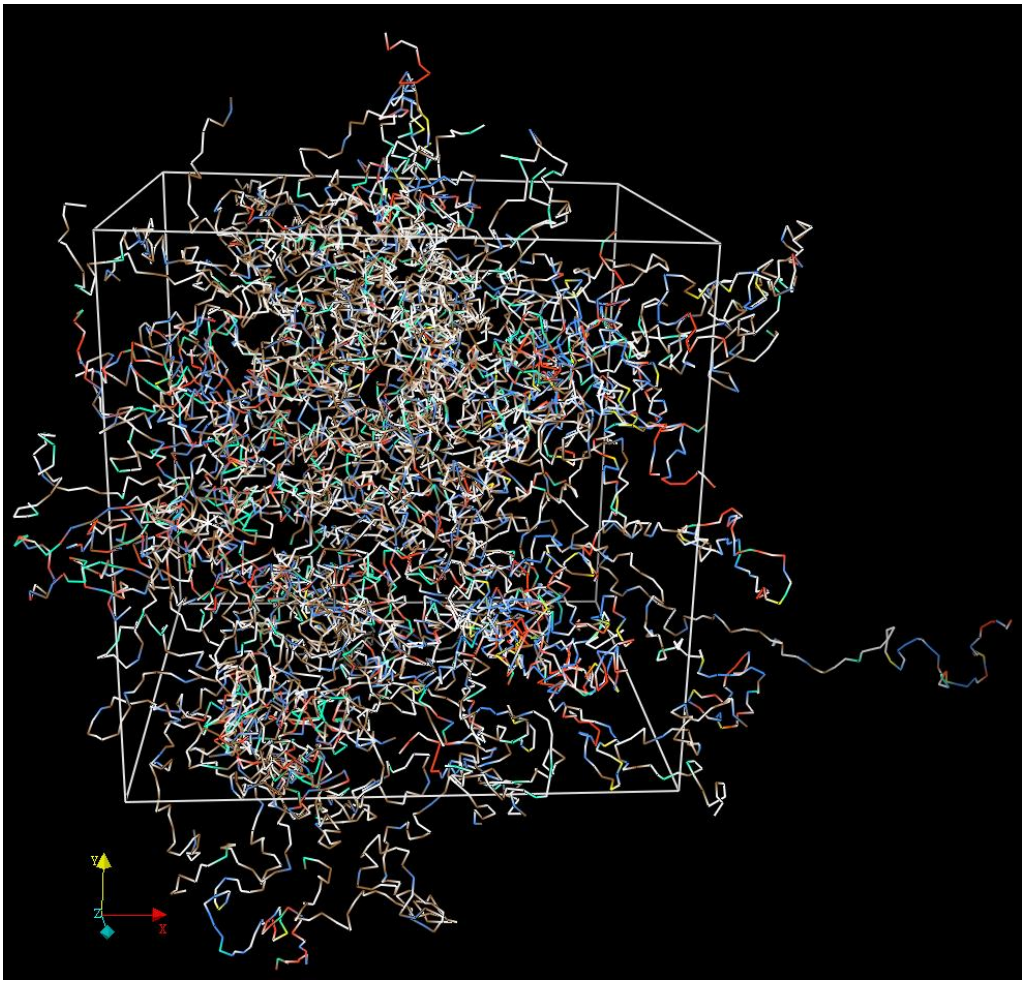


Figure 4.23. The simulation box of free FG-repeats in 3D. As before, different colors have different meaning (see Fig. 3.17).

From what discussed in the long channel experiment, admittedly, the MSD should eventually follow a true free diffusion. It can be seen from the MSD results (Fig. 4.24) that this is indeed the case. Similar to the previous experiment, there are three phases in the patter of the MSD. The motion of free FG-repeats in the first two phases is subdiffusive, with the exponents being 0.87 and 0.53, slightly larger than their corresponding values in the long-channel experiment. The exponent in the last phase is almost one, which is a signature of free diffusion. However, the maximum value of MSD in this experiment reaches an order of magnitude larger than that in the long channel (Fig. 4fxMSDnonthet). This is because the diffusion in x- and y-direction is nonconfined, and equally as z-direction contribute in the overall MSD.

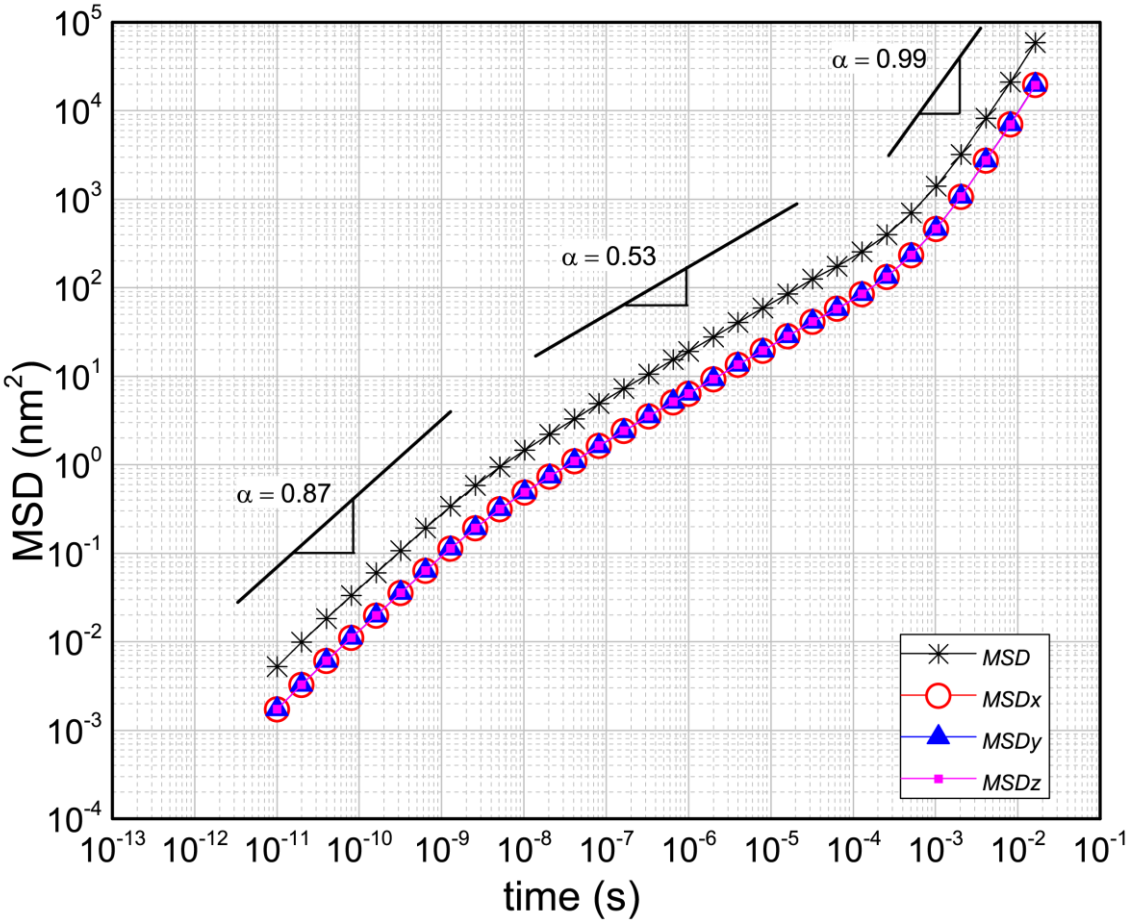


Figure 4.24. MSD for free FG-repeat domains with the same concentration and composition as inside the native NPC. Period boundary condition is used to maintain the concentration.

Expectedly, there is no difference in the MSD of three directions. There are convincing physical and statistical reasons for the strikingly indistinguishable values of MSD in three directions. Physically, three directions are identical and there is no imposed preference for the random movement in any direction, except the temporary preference that is naturally posed by the inter- and intra-chain transient interactions. Statistically, every single point in the MSD curve is averaged over tens of millions of samples of squared displacements (number of samples varies from 10^9 for the shortest lag time to 10^8 for the longest lag time), smoothing out all fleeting disparities. After all, the values of MSD in three directions are not mathematically the same, but the tiny differences cannot be distinguished on a log-log scale plot.

Notably, the subdiffusive power law of $\alpha = \frac{3}{4}$ has been observed in many semiflexible biopolymeric networks, including F-actins (Amblard1996, Palmer1999), microtubules (Caspi2002), and Amoeba proteus (Rogers2008). In two independent and pioneering studies,

Morse (Morse1998) and Gittes (Gittes1998) showed that the power law of $G^*(\omega) \propto \omega^{3/4}$ is a natural characteristic for the complex shear modulus for semiflexible polymers in high frequency domain. According to the GSER (eq. 4.11), one can extend the 3/4 power law to the MSD at short times, $MSD \propto t^{3/4}$ (Palmer1999). It means that in the absence of any subdiffusion-drive factors, the MSD of an aggregation of semiflexible polymers at short times deviates from normal diffusion. This is a natural characteristic of semiflexible polymers undergoing thermal fluctuations.

The disparity in the power law of FG-repeats perhaps arises from mechanical structure of FG-repeats. FG-repeats are very flexible chains, with the finest possible natural structure as a linear sequence of residues (see Table 3.1 and the definition of flexible and semiflexible in the previous Chapter). This extreme structural flexibility shapes the diffusion of FG-repeats in long times differently than that in semiflexible polymers. While this subject is beyond the current dissertation, it is worth a detailed investigation.

Importantly, during the first two phases, the pattern of the MSD remains almost the same across all experiments. Initially, the motion of FG-repeats is subdiffusive with the power law of $\alpha \approx 0.85$. It is then followed by a higher degree of subdiffusion with a reduced power law of $\alpha \approx 0.50$. At very short times, the monomers essentially diffuse so small that they do not feel anything except their adjacent monomers on the same chain. Presumably, the physical linkage between two adjacent monomers is the only factor preventing a monomer from experiencing free diffusion at short times. However, since the time is so small, the preventing effect is not so strong, explaining the exponent of 0.85 in the first phase.

In the second phase, polymers yet do not feel the major subdiffusive factors, like composition, confinement, and end-tethering. This is exactly why the MSD in all experiments reveals almost the same value. Conceivably, the main factor in this phase is the transient entanglement of chains with each other, introducing a significant deviation from linearity in the MSD as manifested by the power law of 0.50.

From the results presented here, it can be seen that the effects of the major subdiffusion-driver factors emerge only at times > 10^{-4} s (this is the onset of the third phase). Believably, FG-repeats begin to disentangle at this time. In the context of the NPC this is important because if the FG-repeats remain entangled, it would compete against an efficient and high throughput of cargo across the channel.

4.10.3.2 The viscoelastic response of FG-repeat domains

Finally, from the computational microrheology based on the many-particle tracking method without external probe as described in previous sections, I calculated the viscoelastic response of FG-repeat domains in different scenarios (see following). Equations 4.13-15 are the main equations used to calcuate the frequency-dependent loss and storage moduli. Instead of measuring the response of an external probe, here I directly measure the response of all 7,520 five-residue long monomers in the FG-repeat domains over a wide range of frequencies (10^1-10^{11} *Hz*).

4.10.3.2.a The viscoelastic response of FG-repeat domains in the native NPC is strongly frequency-dependent.

As the result in Fig. 4.25 shows, the viscoelastic response of FG-repeat meshwork is strongly frequency-dependent with a *gradual* crossover between the loss and storage moduli. The gradual crossover takes place over a long change in the frequency (three orders of magnitude), from 10^5 *Hz* to 10^8 *Hz*, with the middle value of 10^7 *Hz*, which can be taken as the crossover frequency, γ_{c0}. The crossover frequency is used for investigation of structural modifications (da Costa2005). Above this frequency, viscous behavior is predominant in the FG-repeat meshwork, while below γ_{c0} elasticity is dominant (Fig. 4.25). The gradual crossing-over of G' and G'' in the frequency domain suggests that the rheological structure of FG-repeats modifies at a relatively slow pace during the corresponding time period, in contrast with a typical polymer network whose crossover frequency is a single point in the log-log graph (see Fig. 4.35). This implies that in spite of a high rate of transport across the NPC, 1000 per second (Ribbeck2001), FG-repeat domains maintain a rheological stability during the crossover period.

The reciprocal of the crossover frequency is the characteristic relaxation time (Han2007, Kontopoulou2011). Thus, the relaxation time for the FG-repeat meshwork inside the native NPC in the absence of any external probe is $\lambda = 10^{-7} s$. Similarly, the crossover modulus, G_c, is the value of G' and G'' at the crossover frequency where loss and storage moduli are equal. At this frequency, the degrees to which the polymer meshwork behaves elastically or viscously are equal. From Fig. 4.25 it appears that $G_c = 0.1 Pa$ for the FG-repeat meshwork in the native NPC.

Notable, λ is one-fourth of the shortest reconstruction time and 1/80 of the longest reconstruction time (see Table 4.1). Not surprisingly, this means that the FG-repeat meshwork relaxes much faster in the absence of the perturbance introduced by an external agent like a cargo. Given that the reconstruction time is conceptually the same as the relaxation time in the presence of a cargo, the fact that $\lambda \sim \left(1/4 \text{ to } 1/80\right)\tau$ implies that microrheological analysis of FG-repeats based on the external tracer might not provide accurate results. This is because of a potential shift in the relaxation time due to perturbances created by the probe. Indeed, we observed that the shuttling of active cargos across the NPC slightly augments the relaxation time (decreases the crossover frequency) as seen in the Fig. 4.31 (not very visible on the log-log plot). Incidentally, this result also highlights the advantage of the microrheology without external probe, as in this method the thermal fluctuations of the meshwork are *indigenous* to chains, and not biased by any means.

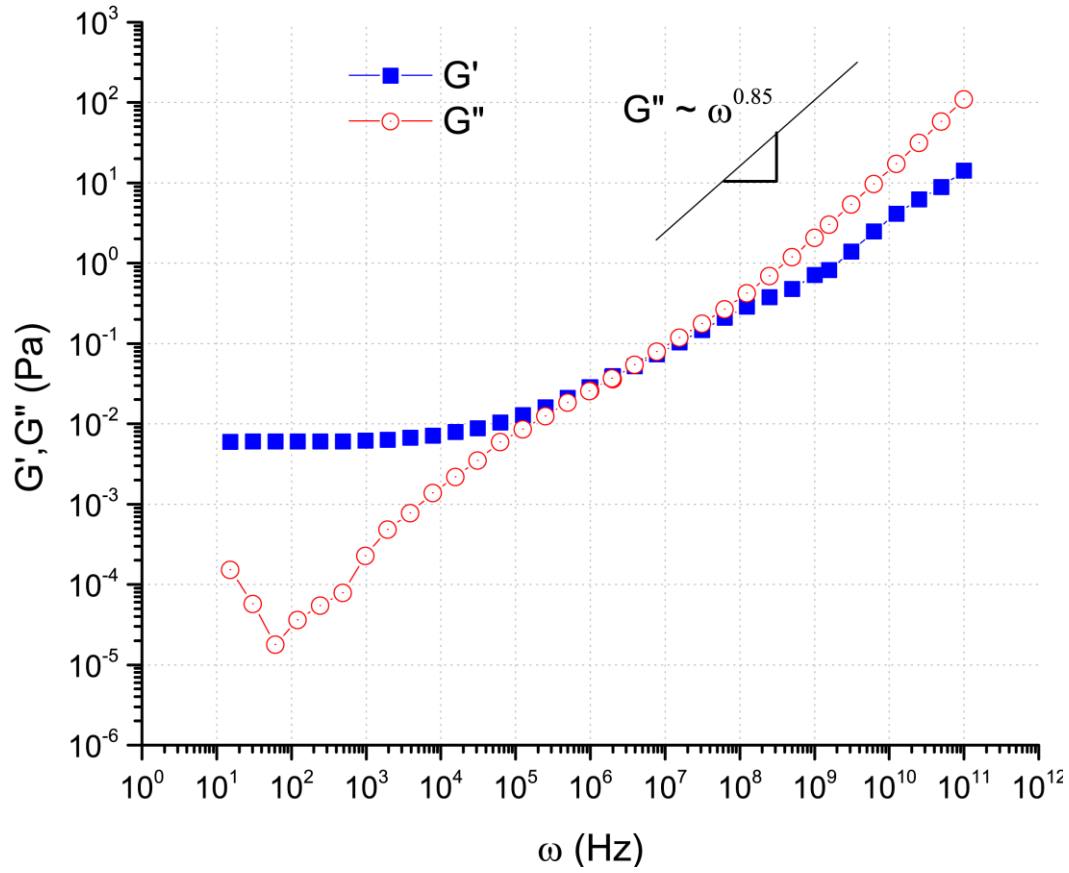


Figure 4.25. The mechanical response of the FG-repeat domains using the many-particle passive microrheology.

Significantly, λ is in the same order of magnitude as the time for nanometer-scale thermal fluctuations of a typical cargo with the diameter of 20 *nm*. Thermal fluctuations of such a cargo over the relevant length-scale of $l = 1-5\,nm$ occur on the timescale of $t = \frac{l}{6D} \sim 10^{-7}\,s$. It thus can be deduced that such a cargo while thermally fluctuating at sub-microseconds, would effectively experience FG-repeats as a *viscoelastic* liquid with the modulus of 0.1 *Pa*.

At frequencies above γ_{c0}, i.e. 10^9 – 10^{11} *Hz*, FG-repeat meshwork behaves more like a viscous liquid as the G'' is 10 times larger than the G'. Thermal fluctuations of ions and water molecules lie precisely within this range of frequency. For an ion or small molecule of *about* 3 *Å* in atomic diameter (the representative size for an ion and water molecule) diffusing in the cytoplasmic viscosity of *5 cP* (Janmey2006, Yang2006), the timescale of thermal fluctuations over the relevant length-scale of $l = 0.1-1nm$ is $10^{-9}-10^{-11}\,s$. This basically means that FG-repeat meshwork behaves as a vicious liquid for thermally agitated small molecules. This is consistent with the free diffusion of small molecules across the NPC. Notably, at this range of high frequencies, G" follows a power-law of $G'' \sim \omega^{0.85}$, which is

above the 3/4 universal power-law, indicating a stronger frequency-dependent response in FG-repeats.

More importantly, at low frequencies a plateau region emerges in the FG-repeats' elastic response and the elasticity dominates. The plateau modulus is a signature of a 'pseudo solid-like' response in FG-repeat domain. It is *pseudo* solid-like, because it does not last over all frequencies (Djabourov2013) and also for the most part, G' is not orders of magnitude larger than G", as it would be in an elastic solid (Krishnamoorti1997). Typically, the plateau region is pronounced in concentration solutions of polymers or in polymer melts (Deshpande2010).

The pseudo solid-like behavior of the FG-repeats is consistent with the functionality of selectivity barrier in that it clogs the central pore against the diffusion of nonspecific cargos. Within the accessibility of the current computational model, the pseudo solid-like response exists over the frequency range of 10^1 - 10^4 *Hz*. The corresponding timescale of 0.1 – 100 *ms* entirely covers the nucleocytoplasmic transport time window (Moussavi-Baygi2011a, Moussavi-Baygi2011b). Curiously, this suggests that over the timescale of nucleocytoplasmic transport, the meshwork of FG-repeats basically behaves as a pseudo solid-like material, i.e., it physically clogs the channel against nonspecific large cargos who diffuse at this timescale. Significantly, a cargo carrying hydrophobic patches on its surface is able to *locally intermesh* with this pseudo solid-like meshwork. Once engaged, however, due to its thermal fluctuations which happens at the timescale of 10^{-7} *s*, the cargo would effectively experience a viscoelastic liquid with a modulus of 0.1 *Pa*. From rheological viewpoint, the disparity between timescales is the key; it allows the active cargo easily penetrate into and diffuse across the FG-repeats. A nonspecific cargo, however, sees FG-meshwork as an elastic medium with pseudo solid-like behavior, because in the first place it cannot intermesh with FG-repeats.

It should also be noted that the 'pseudo solid-like' behavior of FG-repeats at low frequencies is only a comparative response for an otherwise pliable polymeric meshwork that has been shown to have a high degree of structural plasticity (Akey1995, Perez-Terzic1999, Milles2015, Sellés2017). When the mechanical response of FG-repeats at low frequencies is compared to that at high frequencies, the meshwork shows greater solidity than fluidity. Yet, the values of moduli are so small, rendering the network extremely flexible and posing a minimal free energy barrier to the diffusing active cargo.

In a viscoelastic material, the loss tangent, $\tan\delta = \frac{G''}{G'}$ is a measure of the degree of viscosity, in which loss angle δ varies between 0 and $\frac{\pi}{2}$ (Christensen2012). For a pure elastic solid, the stress and stain are completely in phase, and the loss angle is zero. For a pure viscous liquid, the loss angle is $\frac{\pi}{2}$ as the stress and strain are completely out of phase. This simply means that the material cannot store deformation energy, as the stress is entirely converted into the strain.

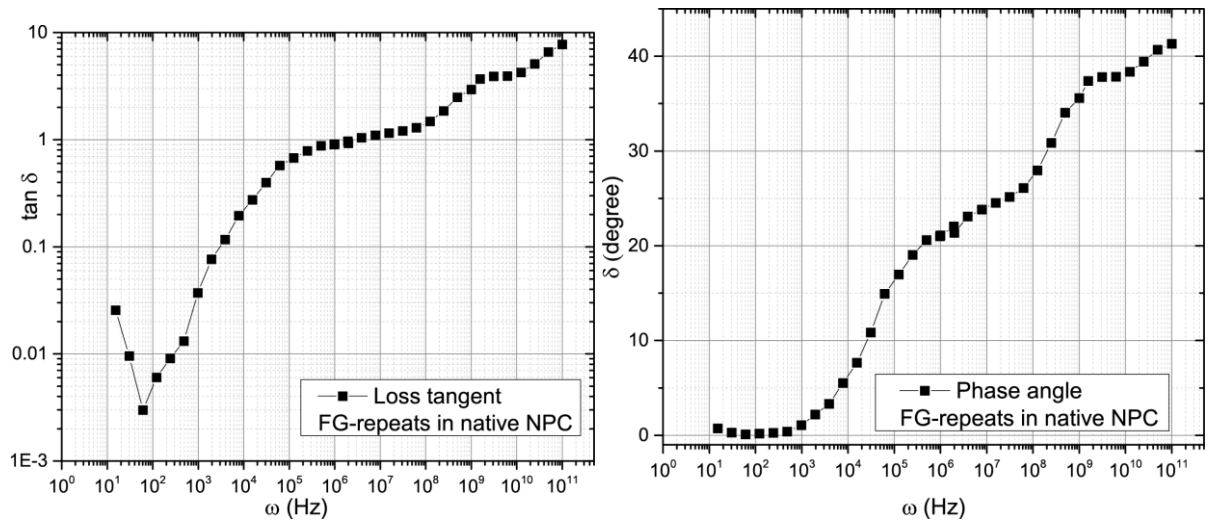


Figure 4.26. Left panel: The loss tangent vs frequency for FG-repeat domains in the native NPC. Right panel: The corresponding loss angle vs frequency.

Shown in Fig. 4.26 are the loss tangent and loss angle of FG-repeats in the native NPC. Clearly, the lower the δ, the more solid-like (elastic) the material is (Kontopoulou2011). At low frequencies the loss angle is almost zero, consistent with the dominant elastic modulus in the viscoelastic response. The maximum out-of-phase angle is $40°$ that happens at highest frequency accessible to this work (10^{11} *Hz*).

The loss tangent is a better representative of viscoelastic behavior, as it is the direct ratio of loss over storage moduli. At all frequencies below 10^7 *Hz* (equivalently, at all times above the relaxation time), the elasticity dominates. As frequency lowers, the elasticity is augmented. Below the frequency of 100 Hz, the loss tangent began to increase, but still the elastic modulus is one-two orders of magnitude larger than the loss modulus. This behavior is consistent with the general viscoelastic response of polymer melts (4.35). It is likely the onset of viscoelastic response at v*ery low* frequencies, which is beyond the capability of the current model and will be discussed later under the general viscoelastic response section.

Overall, with increasing/decreasing frequency, the viscous/elastic redolence of FG-repeats is vigorously amplified, indicating the strongly frequency-dependent viscoelastic nature of FG-repeats.

Noteworthily, using mechanical rheology and single-particle microrheology to study viscoelasticity of semiflexible F-actin filaments, it was found that the viscoelastic response of F-actins was strongly frequency-dependent (Xu1998). At low and high frequencies, F-actins show an elastic plateau modulus and a large loss modulus, respectively, confirming the pseudo solid-like behavior (G' > G") at low frequencies and viscous liquid (G" > G') at high frequencies (Xu1998). There are two hypotheses for the origin of the plateau in the elastic modulus of *semiflexible* polymers at low frequencies.

The first model called the curvature-stress model, is based on the assumption that in a tightly entangled solution, an individual actin filament is restricted to a tube formed by surrounding actin filaments (Xu1998). The dimeter of the tube is much smaller than the actin persistence length (Morse1998). In effect, restricting forces of the tube prevent transversal fluctuations of the actin filament (Xu1998). This restriction on fluctuations would in turn lead to a loss in filament's free energy (Isambert1996), and thus the filament is forced into a more ordered arrangement. In particular, the free energy cost of restricting a semiflexible filament to a tube is the origin of pseudo solid-like behavior, and thus the plateau modulus (Isambert1996). Similarly, the second model, termed the tension-stress model, assumes that the elasticity at low frequencies arises from restricted tangential motion of filaments (MacKintosh1995, Xu1998).

Both models share a similar underlying mechanism for solidification process at low frequencies: it is caused by the restricted motion, and thus more ordered arrangement, of actin filaments. Indeed, the extent of entropic effects on the elastic behavior were later extended and it was shown, based on a detailed microrheological investigation of the cytoskeleton in intact cells, that entropic effects were responsible for the cell elasticity (Deng2006). Now the question arises about the origin of the dominant plateau in the elastic modulus of FG-repeat domains at low frequencies: does a similar mechanism underlie the solidification of FG-repeats?

Remarkably, it has been suggested that in a solution of flexible chains the plateau modulus at low frequencies is due to the entropic elastic stress (Morse1998). It is also known from the theory of polymer dynamics that the elastic stress is dominant in a concentrated solution of polymers, and can be given by the change in the free energy for an instantaneous deformation (Doi1988).

More specifically, however, given the extremely flexible and highly dynamic structure of FG-repeats inside the NPC (Moussavi-Baygi2016), it seems like that neither of curvature-stress and tension-stress models are directly adaptable to FG-repeats. For one thing, the persistence length of FG-repeats is five orders of magnitude smaller than that in the F-actin (Table 3.1). Nevertheless, FG-repeats have restricted motions inside the native NPC due to the end-tethering, significantly reducing their entropies. We thus hypothesize that the end-tethering is the underlying cause for solidification of FG-repeats at low frequencies.

To further dissect this hypothesis, we need to recall from previous sections that: i) microrheology is founded on the basis of information encoded in the inherent thermal fluctuations of particles, ii) the MSD is the most fundamental property used to extract those information, and iii) it is the end-tethering factor that plays the most important role in shaping the MSD pattern of FG-repeats in the native NPC. Therefore, we would investigate the viscoelastic response of nontethered, i.e., free, FG-repeats.

4.10.3.2.b Nontethered FG-repeats in a cubic box with identical concentration as the native NPC form a predominantly viscous liquid

The end-tethering was found in the previous sections to be the most important factor in shaping the MSD of monomers. Thus, naturally the end-tethering should be the most influential factor in shaping microrheology of FG-repeats, as the MSD is the basis of microrheology.

Here I calculated the viscoelastic response of nontethered FG-repeats inside the cubic box (Fig. Fig. 4.23, 4.27, and 4.28). The concentration and compositions of nontethered FG-repeats is identical to that inside the native NPC. The period boundary condition is applied in x-, y-, and z-directions to remove the surface effects and to maintain the concentration inside the box.

As indicated by the loss angle and the loss tangent (Fig. 4.27), for the most part, nontethered FG-repeats behave predominantly as a viscous liquid, which is clearly different from the end-tethered FG-repeats in the native NPC. More interestingly, δ and $\tan\delta$ are less frequency-dependent in free FG-repeats compared to end-tethered FG-repeats in the native NPC (Fig. 4.27 vs 4.26). At all frequencies, $\tan\delta \geq 1$, meaning free FG-repeats always behave more as a viscous liquid rather than elastic pseudo solid. The peak value of the loss angle is $\delta \simeq \frac{\pi}{2}$ that occurs at high frequencies, for which free FG-repeats behave like a pure viscous liquid.

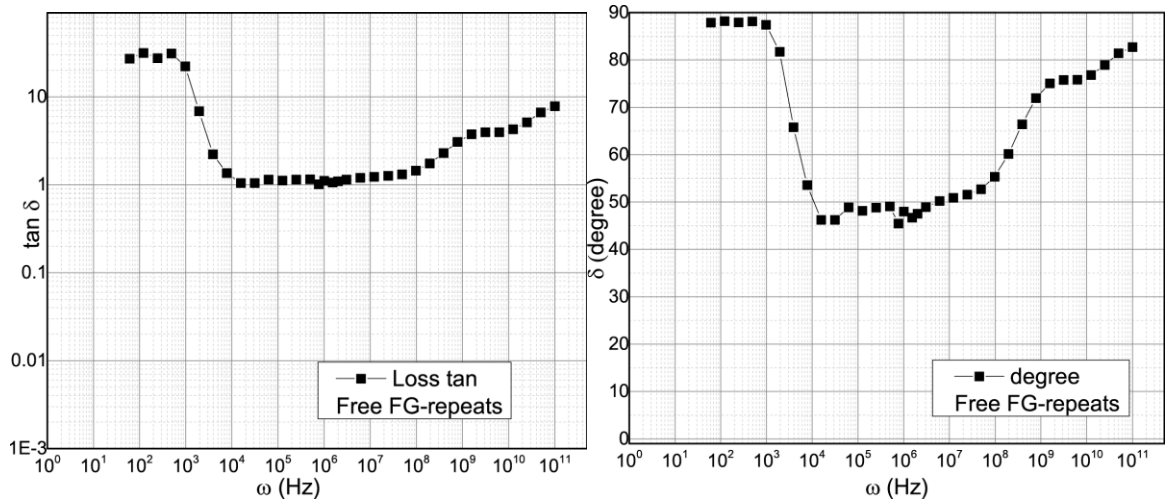


Figure 4.27. Left panel: The loss tangent vs frequency for free FG-repeat domains in the box with periodic boundary condition Right panel: The corresponding loss angle vs frequency.

The elastic modulus is less than or equal to the viscous modulus over all frequencies (Fig. 4.27 and 4.28). Similar to FG-repeats in the native NPC, there is a range of frequencies (10^4-10^8 Hz) over which G' and G" have almost the same values. Expectedly, over this frequency range, the system behaves as a true viscoelastic material. The smallest frequency for the true viscoelastic behavior is $\omega = 10^4\ Hz$, as opposed to FG-repeats in the native NPC where it is $10^5\ Hz$.

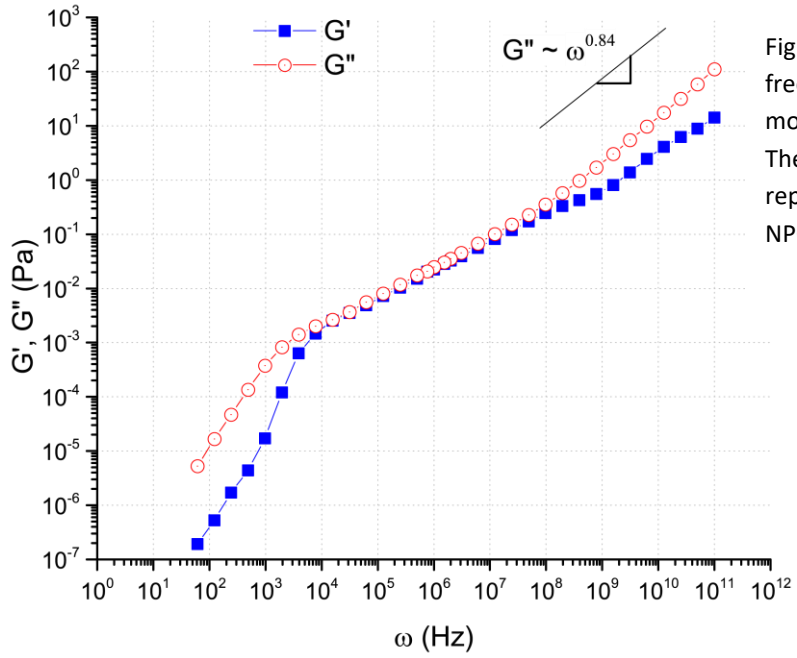


Figure 4.28. The viscoelastic response of free FG-repeats in an infinitely large vessel modeled via periodic boundary condition. The concentration and composition of FG-repeats are the same as those in the native NPC.

Beyond the true viscoelastic range of frequency, G' is always smaller than G" in free FG-repeats. At high frequencies, the viscoelastic response of FG-repeats is similar to that in the native NPC. Notably, almost exactly the same power-law of $G'' \sim \omega^{0.84}$ and the same values of G' ang G" are observed at high frequency domain, and thus, the same explanation is applicable.

More importantly is the viscoelastic response at low-frequency domain. The elasticity drastically drops, leaving the system as a pure viscous liquid. This is in an obvious contrast to end-tethered FG-repeats. For FG-repeats inside the native NPC, the frequency domain below 10^5 *Hz* is exactly where the entropic effects of end-tethering prevail and the entropically-driven elastic response emerges.

In addition to the absence of end-tethering factor in the free FG-repeats inside the box, there is also no confinement effect therein. Therefore, one may argue that it is not clear the predominantly viscous response of free FG-repeats at low frequencies (Fig. 4.28) is to what extent due to each of these two factors. To distinguish between the effects of each of the two factors, here I investigate the viscoelastic response of FG-repeats with only one factor, i.e., the end-tethering factor. This experiment is identical to the grafted FG-repeats on a perfectly permeable, flat wall in which there is no confinement factor, but there is end-tethering (Fig. 4.21).

As can be seen in the Fig. 4.29, the viscoelastic response of nonconfined, but end-tethered, FG-repeats is very similar to that in the native NPC. This further highlights the strong effect of end-tethering in determining the viscoelastic response of FG-repeats at low frequencies. Therefore, the aggregation of structurally identical biopolymers might either behaves as an elastic pseudo-solid or as a viscous liquid, solely depending on the existence or the absence of end-tethering. This

proves our hypothesis about the entropic origin of the dominant elasticity in FG-repeats at low frequencies, manifested as a plateau modulus in the viscoelastic response (Fig. 4.25 and 4.29).

Incidentally, this result also argues against those experiments that express FG-Nups in vitro and then generalize their results to the in-vivo conditions (Frey2006b, Frey2007, Frey2009, Ader2010). The closest computational experiment in the current work mimicking those experiments is free FG-repeats in the box (Fig. 4.23).

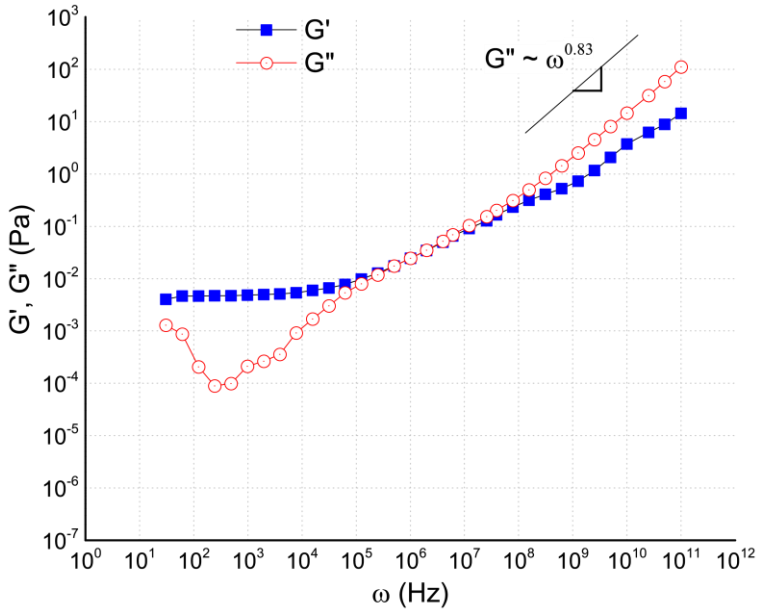


Figure 4.29. The viscoelastic response of end-tethered FG-repeats in a no-confinement setup. The FG-repeats are grafted on a perfectly permeable flat wall (see Fig. 4fxfreew).

Motivated by an in-vitro experiment on the expressed FG-Nups where all of Phe residues were mutated to Ser (Frey2006b, Frey2007), here I performed the same mutation on free FG-repeats in the box. Shown in Fig. 4.30 is the viscoelastic response of mutated (S>F) free FG-repeats (now, literally, SG-repeats!). The overall response is similar to that in wildtype free FG-repeats (Fig. 4.28) in that viscous modulus dominates over almost all frequencies. Noticeably, the elastic modulus undergoes more variability compared to the wildtype free FG-repeats. This can be explained in the context of elevated level of repulsions inter- and intra-FG-repeats compared to wildtype FG-Nups (recall from the last sections that in the absence of hydrophobic affinity, the monomer-monomer repulsions significantly increase).

Overall, microrheological investigation of FG-repeats suggests that most influential factor in determining the viscoelastic response of FG-repeats inside the native NPC is the end-tethering of FG-repeats.

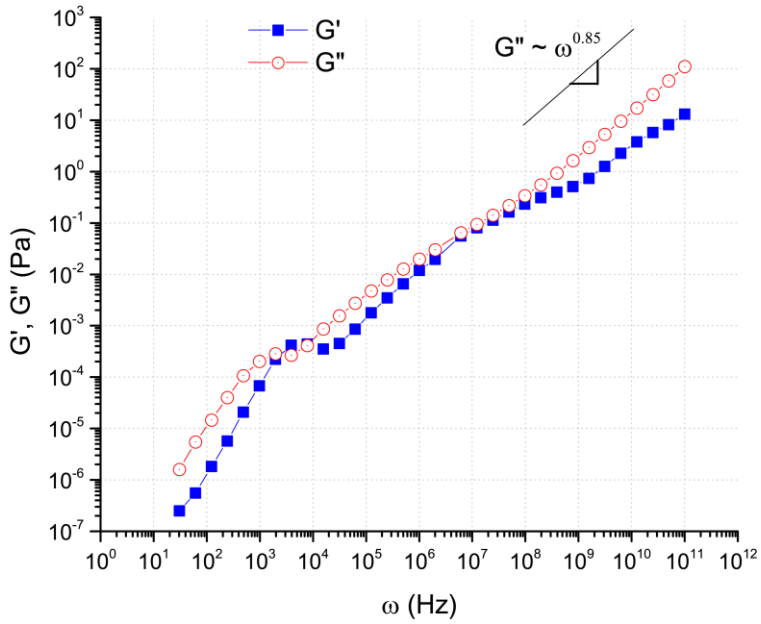


Figure 4.30. The viscoelastic response of the free mutated FG-repeats. All F residues are mutated into S. Therefore, there is no hydrophobic interactions in the mutated domains. The concentration of residues is kept constant via periodic boundary condition. The concentration is the same as that in the native NPC.

4.10.3.2.c Viscoelastic response of FG-repeats in the presence of shuttling active cargos: extending to ions and small proteins in the channel.

In a study by Caggioni et. al. (Caggioni2007) on viscoelasticity of the gellan gum, microrheology and bulk rheology were independently performed to compare the results of each of these two approaches. Particularly, in that study the gelation of the aqueous gellan gum, which has a net negative charge was examined via single-particle-tracking microrheology as well as the bulk rheology (Caggioni2007). It was found that microrheology underestimated the values of moduli compared to the bulk rheology. This was attributed to the electrostatic repulsion between the negatively charged tracer probe and the gel, leading to higher mobility of tracer inside the gel (Caggioni2007). The higher mobility amounts to larger MSD, which in turn predicts smaller values for the viscoelastic moduli (Equations 4.13-15).

The method of many-particle tracking without external probe, however, does not suffer from such a biased mobility as the monomers maintain their unbiased, indigenous thermal fluctuations. Nevertheless, a legit question arises about the microrheological investigation of FG-repeats in the current work: to what extent the continuous shuttling of active cargos inside the NPC might change the predicted viscoelastic response of FG-repeats? This is particularly important because the surface of active cargo is hydrophobically and electrostatically interacting with FG-repeats. Moreover, may the absence of some in-vivo factors, like ions and small proteins in the simulation modify thermal agility of the FG-monomers, resulting in different predicted values for viscoelastic moduli? Indeed, similar observaion has been reported in a similar study on the coarse-grained modeling of F-actin networks, where computational microrheology was employed to evaluate shear moduli of the F-actin network (Kim2009a, Kim2009b).

While in the current coarse-grained model it is not viable to include the solvent molecules as well as small proteins, it is possible to incorporate the continuous shuttling of active cargos, and investigate its effects on microrheology of FG-repeats. To do so, an active cargo with the nominal diameter of 20 nm is nudged into the FG-repeat meshwork from cytoplasmic side (Fig. 4.31). Whilst the active cargo diffuses across the NPC and interacts with FG-repeats, the minimal nudging force is applied on the cargo to steer the transport process in order to speed up the simulation. Once the cargo reaches the nuclear side it is released and the next cargo appears at the cytoplasmic entry, and the same cycle repeats. As this scenario insures a continuous shuttling of active cargos across the pore, it can provide insight into the transport-induced modification in the microrheology of FG-repeat meshwork, if any.

As shown in the Fig. 4.31, the continuous transport of cargos did not change the overall viscoelastic response of FG-repeats. In particular, the elastic modulus is very similar to that in the native NPC. It is remarkable that the behavior of FG-repeats at high frequency is predominantly like a viscous liquid, similar to FG-repeats without cargo and also free FG-repeats. Significantly, almost the same power law of $G'' \sim \omega^{0.86}$ is observed in the presence of nucleocytoplasmic transport. The relaxation time (reciprocal of the crossover frequency) only marginally increases compared to the native NPC in the absence of cargo inside the channel (not visible on the log-log plot).

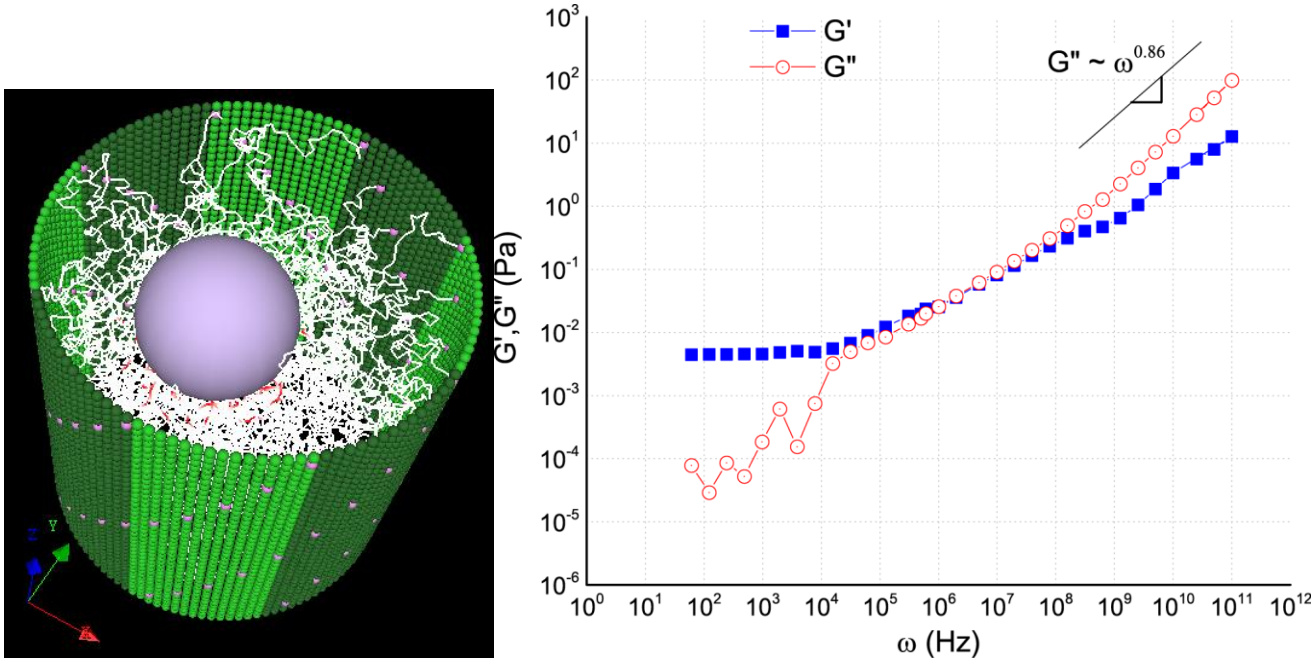


Figure 4.31. Right: The snapshot of the shutting globular cargo across the native NPC. Left: The viscoelastic response of FG-repeat domains in the presence of active cargo. The active cargo continuously shuttling across the channel.

One noticeable difference compared to the native NPC, however, is observed at low frequency domain for the loss modulus (compare Fig. 4.25 and Fig. 4.31): G" does not vary steadily with

frequency. The presence of the active cargo in the channel leads to fluctuations in the viscous modulus, because of transient interactions between NTR binding spots and FG-repeats. As discussed in the 'shear-thinning' in the next section, in a physically aggregated network of polymers, the viscosity arises from numerous transient bonds, collectively resisting the shear force. The active cargo-FG-repeats interactions compete with and transiently disrupt the inter- and intra-FG hydrophobic bonds. In addition to natural hydrophobic interactions of cargo with FG-repeats, because the cargo is nudged in this simulation, its rather fast directional motion would further contribute to a rapider rupture of FG-FG hydrophobic bonds, resulting in more fluctuations in the viscous modulus. Notably, the fluctuations in G" occurs on the same timescale of the cargo transport time (~10^{-3} s).

The active cargo is indeed much larger than ions and small proteins, and expectedly it should have more pronounced effects on the microrheology of FG-repeats than small molecules. Yet, a continuous shuttling of active cargos across the channel does not change the overall viscoelastic response of FG-repeats. This result suggests the presence of ions and small proteins in the channel would have minimal effects on the microrheology inside the NPC, implying that the values of G' ang G" found here should not change dramatically if one considers those factors in a coarse-grained simulation. This further confirms that the elastic response of FG-repeats is driven by entropic effects of the the end-tethering factor, rather than any other factor like the existence of cargos in the pore.

4.10.3.2.d Viscoelastic response of FG-repeat domains might be species-dependent over a specific range of frequency.

Using one-particle and two-particle microrheology for the entangled F-actin filaments of 500 nm long over the frequency range of $10^{-1} - 10^{3}\,Hz$, Liu et. al. found that the frequency-dependent loss and storage moduli varied from $10^{-3}\,Pa$ to $10^{-1}\,Pa$ (Liu2006). By increasing the length of filaments to 2000 nm, the moduli increased two orders of magnitude over the same range of frequency domain, confirming the effects of length of the filaments on its viscoelastic properties (Liu2006).

This observation raises the question of the role of FG-repeats lengths in shaping the microrheology inside the NPC. This is indeed a biologically relevant question, as it is well stablished that the mass and dimensions of the NPC are species-dependent (Peters2006, van der Aa2006, Maimon2012, Bui2013, Eibauer2015, Kosinski2016). It is thus natural to expect disordered domains in FG-Nups have different lengths in different species. Generally, the dimensions of the NPC in vertebrate are larger than those in the yeast, suggesting longer FG-repeat domains in vertebrates. Therefore, I chose the human NPC as the best-structurally-studied vertebrate with ample published data. Yet, unfortunately, there is no data from experiments on the exact locations of FG-Nups along the hNPC. Moreover, no data is available about the sequence of disordered domains within each human FG-Nup.

To overcome these deficiencies, I ran a comparative study based on the functional analogy of FG-Nups in the human NPC vs yeast NPC to predict the locations of FG-Nups in the hNPC (see Appendix B). In addition, I used three online disorder-predictor algorithms to predict the disordered regions of human FG-Nups based on the sequences of the entire FG-Nups (Appendix B).

FG-repeat domains in the human NPC are likely longer than those in the yeast, as the overall linear dimensions of the channel are more than 50% larger than the yeast NPC. It appeared, based on the predicted disordered domains of human FG-Nups that the median of the FG-repeats' length is nearly twice in the human NPC compared to the yNPC (Fig. 4.32).

As shown in the Fig. 4.33 the viscoelastic response of the human FG-repeats at high frequencies is very similar to that in the yeast, again with the same power-law of $G'' \sim \omega^{0.85}$. At low frequencies accessible to the current computational model, however, there is a clear distinction with the yeast FG-repeats in that the elasticity is no longer dominant, and also the plateau modulus is not observed. Note that the lowest frequency attainable here is in the order of 100 Hz, which is an order of magnitude larger than that in the yeast NPC (compare Fig. 4.33 with 4.25). This is because the size, i.e. total number of monomers, of the FG-meshwork in the hNPC is twice as large as that in the yeast NPC, yet the computational time is almost eight times higher.

The larger size of FG-meshwork would likely require longer time for the plateau region to appear in the elastic modulus. Moreover, the disordered domains of human FG-Nups in the simulation are only predicted and are not supported by lab experiment. This adds another layer of uncertainty to the calculated viscoelastic response, which cannot be avoided until the accurate sequence of disordered domains of human FG-Nups are known from wet lab experiments. Nevertheless, the results here can give a big picture of how the chains' length potentially leads to a shift in the viscoelastic response. It also implies that the viscoelastic response of FG-repeats might be species-dependent over a specific range of frequency, although one might not expect a dramatic change.

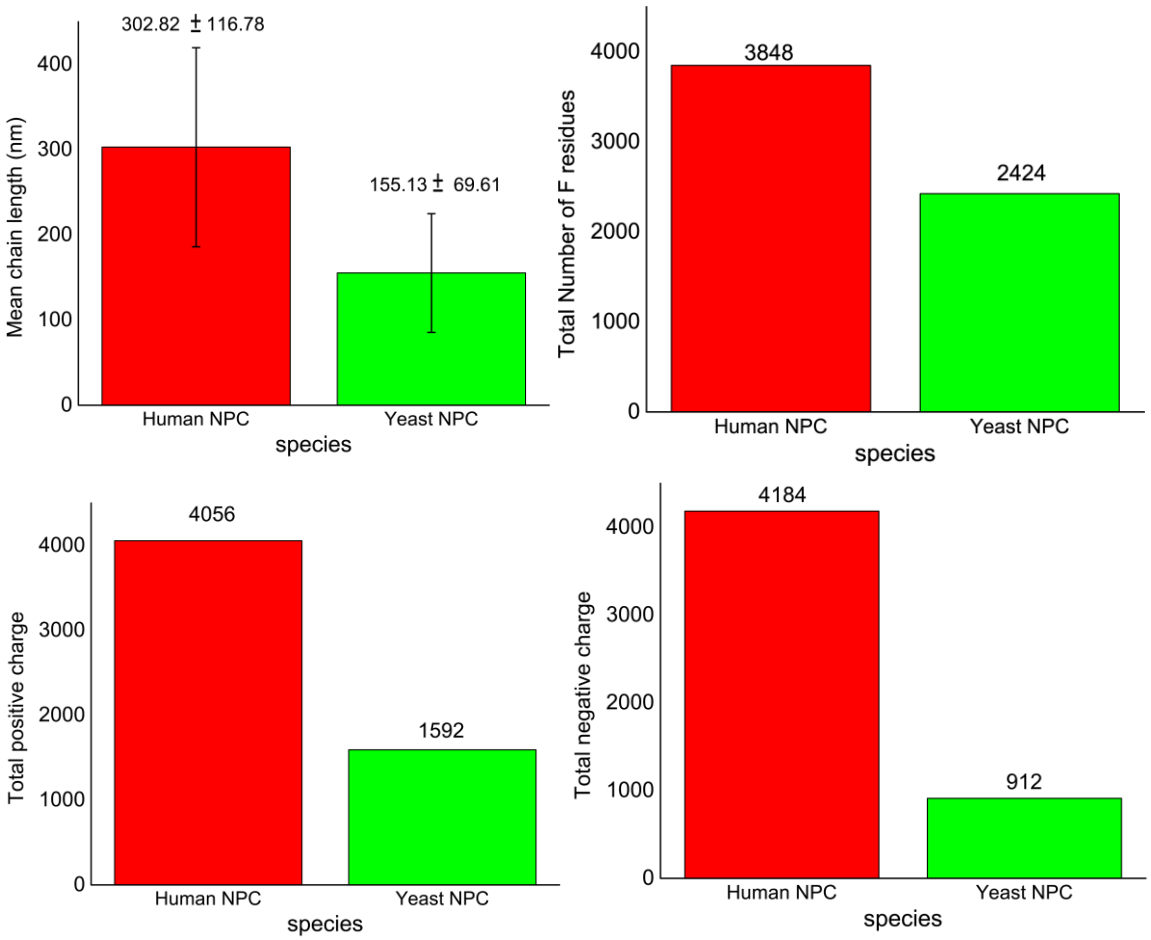


Figure 4.32. The physical properties of predicted disordered regions in the human FG-Nups are compared with those in the yeast.

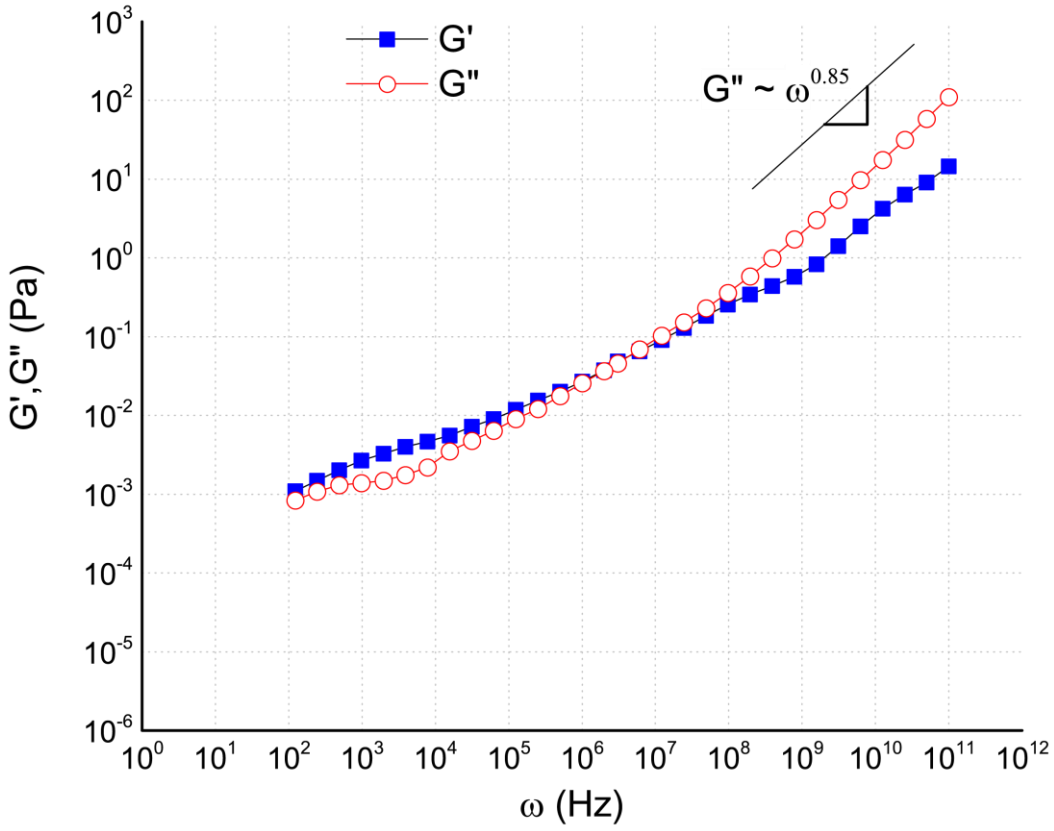


Figure 4.33. The viscoelastic response of predicted FG-repeat domains in human NPC.

4.10.3.2.e The viscosity of FG-repeats shows non-Newtonian behavior with shear-thinning characteristic; the molecular origin of shear-thinning from microrheology viewpoint

Viscosity is probably the most-commonly sought-after quantity in rheology, but it has qualitatively different properties in Newtonian and non-Newtonian fluids (Morrison2001). In simple terms, the viscosity is the resistance of the material flow to the shear stress (McQuarrie1973). For a Newtonian fluid, the shear stress, τ, varies linearly with the velocity gradient (Massey2005), i.e. the rate of shear strain perpendicular to the layers of fluid (Munson2014). The constant of proportionality is called the viscosity, or more precisely, dynamic viscosity $\tau = \eta \frac{\partial u}{\partial y}$ (Massey2005, Munson2014).

For a complex fluid, however, the viscosity is not constant, instead it changes by the rate at which the flow is driven (Morrison2001). It thus also has a more profound definition as (Shenoy2013):

$$\eta^*(\omega) = \eta'(\omega) - i\eta''(\omega), \qquad (4.17)$$

with $\eta'(\omega) = \frac{G''(\omega)}{\omega}$ being the dynamic viscosity as the real part of the complex viscosity, and $\eta''(\omega) = \frac{G'(\omega)}{\omega}$ the imaginary part of the complex viscosity. Fundamentally speaking, the complex viscosity carries the same level of information about the microstructure of the complex fluid as the complex shear modulus does. Nevertheless, since some phenomena, like shear-thinning and shear-thickening (see following) are defined solely based on the viscosity, it is convenient to calcuate the viscosity in addition to shear moduli in a fluid of interest.

The most common characteristic of a non-Newtonian fluid is the shear-thinning, i.e. the decrease in the viscosity with increase in the shear rate (Siginer1999, Morrison2001) (equivalent of frequency in our microrheological analysis). A rarer scenario for non-Newtonian fluid is when the viscosity increases with increased shear rate; this is called shear-thickening.

Shown in Fig. 4.34 are the real and imaginary parts of the complex viscosity of FG-repeat domains in the native NPC as well the free cubic box with periodic boundary condition. It can be seen that FG-repeat domains invariably show shear-thinning behavior, regardless of the physical or geometrical conditions. In other words, whether FG-Nups are end-tethered or free, or whether they are confined or not, the viscosity of FG-repeats generally decreases with increasing frequency. The most direct inference drawn from this observation is that FG-repeats are non-Newtonian in nature with the shear-thinning characteristic. Expectedly, at high frequency regime, the dynamic viscosity approaches a constant value and becomes less frequency-dependent. At this regime, the FG-meshwork behaves like a viscous fluid, as discussed in the previous sections.

Looking deeper at Fig. 4.34 reveals interesting insights into the effects of end-tethering on the viscosity at low-frequency domain. First, free FG-repeats initially show shear-thickening behavior at low frequencies, as opposed to end-tethered FG-repeats inside the native NPC that consistently show shear-thinning. Secondly, consistent with the results obtained from the shear moduli, at low frequencies the elasticity dominates in the FG-repeat meshwork inside the native NPC, while free FG-repeats form a predominantly viscous liquid.

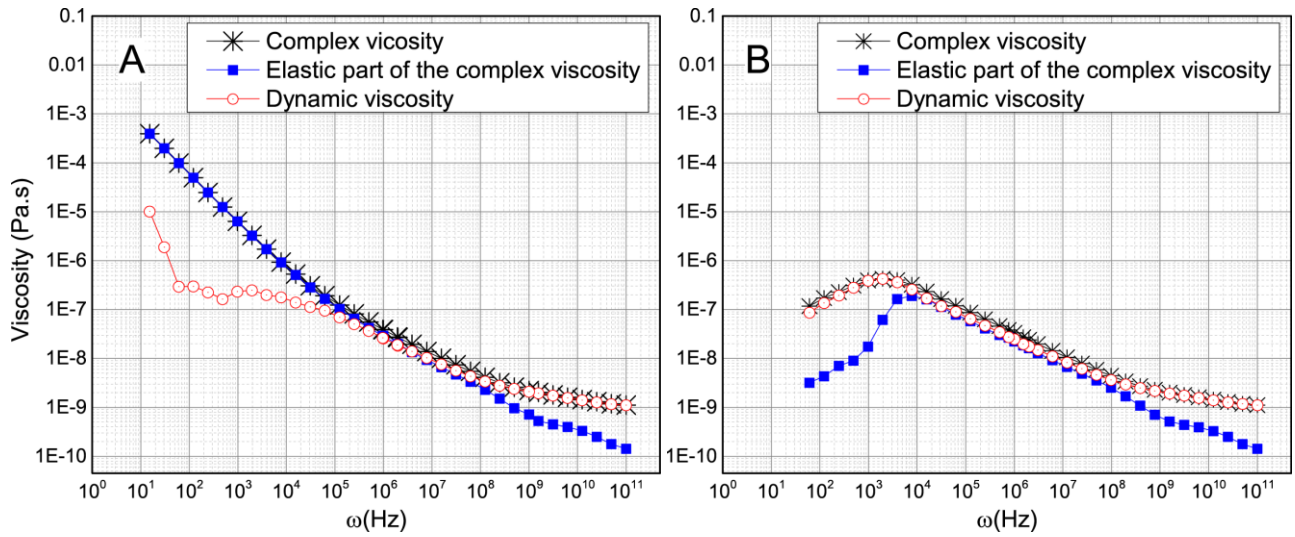


Figure 4.34. The non-Newtonian viscosity of FG-repeat domains shows a shear-thinning behavior with frequency. Shown are the complex viscosity along with its real (dynamic viscosity) and imaginary parts for A) end-tethered FG-repeat domains inside the native NPC, B) free FG-repeat domains in the cubic box with periodic boundary condition.

A biophysical aggregation of polymers, hydrogel or not, is formed by temporal bonds, like hydrophobic interactions and hydrogen bonds (Shenoy2013). Although the bonds are individually weak and short-lived, collectively they can form a stable assembly (Moussavi-Baygi2016). However, because of the dynamic nature of weak physical connections, an applied shear force can break the bonds (Shenoy2013). With an increasing rate of shear more bonds are broken, leading to a decrease in the collective resistance of the meshwork to the shear force, i.e., a reduction in viscosity. This describes the origin of the shear-thinning behavior at the bond scale in a polymer meshwork subject to shear force.

What described above was from a rheology point of view where it is conventional to apply shear force with a varying rate. Equivalently, from the microrheology viewpoint where there is no external shear force, at low frequencies (i.e., long times) more bonds are formed, resulting in an observed higher viscosity. At high frequencies, i.e. short times, fewer bonds are formed, ending up with a smaller viscosity being observed. This, incidentally, touches upon the elegance of microrheology; instead of invasively applying external shear force at varying rates, inherent thermal fluctuations are measured noninvasively at different frequencies.

4.10.3.2.f General viscoelastic response of a polymer melt: a comparison with FG-repeat domains

As mentioned before, obtaining the viscoelastic response over the entire range of frequencies is not possible with a single method or instrument (Chhabra2011). Thus, a common method is to superimpose the results of different experiments and instruments to obtain a mastercurve over all frequencies (Dealy2006). Based on this method, a typical viscoelastic response for a polymer solution is depicted in Fig. 4.35 (Deshpande2010).

Overall, the frequency domain can be divided into four different regions: i) very low, ii) low, iii) high, and iv) very high frequencies. While the first region is beyond the capability of the current computational model, the last region should be avoided in a Brownian dynamic simulation, as it violates the physics of GSER as well as the fundamental coarse-graining assumption of the negligible inertial effects (see Chapter 3 and also section "inertial effect" in this chapter). Therefore, we are limited to a specific range of frequencies in the current computational microrheological analysis so long as the GSER is employed in Brownian dynamics.

The low and high frequencies can be covered by the current model, as highlighted in the Fig. 4.35. It should be mentioned that the highlighted region is not the strict boundary of low and high frequencies; instead, it is just what I could reach in the current dissertation. It definitely can be extended to include more frequencies.

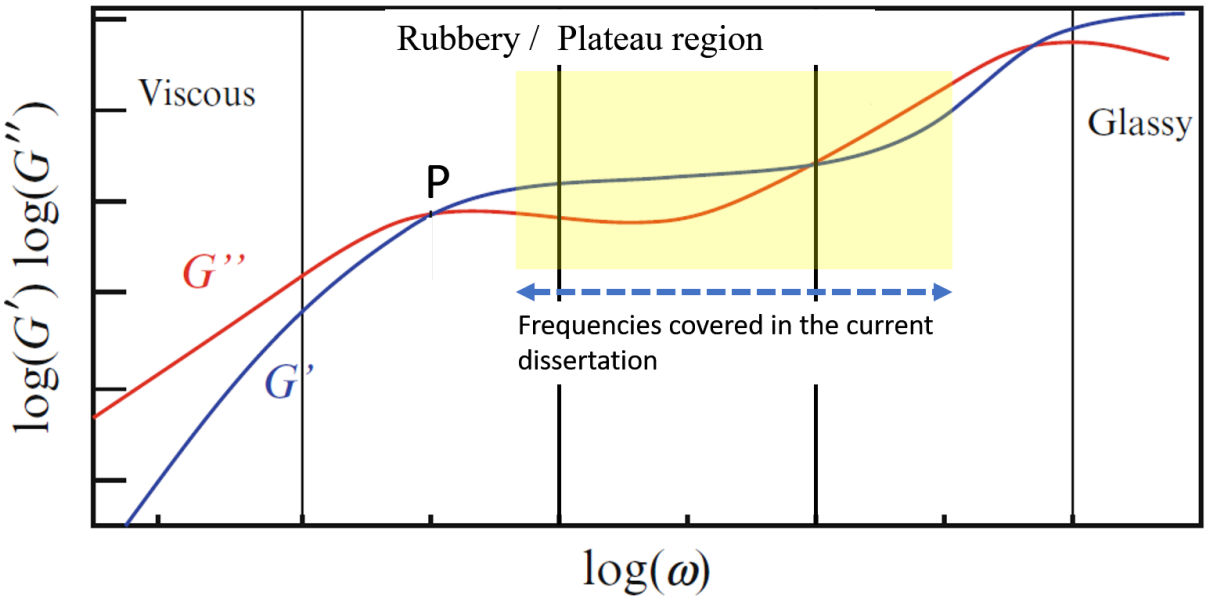


Figure 4.35. The mechanical spectrum of a typical polymer solution/melt over a wide range of frequencies. In this dissertation, we covered frequencies in the yellow zone. Adopted from (Deshpande2010).

At *very low* frequencies G' increases more rapidly than G" (Djabourov2013), but the value of the loss modulus dominates, indicating that the polymeric system is no longer solid or pseudo solid; instead it behaves as a viscous liquid (Osada2016). Dominantly viscous behavior ($G'' \gg G'$) is typical for a physical, but not chemically cross-linked, gel (Osada2016). The change from elastic to viscous can happen over a couple of decades in frequency (Deshpande2010). More broadly speaking, it is the ultimate behavior of every material, including solids, at the limit of $t \to \infty$ (Reiner1964). Incidentally, as we saw in Fig. 4.26, the viscos modulus of FG-repeats starts to grow rapidly at lowest frequency accessible to the simulation.

At higher frequencies, the behavior of the system is truly viscoelastic as a result of competition between viscous and elastic forces, which is largely governed by the entanglement (Deshpande2010). This region called rubbery/plateau region (Deshpande2010, Djabourov2013). The plateau region is perhaps one of the most important regions of interest occurs in this range of frequency.

At *very high* frequencies, the material tends to behave elastically and is referred to as glassy behavior (Deshpande2010), predominantly as a result of the ballistic motion of particles.

The Fig. 4.25 represents a portion of the yet-to-be-determined complete mechanical response of FG-repeats inside the native NPC that is covered in the current dissertation. It is, however, remarkable that the covered region of 10 orders of magnitude of variations, 10^1 – 10^{11} *Hz*, in the frequency domain (Fig. 4.25) matches very well with the highlighted region in Fig. 4.35. In other words, the calculated viscoelastic response in the current work closely follows the viscoelastic response of a general polymer melt/solution over the corresponding frequency domain.

The calculated viscoelastic response in Fig 4.25 can be possibly extrapolated to the very-low frequency region in the Fig. 4.35. The rapid increase in the loss tangent at the lowest frequency accessible to this work (Fig. 4.26) supports the idea of extrapolation. As it is seen from Fig. 4.26, the G" starts to go above G' as frequency decreases, consistent with the point **P** on the Fig. 4.35. It thus can be predicted that FG-repeats would behave as a viscous fluid at very long times (hours/days). An interesting implication is that the conformational behavior of FG-repeats highly depends on the timescale of observation. This in part explains why there are apparently contradictory hypotheses on the conformational behavior of FG-repeats inside the NPC.

4.10.3.2.g Viscoelastic response of FG-repeats does not follow the basic Maxwell model

The famous Maxwell model is a primitive model developed to describe the linear viscoelastic response of a complex fluid under the small amplitude shear stress, τ (Deshpande2010). The differential form of the Maxwell model is (Morrison2001):

$$\tau + \lambda \frac{\partial \tau}{\partial t} = -\eta \dot{\gamma} \,, \qquad (4.18)$$

where λ is the relaxation time, ηcis the zero-shear viscosity, and $\dot{\gamma}$ is the rate of strain. Using this model, the loss and storage moduli for a linear viscoelastic material subject to small amplitude oscillatory shear can be derived as:

$$G' = \frac{\eta \omega^2 \lambda}{1+\omega^2 \lambda^2} \quad \& \quad G'' = \frac{\eta \omega}{1+\omega^2 \lambda^2} \qquad (4.19)$$

Based on the microrheological calculation of the relaxation time of FG-repeat domains, γ_{c0}, we can obtain the prediction of the Maxwell model for the viscoelastic response of FG-repeats in the native NPC (Fig. 4.36). To do so, the value of the cytoplasmic viscosity was taken as the zero-shear viscosity.

An important prediction of the Maxwell model at very low frequencies is that $G' \sim \omega^2$ and $G'' \sim \omega$, signifying the viscous nature of the polymer melts at such frequencies. This general prediction is consistent with overall behavior of polymer melts at very low frequencies (see Fig. 4.35). Yet, the details of the predicted viscoelastic response are not consistent with a general viscoelastic response of a polymeric solution (Fig. 4.35).

It appears that the viscoelastic response of FG-repeats is fundamentally different from what is predicted by the Maxwell model. Aftercall, the response of a polymer melt is way more complex compared to the prediction made by simplistic models like Maxwell (Deshpande2010). This is because of the Maxwell model is oversimplified and does not take into account many factors like hydrophobic affinity, nonhomogeneous distribution of interacting groups along the chains, chains' concentration and length, and geometrical confinement and end-tethering. As we show, the latter is the major factor in shaping the viscoelastic response of the FG-repeats inside the NPC.

Curiously, in the absence of the end-tethering factor (Fig. 4.28), the prediction of the Maxwell model is consistent with the viscoelastic response at low frequencies. Consistent with the simulation results in Fig. 4.28, the Maxwell model predicts a predominantly viscous liquid at low frequencies. Again, this is a reiteration of the importance of the end-tethering in shaping the viscoelastic response of FG-repeats in the native NPC. The crude Maxwell model that is developed based on the linear viscoelastic response of simple polymers can be partially fitted to the free FG-repeats, but not to the end-tethered FG-repeats.

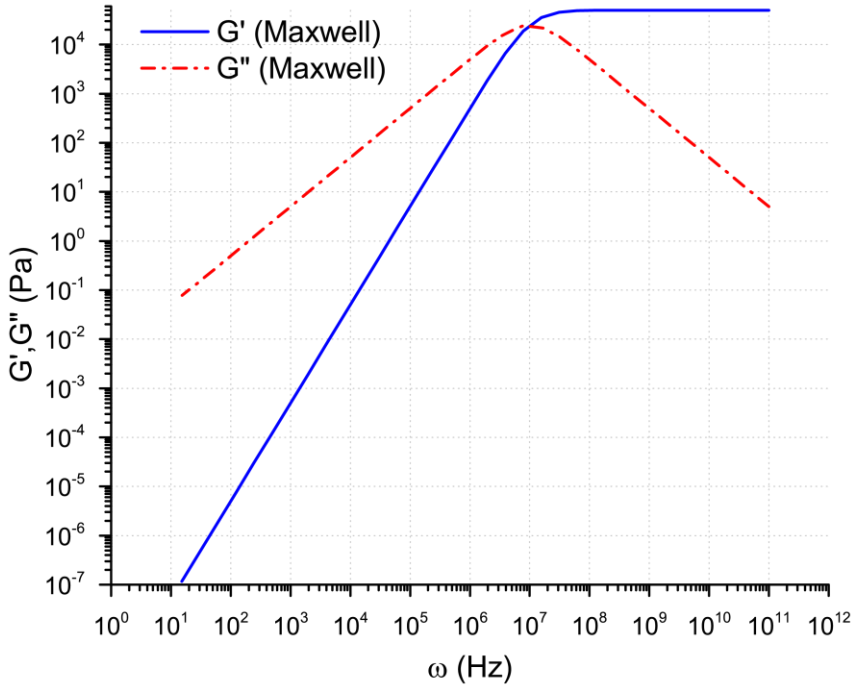


Figure 4.36. The viscoelastic response of FG-repeat domains in the native NPC as predicted by the Maxwell model.

There are also extensions of the Maxwell model with several relaxation times to give a better fit to the real viscoelastic response (Morrison2001). The number of relaxation times in the extended

Maxwell is arbitrary and depending on the desired quality of the fit, as many relaxation times as needed can be used. With an appropriate choice of N relaxation times, almost every data set on the G' and G" in the frequency domain (or the time domain) can be fitted (Morrison2001). However, with too many parameters the problem is basically reduced to solving a numerical equation and looks like a curve-fitting problem. Therefore, from physics and biology viewpoints, finding an accurate mathematical equation with many parameters to fit the rheological data set does not necessarily have a meaningful interpretation.

4.10.3.2.h Hydrogel or not?

> *"A gel is a gel, as long as one cannot prove that it is not a gel!" (Nishinari, 2009)*

In a gross definition, gelation is the conversion of a liquid solution, 'sol' state, to a disordered solid, 'gel' state, through formation of some type of network of physical or chemical bonds to form a physical or chemical gel, respectively (Larson1999). Chemical gels are constructed as a result of permanent network formation like covalent bonds, while in physical gels, the bonds are transient and reversible, like hydrophobic bonds, hydrogen bonds, and attractive electrostatic interactions (Guvendiren2012, Csizmok2016). Each group can be further divided into weak or strong gel, depending on the mechanical properties, arising from the nature of bonds. Physical gels are much more diverse because of the large diversity in physical bonds and their transient nature. Gelation inside the NPC, if any, should be investigated under the physical gel category.

More strictly speaking, however, there have been, and still are, long debates in rheology about a formal definition of the gel since the twenties, after Lloyd (Lloyd1926) stated that "the gel is one which it is easier to recognize than to define"[2]. It is not a straightforward task to answer the question of "What is a gel?" (Nishinari2009). As such, it is not surprising that for a long time rheologists have proposed different definitions, either from phenomenological, structural, rheological, or operational viewpoints (Almdal1993, Nishinari2009, Djabourov2013).

While from a rheological viewpoint, a gel can be thought of as a system that does not flow and also has a stable plateau in the storage modulus over wide range of applicable frequencies, from a structural viewpoint a gel is based on the extensive connectivity inside the system (Nishinari2009). Similarly, from phenomenological point of view, a gel is a soft, solid-like, or liquid-like material composed of two or more components, one of which is a liquid element (Almdal1993). If the liquid is water, the gel is called hydrogel. In this definition, the gel characteristic is defined in terms of its response over time.

Yet, there is not an all-inclusive definition for the gel from quantitative viewpoint. This is perhaps due to the fact that in practice there exist numerous types of gels, either synthetic, natural, or biological, that are different from each other in many ways, including the formation mechanism,

[2] This reminds me of a time when I was asking from a polymer scientist professor in her office on UC Berkeley campus about the 'true' definition of the gel. Her final response, while she was sitting behind her desk, was: "Look! if I put it on my desk and I see it doesn't flow, I would call it a gel"

reversibility, structural elements, flow lifetime, length scale, etc. It is not thus easy to propose a comprehensive definition for a gel, because there is always possibility to find exception.

On a more *rheologically* relevant basis for the definition of a physical gel, it has been stated that the mechanical spectrum of a gel, as a viscoelastic solid, has a finite elastic modulus whose value is higher, usually 5–50 times, than the loss modulus *at all frequencies* (Fig. 4.37)(Djabourov2013). Such a physical gel maybe referred to as a 'well-cured' gel (Ross-Murphy1998). The plateau in this definition is similar to the plateau region in a concentrated polymer solution or melt (Fig. 4.35), except that the plateau region in a polymer solution or melt is observed only over a limited range of frequencies. While it is generally accepted that in any kind of gel the elasticity dominates, the key point in the rheological definition is that the dominance of the elasticity should maintain over all frequencies, although its values may vary at different frequencies. This is not the case for the FG-repeats in the NPC (Fig. 4.25), and obviously not for free FG-repeats (Fig. 4.28).

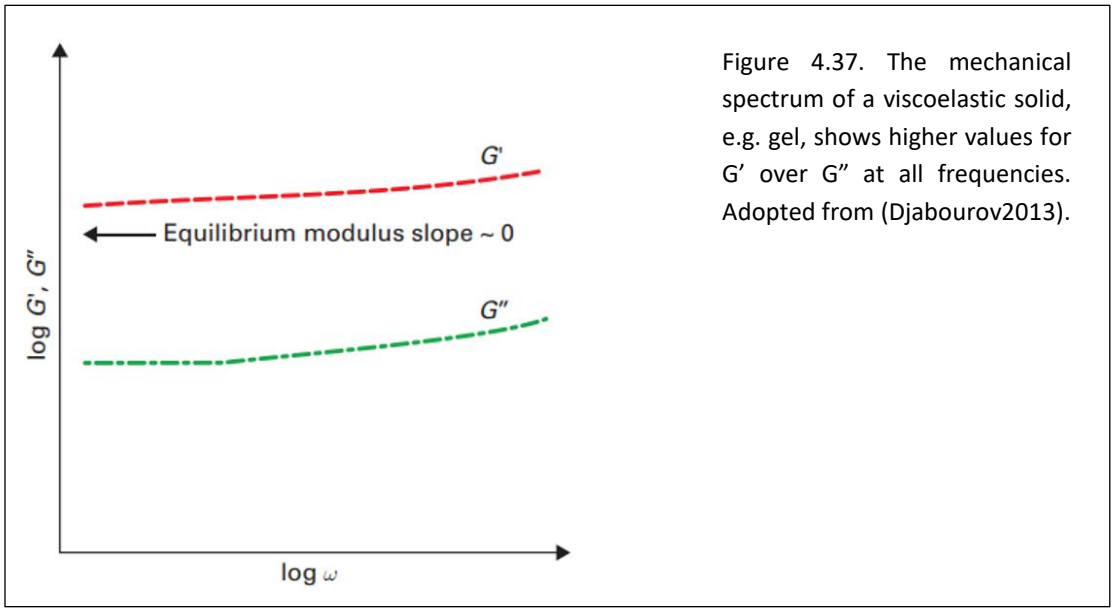


Figure 4.37. The mechanical spectrum of a viscoelastic solid, e.g. gel, shows higher values for G' over G" at all frequencies. Adopted from (Djabourov2013).

It is known that for an entangled polymer system, like a polymer melt or a concentrated polymer solution, a cross-over in G' and G" takes place as the frequency decreases (Ross-Murphy1998). When the frequency decreases beyond the cross-over point, polymer solution/melt behaves like a viscous liquid, as discussed earlier (see Fig. 4.35). In previous sections we discussed about the behavior of FG-repeats at low and *very low* frequencies, and argued that FG-repeats, either end-tethered or free, behave as viscous liquids at long times (very low very frequencies). This is in contrast to a true gel in which both loss and storage moduli plateau over long times in the order of seconds (very low frequencies) and also $G' \gg G''$ (Ross-Murphy1998). Thus, as far as the computational microrheological-driven viscoelastic response within the limitations of the current model can predict, the associating aggregation of end-tethered, and to a greater extent, free, FG-repeats unambiguously is not a *true gel.* Instead, FG-repeat meshwork tends to show the characteristics of a concentrated polymer solution or a polymer melt. We also hypothesize that the

gelation of single-type free FG-repeats *in vivo* (Frey2006b, Frey2007, Frey2009, Mohr2009, Ader2010) is likely due to the extreme pH conditions in the experiment.

By definition a polymer solution forms when polymer chains are dissolved in a solvent, while a polymer melt is a bulk liquid state of the aggregation of polymer chains in the absence of solvent (Rubinstein2003). Now, the question is that if FG-repeats do not form a true gel, which conformation do they take? A concentrated polymer solution or a polymer melt?

At first glance, it sounds plausible to state that FG-repeats form a concentrated polymer solution, because of the abundance of water and ions, acting as a solvent inside the cell. However, one should look deeper into the role and characteristics of waters inside the cell. Intrecellular water has inevitably different characteristics than extracellular water. For one thing, the intracellular water is more ordered than the extracellular water (Hazlewood1975). Moreover, inside the cell, among other things the water molecules mainly play the role of metabolic reactants, products, regulators, chaperons, and catalysts. It is thus suggested that only small portion of intracellular water serve solely as a medium (Bagatolli2016). Consequently, it seems unconvincing that the intracellular water act as a solvent for FG-repeats inside the NPC. In particular, FG-repeats possess a molar concentration inside the NPC. These suggest that FG-repeats likely form a polymer melt within the NPC.

Finally, it should be mentioned that the concept of 'temporary gel' (Kontopoulou2011) might challenge this conclusion in that a physical aggregation of transiently associating polymers might form a gel over a limited frequency range. From this viewpoint, it can be said that FG-repeats reveal the characteristics of a temporary gel that is 'reversible' in nature.

At conclusion, we suggest that FG-repeats behave as a *polymer melt*, when one considers their viscoelastic response over the entire frequency domain accessible to the current mode. However, over the limited range of frequency where the plateau modulus is dominant, FG-repeats temporarily behave as a *reversible gel*. After all, it is the *combination* of the viscoelastic behavior with other analytical techniques that can provide deep insights into the gelation process (te Nijenhuis1997).

Conclusion and future works.

In summary, FG-repeat domains inside the native NPC are collectively extremely dynamic on the timescale of submicrosecond. They form a self-healing meshwork, reconstructing itself in response to any lesion made by an active cargo. The reconstruction time depends on the size and the shape of the cargo.

From a microrheological point of view, the conformational behavior of FG-repeats can be directly extracted from their MSD. Traditionally, in a single-particle tracking microrheology a probe is sent into the polymeric medium, and from its MSD then microdynamics of the medium is inferred. Many-particle tracking microrheology extends the method to more than one probe. Nevertheless, the indigenous thermal fluctuations of the meshwork are biased upon interacting with the external probe. In this dissertation I used many-particle tracking without external probe in which the indigenous thermal fluctuations of monomers are directly measured via their MSD, quantifying the conformational behavior.

FG-repeats inside the NPC possess a peculiar situation, accompanied by several major factors, including end-tethering, confinement, crowdedness, and inter- and intra-monomer interactions. It appears that the end-tethering has a dominant effect on the conformational behavior of FG-repeats. The viscoelastic response in the frequency domain is an effective way to quantify the conformational behavior. The viscoelastic response of FG-repeats is strongly frequency-dependent. Consistent with the purpose and function of the selectivity barrier, FG-repeats behave as a viscous liquid at high frequencies, i.e. the timescale of the diffusion of small molecules. The liquid is non-Newtonian with shear-thinning behavior, as manifested in the frequency-dependent dynamic viscosity. Moreover, at high frequencies, FG-repeats regardless of their physical, geometrical, and compositional conditions invariably show the same frequency-dependent power-law.

At low frequencies, the viscoelastic response of FG-repeats is strongly affected by the end-tethering. At these frequencies, FG-repeats form a predominantly elastic medium inside the NPC. Importantly, the equivalent timescale is consistent with the nucleocytoplasmic transport time on the order of milliseconds. In the absence of end-tethering factor, free FG-repeats behave like a non-Newtonian viscous liquid over all frequencies.

Finally, whether FG-repeats form a gel inside the NPC channel or not, depends on the timescale of the experiment. From a microrheology point of view, FG-repeats behave more like a polymer melt, rather than a gel. Unambiguously, FG-repeats do not form a true gel over all frequency domain. However, it may be said that FG-repeats form a temporary gel over a limited range of frequencies. Significantly, this range of frequency is equivalent to the nucleocytoplasmic transport timescale.

Future works

This work can be further extended to study other major effects on the conformational behavior of polymeric systems, and in particular intrinsically disordered proteins. One interesting, less-

explored area would be the combination of bioinformatics and biophysical coarse-grained model to further scrutinize the effects of conserved and non-conserved amino-acid sequences on the microrheology of FG-repeats.

Another interesting future work is to extend the viscoelastic response of FG-repeats over all range of frequencies, from very low to very high. The latter can be conducted via molecular dynamics approach. For very low frequencies, however, while it is not practical to use the current level of coarse-graining, it is possible to develop a coarser model. Alternatively, one may implement dynamical information into a probabilistic approach, like ABM, to reach to the very low frequency regime. Finally, the results from different scales should be carefully superimposed with appropriate shift factors. A complete viscoelastic response of FG-repeats over all range of frequencies can provide excellent insight into their true conformational behavior. The many-particle tracking microrheology without external probe can be used to investigate any polymeric system of interest.

Finally, the level of coarse-graining justified in this work on the basis of physical foundation and sensitivity analysis, may be used as a criterion in finding the optimal granularity level in similar works.

References

Addas, K. M., C. F. Schmidt and J. X. Tang (2004). "Microrheology of solutions of semiflexible biopolymer filaments using laser tweezers interferometry." Physical Review E **70**(2): 021503.

Ader, C., S. Frey, W. Maas, H. B. Schmidt, D. Gorlich and M. Baldus (2010). "Amyloid-like interactions within nucleoporin FG hydrogels." PNAS **107**(14): 6281-6285.

Aitchison, J. D. and M. P. Rout (2012a). "The yeast nuclear pore complex and transport through it." Genetics **190**(3): 855-883.

Aitchison, J. D. and M. P. Rout (2012b). "The yeast nuclear pore complex and transport through it." Genetics **190**(3): 855-883.

Akey, C. W. (1995). "Structural plasticity of the nuclear pore complex." Journal of molecular biology **248**(2): 273-293.

Alber, F., S. Dokudovskaya, L. M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams and B. T. Chait (2007a). "The molecular architecture of the nuclear pore complex." Nature **450**(7170): 695.

Alber, F., S. Dokudovskaya, L. M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B. T. Chait, M. P. Rout and A. Sali (2007b). "Determining the architectures of macromolecular assemblies." Nature **450**(7170): 683-694.

Allen, M. P. and D. Tildesley (1987). Computer Simulation of Liquids. . Oxford Clarendon Press.

Allen, M. P. and D. J. Tildesley (1989). Computer simulation of liquids, Oxford university press.

Allison, S. A. (1986). "Brownian Dynamics Simulation of Wormlike Chains - Fluorescence Depolarization and Depolarized Light-Scattering." Macromolecules **19**(1): 118-124.

Almdal, K., J. Dyre, S. Hvidt and O. Kramer (1993). "Towards a phenomenological definition of the term 'gel'." Polymer gels and networks **1**(1): 5-17.

Amblard, F., A. C. Maggs, B. Yurke, A. N. Pargellis and S. Leibler (1996). "Subdiffusion and anomalous local viscoelasticity in actin networks." Physical review letters **77**(21): 4470.

Ando, D., R. Zandi, Y. W. Kim, M. Colvin, M. Rexach and A. Gopinathan (2014). "Nuclear pore complex protein sequences determine overall copolymer brush structure and function." Biophysical journal **106**(9): 1997-2007.

Ando, T. and J. Skolnick (2010). "Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion." Proceedings of the National Academy of Sciences **107**(43): 18457-18462.

Andrews, N., A. Doufas and A. McHugh (1998). "Effect of solvent quality on the rheological and rheooptical properties of flexible polymer solutions." Macromolecules **31**(9): 3104-3108.

Ashurst, W.-T. and W. Hoover (1975). "Dense-fluid shear viscosity via nonequilibrium molecular dynamics." Physical Review A **11**(2): 658.

Atakhorrami, M., G. Koenderink, C. Schmidt and F. MacKintosh (2005). "Short-time inertial response of viscoelastic fluids: observation of vortex propagation." Physical review letters **95**(20): 208302.

Azimi, M., E. Bulat, K. Weis and M. R. Mofrad (2014). "An agent-based model for mRNA export through the nuclear pore complex." Molecular biology of the cell **25**(22): 3643-3653.

Azimi, M., Y. Jamali and M. R. Mofrad (2011). "Accounting for diffusion in agent based models of reaction-diffusion systems with application to cytoskeletal diffusion." PLoS One **6**(9): e25306.

Azimi, M. and M. R. Mofrad (2013). "Higher nucleoporin-Importinβ affinity at the nuclear basket increases nucleocytoplasmic import." PloS one **8**(11): e81741.

Babu, M. M., R. W. Kriwacki and R. V. Pappu (2012). "Versatility from protein disorder." Science **337**(6101): 1460-1461.

Bagatolli, L. and R. Stock (2016). "The cell as a gel: material for a conceptual discussion." Physiol. Mini Rev **9**: 38-49.

Bah, A., R. M. Vernon, Z. Siddiqui, M. Krzeminski, R. Muhandiram, C. Zhao, N. Sonenberg, L. E. Kay and J. D. Forman-Kay (2015). "Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch." Nature **519**(7541): 106-109.

Bangs, P., B. Burke, C. Powers, R. Craig, A. Purohit and S. Doxsey (1998). "Functional analysis of Tpr: identification of nuclear pore complex association and nuclear localization domains and a role in mRNA export." J Cell Biol **143**(7): 1801-1812.

Bao, G. (2009). "Protein mechanics: a new frontier in biomechanics." Experimental mechanics **49**(1): 153.

Bauer, D. W., J. B. Huffman, F. L. Homa and A. Evilevitch (2013). "Herpes virus genome, the pressure is on." Journal of the American Chemical Society **135**(30): 11216-11221.

Bayliss, R., A. H. Corbett and M. Stewart (2000a). "The molecular mechanism of transport of macromolecules through nuclear pore complexes." Traffic **1**(6): 448-456.

Bayliss, R., S. W. Leung, R. P. Baker, B. B. Quimby, A. H. Corbett and M. Stewart (2002a). "Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats." The EMBO journal **21**(12): 2843-2853.

Bayliss, R., T. Littlewood and M. Stewart (2000b). "Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking." Cell **102**(1): 99-108.

Bayliss, R., T. Littlewood, L. A. Strawn, S. R. Wente and M. Stewart (2002b). "GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta." J Biol Chem **277**(52): 50597-50606.

Beck, M., F. Forster, M. Ecke, J. M. Plitzko, F. Melchior, G. Gerisch, W. Baumeister and O. Medalia (2004). "Nuclear pore complex structure and dynamics revealed by cryoelectron tomography." Science **306**(5700): 1387-1390.

Beck, M., V. Lučić, F. Förster, W. Baumeister and O. Medalia (2007). "Snapshots of nuclear pore complexes in action captured by cryo-electron tomography." Nature **449**(7162): 611-615.

Bednenko, J., G. Cingolani and L. Gerace (2003). "Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport." J Cell Biol **162**(3): 391-401.

Ben-Efraim, I., P. D. Frosst and L. Gerace (2009). "Karyopherin binding interactions and nuclear import mechanism of nuclear pore complex protein Tpr." BMC cell biology **10**(1): 74.

Ben-Efraim, I. and L. Gerace (2001a). "Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import." J Cell Biol **152**(2): 411-417.

Ben-Efraim, I. and L. Gerace (2001b). "Gradient of increasing affinity of importin β for nucleoporins along the pathway of nuclear import." The Journal of cell biology **152**(2): 411-418.

Benmouna, F. and D. Johannsmann (2004). "Viscoelasticity of gelatin surfaces probed by AFM noise analysis." Langmuir **20**(1): 188-193.

Berendsen, H. J. (2007). Simulating the physical world: hierarchical modeling from quantum mechanics to fluid dynamics, Cambridge University Press.

Berg, H. C. (1983). Random walks in biology, Princeton University Press.

Berg, O. G. (1979). "Brownian motion of the wormlike chain and segmental diffusion of DNA." Biopolymers **18**(11): 2861-2874.

Bergenholtz, J. and M. Fuchs (1999). "Nonergodicity transitions in colloidal suspensions with attractive interactions." Physical Review E **59**(5): 5706.

Binder, K. (1995). Monte Carlo and molecular dynamics simulations in polymer science, Oxford University Press.

Binder, K., D. Heermann, L. Roelofs, A. J. Mallinckrodt and S. McKay (1993). "Monte Carlo simulation in statistical physics." Computers in Physics **7**(2): 156-157.

Bird, R. B., C. F. Curtiss, R. C. Armstrong and O. Hassager (1987). Dynamics of polymeric liquids. Vol. 2: Kinetic Theory Wiley.

Block, J., V. Schroeder, P. Pawelzyk, N. Willenbacher and S. Koster (2015). "Physical properties of cytoplasmic intermediate filaments." Biochim Biophys Acta **1853**(11 Pt B): 3053-3064.

Boal, D. and D. H. Boal (2012). Mechanics of the Cell, Cambridge University Press.

Bonabeau, E. (2002). "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the National Academy of Sciences **99**(suppl 3): 7280-7287.

Bouchaud, J.-P. and A. Georges (1990). "Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications." Physics reports **195**(4-5): 127-293.

Box, G. E. and M. E. Muller (1958). "A note on the generation of random normal deviates." The annals of mathematical statistics **29**(2): 610-611.

Bressloff, P. C. and J. M. Newby (2013). "Stochastic models of intracellular transport." Reviews of Modern Physics **85**(1): 135.

Brickner, J. H. (2010). "Transcriptional memory: staying in the loop." Current Biology **20**(1): R20-R21.

Broedersz, C. P. and F. C. MacKintosh (2014). "Modeling semiflexible polymer networks." Reviews of Modern Physics **86**(3): 995.

Brown, R. (1828). A Brief Account of Microscopical Observations Made... on the Particles Contained in the Pollen of Plants, and on the General Existence of Active Molecules in Organic and Inorganic Bodies.

Buehler, M. J. and Y. C. Yung (2009). "Deformation and failure of protein materials in physiologically extreme conditions and disease." Nat Mater **8**(3): 175-188.

Bui, K. H., A. von Appen, A. L. DiGuilio, A. Ori, L. Sparks, M.-T. Mackmull, T. Bock, W. Hagen, A. Andrés-Pons and J. S. Glavy (2013). "Integrated structural analysis of the human nuclear pore complex scaffold." Cell **155**(6): 1233-1243.

Bustamante, C., J. F. Marko, E. D. Siggia and S. Smith (1994). "Entropic elasticity of lambda-phage DNA." Science **265**(5178): 1599-1600.

Caggioni, M., P. Spicer, D. Blair, S. Lindberg and D. Weitz (2007). "Rheology and microrheology of a microstructured fluid: The gellan gum case." Journal of Rheology **51**(5): 851-865.

Callan, H. and S. Tomlin (1950). "Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope." Proceedings of the Royal Society of London B: Biological Sciences **137**(888): 367-378.

Capelson, M., Y. Liang, R. Schulte, W. Mair, U. Wagner and M. W. Hetzer (2010). "Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes." Cell **140**(3): 372-383.

Carignano, M. and I. Szleifer (1994). "Pressure isotherms, phase transition, instability, and structure of tethered polymers in good, Θ, and poor solvents." The Journal of chemical physics **100**(4): 3210-3223.

Carignano, M. and I. Szleifer (1995). "Structural and thermodynamic properties of end-grafted polymers on curved surfaces." The Journal of chemical physics **102**(21): 8662-8669.

Carter, E. (1994). "Generation and Application of Random Numbers." Forth Dimensions **16**: 12-12.

Cascales, J. L. and J. G. de la Torre (1991). "Simulation of polymer chains in elongational flow. Steady-state properties and chain fracture." The Journal of chemical physics **95**(12): 9384-9392.

Casolari, J. M., C. R. Brown, S. Komili, J. West, H. Hieronymus and P. A. Silver (2004). "Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization." Cell **117**(4): 427-439.

Caspi, A., R. Granek and M. Elbaum (2002). "Diffusion and directed motion in cellular transport." Physical Review E **66**(1): 011916.

Chakrabarti, R., A. Debnath and K. Sebastian (2014). "Diffusion in an elastic medium: A model for macromolecule transport across the nuclear pore complex." Physica A: Statistical Mechanics and its Applications **404**: 65-78.

Chandran, P. L. and M. R. Mofrad (2010). "Averaged implicit hydrodynamic model of semiflexible filaments." Physical Review E **81**(3): 031920.

Chatterjee, D. and B. J. Cherayil (2011). "Subdiffusion as a model of transport through the nuclear pore complex." The Journal of chemical physics **135**(15): 10B612.

Chen, D., E. Weeks, J. C. Crocker, M. Islam, R. Verma, J. Gruber, A. Levine, T. C. Lubensky and A. Yodh (2003). "Rheological microscopy: local mechanical properties from microrheology." Physical review letters **90**(10): 108301.

Chen, D. T., Q. Wen, P. A. Janmey, J. C. Crocker and A. G. Yodh (2010). "Rheology of soft materials."

Chhabra, R. P. and J. F. Richardson (2011). Non-Newtonian flow and applied rheology: engineering applications, Butterworth-Heinemann.

Chi, N. C. and S. A. Adam (1997). "Functional domains in nuclear import factor p97 for binding the nuclear localization sequence receptor and the nuclear pore." Molecular Biology of the Cell **8**(6): 945-956.

Chook, Y. M. and G. Blobel (2001). "Karyopherins and nuclear import." Curr Opin Struct Biol **11**(6): 703-715.

Christensen, K. (2002). "Percolation theory." Imperial College London **1**.

Christensen, R. (2012). Theory of viscoelasticity: an introduction, Elsevier.

Cifre, J. G. H. and J. G. de la Torre (1999). "Steady-state behavior of dilute polymers in elongational flow. Dependence of the critical elongational rate on chain length, hydrodynamic interaction, and excluded volume." Journal of Rheology **43**(2): 339-358.

Claesson, P. M. and H. K. Christenson (1988). "Very long range attractive forces between uncharged hydrocarbon and fluorocarbon surfaces in water." The Journal of Physical Chemistry **92**(6): 1650-1655.

Coalson, R. D., A. Eskandari Nasrabad, D. Jasnow and A. Zilman (2015). "A Polymer-Brush-Based Nanovalve Controlled by Nanoparticle Additives: Design Principles." The Journal of Physical Chemistry B **119**(35): 11858-11866.

Cohen, A. (1991). "A Padé approximant to the inverse Langevin function." Rheologica acta **30**(3): 270-273.

Cohen, M., N. Feinstein, K. L. Wilson and Y. Gruenbaum (2003). "Nuclear pore protein gp210 is essential for viability in HeLa cells and Caenorhabditis elegans." Mol Biol Cell **14**(10): 4230-4237.

Colwell, L. J., M. P. Brenner and K. Ribbeck (2010a). "Charge as a Selection Criterion for Translocation through the Nuclear Pore Complex." Plos Computational Biology **6**(4): -.

Colwell, L. J., M. P. Brenner and K. Ribbeck (2010b). "Charge as a Selection Criterion for Translocation through the Nuclear Pore Complex." Plos Computational Biology **6**(4).

Cook, A., E. Fernandez, D. Lindner, J. Ebert, G. Schlenstedt and E. Conti (2005). "The structure of the nuclear export receptor Cse1 in its cytosolic state reveals a closed conformation incompatible with cargo binding." Molecular cell **18**(3): 355-367.

Cramer, H. (1950). "Mathematical methods of statistics, Princeton, 1946." Mathematical Reviews (MathSciNet): MR16588 Zentralblatt MATH **63**: 300.

Creighton, T. E. (1993). Proteins : structures and molecular properties. New York, W.H. Freeman.

Crocker, J. C. and B. D. Hoffman (2007). "Multiple-particle tracking and two-point microrheology in cells." Methods in cell biology **83**: 141-178.

Crocker, J. C., M. T. Valentine, E. R. Weeks, T. Gisler, P. D. Kaplan, A. G. Yodh and D. A. Weitz (2000). "Two-point microrheology of inhomogeneous soft materials." Physical Review Letters **85**(4): 888.

Cronshaw, J. M., A. N. Krutchinsky, W. Zhang, B. T. Chait and M. J. Matunis (2002). "Proteomic analysis of the mammalian nuclear pore complex." J Cell Biol **158**(5): 915-927.

Cronshaw, J. M. and M. J. Matunis (2004). "The nuclear pore complex: disease associations and functional correlations." Trends Endocrinol Metab **15**(1): 34-39.

Csizmok, V., A. V. Follis, R. W. Kriwacki and J. D. Forman-Kay (2016). "Dynamic protein interaction networks and new structural paradigms in signaling." Chemical reviews **116**(11): 6424-6462.

Curtis, R., C. Steinbrecher, M. Heinemann, H. Blanch and J. Prausnitz (2002). "Hydrophobic forces between protein molecules in aqueous solutions of concentrated electrolyte." Biophysical chemistry **98**(3): 249-265.

D'Angelo, M. A. and M. W. Hetzer (2008). "Structure, dynamics and function of nuclear pore complexes." Trends Cell Biol **18**(10): 456-466.

da Costa, H. M., V. D. Ramos and M. C. Rocha (2005). "Rheological properties of polypropylene during multiple extrusion." Polymer Testing **24**(1): 86-93.

Dahl, K. N., S. M. Kahn, K. L. Wilson and D. E. Discher (2004). "The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber." Journal of Cell Science **117**(20): 4779-4786.

Damelin, M. and P. A. Silver (2000). "Mapping interactions between nuclear transport factors in living cells reveals pathways through the nuclear pore complex." Molecular cell **5**(1): 133-140.

Dange, T., D. Grunwald, A. Grunwald, R. Peters and U. Kubitscheck (2008a). "Autonomy and robustness of translocation through the nuclear pore complex: a single-molecule study." THE JOURNAL OF CELL BIOLOGY **183**(1): 77-86.

Dange, T., D. Grünwald, A. Grünwald, R. Peters and U. Kubitscheck (2008b). "Autonomy and robustness of translocation through the nuclear pore complex: a single-molecule study." J Cell Biol **183**(1): 77-86.

Dasgupta, B. R., S.-Y. Tee, J. C. Crocker, B. Frisken and D. Weitz (2002). "Microrheology of polyethylene oxide using diffusing wave spectroscopy and single scattering." Physical Review E **65**(5): 051505.

Davis, M. E. and J. A. McCammon (1990). "Electrostatics in biomolecular structure and dynamics." Chemical Reviews **90**(3): 509-521.

de Gennes, P.-G. (1979a). Scaling concepts in polymer physics. Ithaca, N.Y., Cornell University Press.

De Gennes, P. (1976). "Dynamics of entangled polymer solutions. II. Inclusion of hydrodynamic interactions." Macromolecules **9**(4): 594-598.

De Gennes, P. G. (1979b). Scaling concepts in polymer physics, Cornell university press.

Dealy, J. (2007). "Questions about relaxation spectra submitted by a reader." Rheology Bulletin **76**(1): 14.

Dealy, J. M. and R. G. Larson (2006). Structure and rheology of molten polymers, Hanser.

Deng, L., X. Trepat, J. P. Butler, E. Millet, K. G. Morgan, D. A. Weitz and J. J. Fredberg (2006). "Fast and slow dynamics of the cytoskeleton." Nature materials **5**(8): 636.

Denning, D., S. Patel, V. Uversky, A. Fink and M. Rexach (2003a). "Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded." PNAS **100**: 2450–2455.

Denning, D. P., S. S. Patel, V. Uversky, A. L. Fink and M. Rexach (2003b). "Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded." Proc Natl Acad Sci U S A **100**(5): 2450-2455.

Deshpande, A., J. M. Krishnan and P. B. S. Kuma, Eds. (2010). Rheology of Complex Fluids, Springer.

Devos, D., S. Dokudovskaya, R. Williams, F. Alber, N. Eswar, B. T. Chait, M. P. Rout and A. Sali (2006). "Simple fold composition and modular architecture of the nuclear pore complex." PNAS **103**: 2172–2177.

Dhont, J. K. and W. Briels (2003). "Inhomogeneous suspensions of rigid rods in flow." The Journal of chemical physics **118**(3): 1466-1478.

Diesinger, P. M. and D. W. Heermann (2010). "Hydrophobicity as a possible reason for gelation of FG-rich nucleoporins." European Biophysics Journal **39**(2): 299.

Djabourov, M., K. Nishinari and S. B. Ross-Murphy (2013). Physical gels from biological and synthetic polymers, Cambridge University Press.

Dlugosz, M. and J. Trylska (2011). "Diffusion in crowded biological environments: applications of Brownian dynamics." BMC biophysics **4**: 3.

Dogic, Z., J. Zhang, A. Lau, H. Aranda-Espinoza, P. Dalhaimer, D. E. Discher, P. Janmey, R. D. Kamien, T. C. Lubensky and A. Yodh (2004). "Elongation and fluctuations of semiflexible polymers in a nematic solvent." Physical review letters **92**(12): 125503.

Doi, M. and S. F. Edwards (1988). The theory of polymer dynamics, oxford university press.

Dolker, N., U. Zachariae and H. Grubmuller (2010). "Hydrophilic linkers and polar contacts affect aggregation of FG repeat peptides." Biophys J **98**(11): 2653-2661.

Donaldson, S. H., C. T. Lee, B. F. Chmelka and J. N. Israelachvili (2011). "General hydrophobic interaction potential for surfactant/lipid bilayers from direct force measurements between light-modulated bilayers." Proceedings of the National Academy of Sciences **108**(38): 15699-15704.

Doyle, P. S., E. S. Shaqfeh and A. P. Gast (1997). "Dynamic simulation of freely draining flexible polymers in steady linear flows." Journal of Fluid Mechanics **334**: 251-291.

Eastman, P. (2015). Introduction to Statistical Mechanics, Free distribution

Eibauer, M., M. Pellanda, Y. Turgay, A. Dubrovsky, A. Wild and O. Medalia (2015). "Structure and gating of the nuclear pore complex." Nature Communications **6**.

Einstein, A. (1905). "On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat." Annalen der physik **17**: 549-560.

Ermak, D. L. and J. A. Mccammon (1978). "Brownian Dynamics with Hydrodynamic Interactions." Journal of Chemical Physics **69**(4): 1352-1360.

Fahrenkrog, B. and U. Aebi (2002). The vertebrate nuclear pore complex: from structure to function. Nuclear Transport, Springer**:** 25-48.

Fahrenkrog, B., E. C. Hurt, U. Aebi and N. Pante (1998). "Molecular architecture of the yeast nuclear pore complex: localization of Nsp1p subcomplexes." J Cell Biol **143**(3): 577-588.

Fakhri, N., D. A. Tsyboulski, L. Cognet, R. B. Weisman and M. Pasquali (2009). "Diameter-dependent bending dynamics of single-walled carbon nanotubes in liquids." Proceedings of the National Academy of Sciences **106**(34): 14219-14223.

Fall, C. (2002). Computational Cell Biology. New Yrok, Springer.

Ferry, J. D. (1980). Viscoelastic properties of polymers, John Wiley & Sons.

Feuerbach, F., V. Galy, E. Trelles-Sticken, M. Fromont-Racine, A. Jacquier, E. Gilson, J.-C. Olivo-Marin, H. Scherthan and U. Nehrbass (2002). "Nuclear architecture and spatial positioning help establish transcriptional states of telomeres in yeast." Nature cell biology **4**(3): 214-221.

Fick, A. (1855). "V. On liquid diffusion." Philosophical Magazine Series 4 **10**(63): 30-39.

Fikus, M. and P. Pawlowski (1989). "Bioelectrorheological model of the cell. 2. Analysis of creep and its experimental verification." Journal of theoretical biology **137**(4): 365-373.

Fischer, H., I. Polikarpov and A. F. Craievich (2004). "Average protein density is a molecular-weight-dependent function." Protein Sci **13**(10): 2825-2828.

Fixman, M. (1978a). "Simulation of polymer dynamics. I. General theory." The Journal of Chemical Physics **69**(4): 1527-1537.

Fixman, M. (1978b). "Simulation of polymer dynamics. II. Relaxation rates and dynamic viscosity." The Journal of Chemical Physics **69**(4): 1538-1545.

Flory, P. J. (1941). "Molecular size distribution in three dimensional polymers. I. Gelation1." Journal of the American Chemical Society **63**(11): 3083-3090.

Flory, P. J. (1953). Principles of polymer chemistry, Cornell University Press.

Fogolari, F., A. Brigo and H. Molinari (2002). "The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology." Journal of Molecular Recognition **15**(6): 377-392.

Francis, N. J. and R. E. Kingston (2001). "Mechanisms of transcriptional memory." Nature Reviews Molecular Cell Biology **2**(6): 409-421.

Freed, K. F. and S. Edwards (1974). "Polymer viscosity in concentrated solutions." The Journal of Chemical Physics **61**(9): 3626-3633.

Frenkel, D. and B. Smit (2001). Understanding molecular simulation: from algorithms to applications, Academic press.

Frey, S. and D. Gorlich (2007). "A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes." Cell **130**(3): 512-523.

Frey, S. and D. Gorlich (2009). "FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties." EMBO J **28**(17): 2554-2567.

Frey, S., R. P. Richter and D. Goerlich (2006a). "FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties." Science **314**(5800): 815-817.

Frey, S., R. P. Richter and D. Gorlich (2006b). "FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties." Science **314**(5800): 815-817.

Frosst, P., T. Guan, C. Subauste, K. Hahn and L. Gerace (2002). "Tpr is localized within the nuclear basket of the pore complex and has a role in nuclear protein export." The Journal of cell biology **156**(4): 617-630.

Gall, J. (1967). "Octagonal nuclear pores." J. Cell Biol. **32**: 391 - 399.

Gallardo, P., S. Salas-Pino and R. R. Daga (2017). "A new role for the nuclear basket network." Microbial Cell **4**(12): 423.

Gamini, R., W. Han, J. E. Stone and K. Schulten (2014). "Assembly of Nsp1 nucleoporins provides insight into nuclear pore complex gating." PLoS computational biology **10**(3): e1003488.

Gardel, M., M. Valentine, J. C. Crocker, A. Bausch and D. Weitz (2003). "Microrheology of entangled F-actin solutions." Physical review letters **91**(15): 158302.

Geng, Y., P. Dalhaimer, S. Cai, R. Tsai, M. Tewari, T. Minko and D. E. Discher (2007). "Shape effects of filaments versus spherical particles in flow and drug delivery." Nature Nanotechnology **2**(4): 249-255.

Ghatak, A. K. and S. Lokanathan (2004). Quantum Mechanics: Theory & Applications, Macmillan.

Ghavami, A., E. van der Giessen and P. R. Onck (2016). "Energetics of transport through the nuclear pore complex." PloS one **11**(2): e0148876.

Ghavami, A., L. M. Veenhoff, E. van der Giessen and P. R. Onck (2014). "Probing the disordered domain of the nuclear pore complex through coarse-grained molecular dynamics simulations." Biophysical journal **107**(6): 1393-1402.

Giannelis, E. P., R. Krishnamoorti and E. Manias (1999). "Polymer-silicate nanocomposites: Model systems for confined polymers and polymer brushes." Polymers in Confined Environments **138**: 107-147.

Gisler, T. and D. A. Weitz (1999). "Scaling of the microrheology of semidilute F-actin solutions." Physical review letters **82**(7): 1606.

Gittes, F. and F. MacKintosh (1998). "Dynamic shear modulus of a semiflexible polymer network." Physical Review E **58**(2): R1241.

Gittes, F., B. Mickey, J. Nettleton and J. Howard (1993). "Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape." J Cell Biol **120**(4): 923-934.

Gittes, F., B. Schnurr, P. Olmsted, F. C. MacKintosh and C. F. Schmidt (1997). "Microscopic viscoelasticity: shear moduli of soft materials determined from thermal fluctuations." Physical review letters **79**(17): 3286.

Gorlich, D., M. J. Seewald and K. Ribbeck (2003). "Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation." EMBO J **22**(5): 1088-1100.

Görlich, D., M. J. Seewald and K. Ribbeck (2003). "Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation." The EMBO journal **22**(5): 1088-1100.

Goryaynov, A. and W. Yang (2014). "Role of molecular charge in nucleocytoplasmic transport." PLoS One **9**(2): e88792.

Gottlieb, M. and R. B. Bird (1976). "A molecular dynamics calculation to confirm the incorrectness of the random-walk distribution for describing the Kramers freely jointed bead–rod chain." The Journal of Chemical Physics **65**(6): 2467-2468.

Graessley, W. W., R. C. Hayward and G. S. Grest (1999). "Excluded-volume effects in polymer solutions. 2. Comparison of experimental results with numerical simulation data." Macromolecules **32**(10): 3510-3517.

Grant, M. and W. Russel (1993). "Volume-fraction dependence of elastic moduli and transition temperatures for colloidal silica gels." Physical Review E **47**(4): 2606.

Griffis, E. R., B. Craige, C. Dimaano, K. S. Ullman and M. A. Powers (2004). "Distinct functional domains within nucleoporins Nup153 and Nup98 mediate transcription-dependent mobility." Molecular Biology of the Cell **15**(4): 1991-2002.

Groot, R. D. and T. J. Madden (1998). "Dynamic simulation of diblock copolymer microphase separation." The Journal of chemical physics **108**(20): 8713-8724.

Grotendorst, J., N. Attig, S. Blügel and D. Marx (2009). "Multiscale simulation methods in molecular sciences." Lecture Notes, NIC Series **42**.

Guilak, F., J. R. Tedrow and R. Burgkart (2000). "Viscoelastic properties of the cell nucleus." Biochemical and Biophysical Research Communications **269**(3): 781-786.

Guo, L. and E. Luijten (2005). "Reversible gel formation of triblock copolymers studied by molecular dynamics simulation." Journal of Polymer Science Part B: Polymer Physics **43**(8): 959-969.

Gustin, K. E. and P. Sarnow (2001). "Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition." EMBO J **20**(1-2): 240-249.

Guvendiren, M., H. D. Lu and J. A. Burdick (2012). "Shear-thinning hydrogels for biomedical applications." Soft matter **8**(2): 260-272.

Haile, J. (1992). Molecular dynamics simulation: elementary methods, JOHN WILEY & SONS, INC.

Hakansson, C. and C. Elvingson (1994). "Semiflexible Chain Molecules with Nonuniform Curvature .1. Structural-Properties." Macromolecules **27**(14): 3843-3849.

Hakata, Y., M. Yamada and H. Shida (2003). "A multifunctional domain in human CRM1 (exportin 1) mediates RanBP3 binding and multimerization of human T-cell leukemia virus type 1 Rex protein." Molecular and cellular biology **23**(23): 8751-8761.

Han, C. D. (2007). Rheology and processing of polymeric materials: Volume 1: Polymer Rheology, Oxford University Press on Demand.

Hazlewood, C. F. (1975). "A Role for Water in the Exclusion of Cellular Sodium—Is a Sodium Pump Needed?" Cardiovascular diseases **2**(1): 83.

Helbing, D., I. Farkas and T. Vicsek (2000). "Simulating dynamical features of escape panic." Nature **407**(6803): 487-490.

Hellekalek, P. (1998). "Good random number generators are (not so) easy to find." Mathematics and Computers in Simulation **46**(5): 485-505.

Herrmann, M., N. Neuberth, J. r. Wissler, J. Pérez, D. Gradl and A. Naber (2009). "Near-field optical study of protein transport kinetics at a single nuclear pore." Nano letters **9**(9): 3330-3336.

Hess, B., C. Kutzner, D. Van Der Spoel and E. Lindahl (2008). "GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation." Journal of chemical theory and computation **4**(3): 435-447.

Hess, W., T. A. Vilgis and H. H. Winter (1988). "Dynamical critical behavior during chemical gelation and vulcanization." Macromolecules **21**(8): 2536-2542.

Honig, B. and A. Nicholls (1995). "Classical electrostatics in biology and chemistry." Science **268**(5214): 1144-1149.

Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling (2006). "Comparison of multiple Amber force fields and development of improved protein backbone parameters." Proteins: Structure, Function, and Bioinformatics **65**(3): 712-725.

Houchmandzadeh, B. and S. Dimitrov (1999). "Elasticity measurements show the existence of thin rigid cores inside mitotic chromosomes." The Journal of cell biology **145**(2): 215-223.

Houchmandzadeh, B., J. F. Marko, D. Chatenay and A. Libchaber (1997). "Elasticity and structure of eukaryote chromosomes studied by micromanipulation and micropipette aspiration." The Journal of cell biology **139**(1): 1-12.

Howard, J. (2001). "Mechanics of motor proteins and the cytoskeleton."

Hummer, G. (1999). "Hydrophobic force field as a molecular alternative to surface-area models." Journal of the American Chemical Society **121**(26): 6299-6305.

Hur, J. S., E. S. Shaqfeh and R. G. Larson (2000). "Brownian dynamics simulations of single DNA molecules in shear flow." Journal of Rheology **44**(4): 713-742.

Hurt, E. and M. Beck (2015). "Towards understanding nuclear pore complex architecture and dynamics in the age of integrative structural analysis." Current Opinion in Cell Biology **34**: 31-38.

Iovine, M. K., J. L. Watkins and S. R. Wente (1995). "The GLFG repetitive region of the nucleoporin Nup116p interacts with Kap95p, an essential yeast nuclear import factor." The Journal of Cell Biology **131**(6): 1699-1713.

Isambert, H. and A. Maggs (1996). "Dynamics and rheology of actin solutions." Macromolecules **29**(3): 1036-1040.

Isgro, T. and K. Schulten (2007a). "Association of nuclear pore FG-repeat domains to NTF2 import and export complexes." J Mol Biol **366**: 330–345.

Isgro, T. A. and K. Schulten (2005a). "Binding Dynamics of Isolated Nucleoporin Repeat Regions to Importin-b." Structure **13**: 1869–1879.

Isgro, T. A. and K. Schulten (2005b). "Binding dynamics of isolated nucleoporin repeat regions to importin-beta." Structure **13**(12): 1869-1879.

Isgro, T. A. and K. Schulten (2007). "Cse1p-binding dynamics reveal a binding pattern for FG-repeat nucleoporins on transport receptors." Structure **15**(8): 977-991.

Israelachvili, J. N. (1992). Intermolecular and surface forces. London Academic Press.

Israelachvili, J. N. (2011). Intermolecular and surface forces, Academic press.

Jahed, Z., M. Soheilypour, M. Peyro and M. R. Mofrad (2016). "The LINC and NPC relationship–it's complicated!" J Cell Sci **129**(17): 3219-3229.

Jamali, T., Y. Jamali, M. Mehrbod and M. R. K. Mofrad (2011). "Nuclear Pore Complex: Biochemistry and Biophysics of Nucleocytoplasmic Transport in Health and Disease." nternational Reviews of Cell and Molecular Biology. **28706** 233.

Janmey, P. and C. Schmidt (2006). Experimental measurements of intracellular mechanics. Cytoskeletal Mechanics; Models and Measuements. M. Mofrad and R. D. Kamm.

Jovanovic-Talisman, T., J. Tetenbaum-Novatt, A. S. McKenney, A. Zilman, R. Peters, M. P. Rout and B. T. Chait (2009a). "Artificial nanopores that mimic the transport selectivity of the nuclear pore complex." Nature **457**(7232): 1023-1027.

Jovanovic-Talisman, T., J. Tetenbaum-Novatt, A. S. McKenney, A. Zilman, R. Peters, M. P. Rout and B. T. Chait (2009b). "Artificial nanopores that mimic the transport selectivity of the nuclear pore complex." Nature **457**(7232): 1023-1027.

Kabachinski, G. and T. U. Schwartz (2015). "The nuclear pore complex–structure and function at a glance." Journal of Cell Science **128**(3): 423-429.

Kapon, R., A. Topchik, D. Mukamel and Z. Reich (2008). "A possible mechanism for self-coordination of bidirectional traffic across nuclear pores." Physical biology **5**(3): 036001.

Kasper, L. H., P. K. Brindle, C. A. Schnabel, C. E. Pritchard, M. L. Cleary and J. M. van Deursen (1999). "CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity." Mol Cell Biol **19**(1): 764-776.

Kehlenbach, R. H., A. Dickmanns, A. Kehlenbach, T. Guan and L. Gerace (1999). "A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export." The Journal of cell biology **145**(4): 645-657.

Kellogg, G. E., S. F. Semus and D. J. Abraham (1991). "HINT: a new method of empirical hydrophobic field calculation for CoMFA." Journal of computer-aided molecular design **5**(6): 545-552.

Kim, M., W. G. Chen, J. W. Kang, M. J. Glassman, K. Ribbeck and B. D. Olsen (2015). "Artificially Engineered Protein Hydrogels Adapted from the Nucleoporin Nsp1 for Selective Biomolecular Transport." Advanced Materials **27**(28): 4207-4212.

Kim, T., M. L. Gardel and E. Munro (2014). "Determinants of fluidlike behavior and effective viscosity in cross-linked actin networks." Biophysical journal **106**(3): 526-534.

Kim, T., W. Hwang and R. Kamm (2009a). "Computational analysis of a cross-linked actin-like network." Experimental Mechanics **49**(1): 91-104.

Kim, T., W. Hwang, H. Lee and R. D. Kamm (2009b). "Computational analysis of viscoelastic properties of crosslinked actin networks." PLoS computational biology **5**(7): e1000439.

Kiseleva, E., T. D. Allen, S. Rutherford, M. Bucci, S. R. Wente and M. W. Goldberg (2004). "Yeast nuclear pore complexes have a cytoplasmic ring and internal filaments." Journal of Structural Biology **145**(3): 272-288.

Kiseleva, E., M. W. Goldberg, T. D. Allen and C. W. Akey (1998). "Active nuclear pore complexes in Chironomus: visualization of transporter configurations related to mRNP export." Journal of Cell Science **111**: 223-236.

Kiseleva, E., M. W. Goldberg, B. Daneholt and T. D. Allen (1996). "RNP export is mediated by structural reorganization of the nuclear pore basket." Journal of Molecular Biology **260**(3): 304-311.

Klafter, J. and I. M. Sokolov (2005). "Anomalous diffusion spreads its wings." Physics world **18**(8): 29.

Klafter, J. and I. M. Sokolov (2011). First steps in random walks: from tools to applications, Oxford University Press.

Klenin, K., H. Merlitz and J. Langowski (1998). "A Brownian dynamics program for the simulation of linear and circular DNA and other wormlike chain polyelectrolytes." Biophysical Journal **74**(2): 780-788.

Kneller, J. M., C. Elvingson and G. A. Arteca (2005). "Shape transitions induced by mechanical external stretching of grafted self-attractive wormlike chains." Chemical physics letters **407**(4): 384-390.

Knudsen, K., J. G. de la Torre and A. Elgsaeter (1996). "Gaussian chains with excluded volume and hydrodynamic interaction: shear rate dependence of radius of gyration, intrinsic viscosity and flow birefringence." Polymer **37**(8): 1317-1322.

Koh, J. and G. Blobel (2015). "Allosteric regulation in gating the central channel of the nuclear pore complex." Cell **161**(6): 1361-1373.

Kolhar, P., A. C. Anselmo, V. Gupta, K. Pant, B. Prabhakarpandian, E. Ruoslahti and S. Mitragotri (2013). "Using shape effects to target antibody-coated nanoparticles to lung and brain endothelium." Proceedings of the National Academy of Sciences of the United States of America **110**(26): 10753-10758.

Kontopoulou, M. (2011). Applied polymer rheology: polymeric fluids with industrial applications, John Wiley & Sons.

Kose, S., N. Imamoto, T. Tachibana, M. Yoshida and Y. Yoneda (1999). "β-subunit of nuclear pore-targeting complex (importin-β) can be exported from the nucleus in a Ran-independent manner." Journal of Biological Chemistry **274**(7): 3946-3952.

Kosinski, J., S. Mosalaganti, A. von Appen, R. Teimer, A. L. DiGuilio, W. Wan, K. H. Bui, W. J. Hagen, J. A. Briggs and J. S. Glavy (2016). "Molecular architecture of the inner ring scaffold of the human nuclear pore complex." Science **352**(6283): 363-365.

Kotwaliwale, C. V. and A. F. Dernburg (2009). "Old Nuclei Spring New Leaks." Cell **136**(2): 211-212.

Kowalczyk, S. W., L. Kapinos, T. R. Blosser, T. Magalhães, P. van Nies, R. Y. Lim and C. Dekker (2011). "Single-molecule transport across an individual biomimetic nuclear pore complex." Nature Nanotechnology **6**(7): 433-438.

Kranbuehl, D. E. and P. H. Verdier (1979). "Simulation of the relaxation of random-coil polymer chains by lattice models with excluded volume: The effect of bead movement rules." The Journal of Chemical Physics **71**(6): 2662-2670.

Krishnamoorti, R. and E. P. Giannelis (1997). "Rheology of end-tethered polymer layered silicate nanocomposites." Macromolecules **30**(14): 4097-4102.

Kröger, M. (2005). Models for polymeric and anisotropic liquids, Springer Science & Business Media.

Kubitscheck, U., D. Grunwald, A. Hoekstra, D. Rohleder, T. Kues, J. P. Siebrasse and R. Peters (2005). "Nuclear transport of single molecules: dwell times at the nuclear pore complex." J Cell Biol **168**(2): 233-243.

Kubo, R. (1966). "The fluctuation-dissipation theorem." Reports on progress in physics **29**(1): 255.

Kuhn, v. W. and F. Grun (1942). "Relationships between elastic constants and stretching double refraction of highly elastic substances." Kolloid-Z **101**: 248-271.

Kumar, S. K. and J. F. Douglas (2001). "Gelation in physically associating polymer solutions." Physical review letters **87**(18): 188301.

Kuriyan, J., B. Konforti and D. Wemmer (2012). The molecules of life: Physical and chemical principles, Garland Science.

Kustanovich, T. and Y. Rabin (2004). "Metastable network model of protein transport through nuclear pores." Biophysical journal **86**(4): 2008-2016.

Larson, R. G. (1999). The Structure and Rheoloy of Complex Fluids. New York, Oxford University Press, Inc.

Lau, A. W., B. D. Hoffman, A. Davies, J. C. Crocker and T. C. Lubensky (2003). "Microrheology, stress fluctuations, and active behavior of living cells." Physical review letters **91**(19): 198101.

Leach, A. R. (2001). Molecular modelling principles and applications, Pearson Education.

Lee, G., P. G. Arscott, V. A. Bloomfield and D. F. Evans (1989). "Scanning tunneling microscopy of nucleic acids." Science **244**(4903): 475-477.

Lee, J. K., J. Barker and F. F. Abraham (1973). "Theory and Monte Carlo simulation of physical clusters in the imperfect vapor." The Journal of Chemical Physics **58**(8): 3166-3180.

Lemons, D. S. (2002). An introduction to stochastic processes in physics, JHU Press.

Lemons, D. S. and A. Gythiel (1997). "Paul Langevin's 1908 paper "on the theory of Brownian motion"["Sur la théorie du mouvement brownien," CR Acad. Sci.(Paris) 146, 530–533 (1908)]." American Journal of Physics **65**(11): 1079-1081.

Levine, A. J. and T. C. Lubensky (2000). "One- and two-particle microrheology." Phys Rev Lett **85**(8): 1774-1777.

Levitt, M. and A. Warshel (1975). "Computer simulation of protein folding." Nature **253**(5494): 694-698.

Lezon, T. R., A. Sali and I. Bahar (2009). "Global motions of the nuclear pore complex: insights from elastic network models." PLoS computational biology **5**(9): e1000496.

Li, L. and R. Larson (2000). "Excluded volume effects on the birefringence and stress of dilute polymer solutions in extensional flow." Rheologica acta **39**(5): 419-427.

Light, W. H., D. G. Brickner, V. R. Brand and J. H. Brickner (2010). "Interaction of a DNA zip code with the nuclear pore complex promotes H2A. Z incorporation and INO1 transcriptional memory." Molecular cell **40**(1): 112-125.

Light, W. H., J. Freaney, V. Sood, A. Thompson, A. D'Urso, C. M. Horvath and J. H. Brickner (2013). "A conserved role for human Nup98 in altering chromatin structure and promoting epigenetic transcriptional memory." PLoS biology **11**(3): e1001524.

Lim, R. Y., J. Koser, N. P. Huang, K. Schwarz-Herion and U. Aebi (2007a). "Nanomechanical interactions of phenylalanine-glycine nucleoporins studied by single molecule force-volume spectroscopy." J Struct Biol **159**(2): 277-289.

Lim, R. Y. H., B. Fahrenkrog, J. Koser, K. Schwarz-Herion, J. Deng and U. Aebi (2007b). "Nanomechanical Basis of Selective Gating by the Nuclear Pore Complex." Science **318**: 640-643.

Lim, R. Y. H., N. P. Huang, J. Koser, J. L. Deng, K.H.A., K. Schwarz-Herion, B. Fahrenkrog and U. Aebi (2006). "Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport." PNAS **103**: 9512–9517.

Liu, J., M. Gardel, K. Kroy, E. Frey, B. D. Hoffman, J. C. Crocker, A. Bausch and D. Weitz (2006). "Microrheology probes length scale dependent rheology." Physical review letters **96**(11): 118104.

Liverpool, T. and F. MacKintosh (2005). "Inertial effects in the response of viscous and viscoelastic fluids." Physical review letters **95**(20): 208303.

Lloyd, D. J. (1926). "The problem of gel structure." Colloid chemistry **1**: 767-782.

Ma, J., Z. Liu, N. Michelotti, S. Pitchiaya, R. Veerapaneni, J. R. Androsavich, N. G. Walter and W. Yang (2013). "High-resolution three-dimensional mapping of mRNA export through the nuclear pore." Nature communications **4**: 2414.

Macara, I. G. (2001a). "Transport into and out of the Nucleus." MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS **65**: 570-594.

Macara, I. G. (2001b). "Transport into and out of the nucleus." Microbiol Mol Biol Rev **65**(4): 570-594.

MacKintosh, F., J. Käs and P. Janmey (1995). "Elasticity of semiflexible biopolymer networks." Physical review letters **75**(24): 4425.

MacKintosh, F. and C. Schmidt (1999). "Microrheology." Current Opinion in Colloid & Interface Science **4**(4): 300-307.

Macosko, C. (1994). Rheology Principles, Measurements and Applications. New York, Wiley.

Maeda, S., Y. Hara, R. Yoshida and S. Hashimoto (2008). "Peristaltic motion of polymer gels." Angewandte Chemie **120**(35): 6792-6795.

Maimon, T., N. Elad, I. Dahan and O. Medalia (2012). "The human nuclear pore complex as revealed by cryo-electron tomography." Structure **20**(6): 998-1006.

Mao, B., M. Pear, J. McCammon and F. Quiocho (1982). "Hinge-bending in L-arabinose-binding protein. The" Venus's-flytrap" model." Journal of Biological Chemistry **257**(3): 1131-1133.

Marko, J. and E. Siggia (1995). "Stretching DNA." Macromolecules **28**(26): 8759-8770.

Marrink, S. J., A. H. De Vries and A. E. Mark (2004). "Coarse grained model for semiquantitative lipid simulations." The Journal of Physical Chemistry B **108**(2): 750-760.

Marrink, S. J., H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. De Vries (2007). "The MARTINI force field: coarse grained model for biomolecular simulations." The journal of physical chemistry B **111**(27): 7812-7824.

Marsaglia, G. and W. W. Tsang (2000). "The ziggurat method for generating random variables." Journal of statistical software **5**(8): 1-7.

Martínez, N., A. Alons, M. D. Moragues, J. Pontón and J. Schneider (1999). "The nuclear pore complex protein Nup88 is overexpressed in tumor cells." Cancer research **59**(21): 5408-5411.

Maruyama, K. (1997). "Connectin/titin, giant elastic protein of muscle." The FASEB journal **11**(5): 341-345.

Mason, T., K. Ganesan, J. Van Zanten, D. Wirtz and S. Kuo (1997). "Particle tracking microrheology of complex fluids." Physical Review Letters **79**(17): 3282.

Mason, T. and D. Weitz (1995). "Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids." Physical review letters **74**(7): 1250.

Mason, T. G. (2000). "Estimating the viscoelastic moduli of complex fluids using the generalized Stokes–Einstein equation." Rheologica Acta **39**(4): 371-378.

Massey, B. and J. Ward-Smith (2005). Mechanics of fluids 8th ed, Taylor & Francis e-Library.

Matsuura, Y. and M. Stewart (2004). "Structural basis for the assembly of a nuclear export complex." Nature **432**(7019): 872-877.

Maul, G. G. (1971). "On the octagonality of the nuclear pore complex." The Journal of cell biology **51**(2): 558.

McCammon, J. A., B. R. Gelin, M. Karplus and P. G. WOLYNES (1976). "The hinge-bending mode in lysozyme." Nature **262**(5566): 325-326.

McQuarrie, D. A. (1973). "Statistical thermodynamics."

Medina-Noyola, M. and D. A. McQuarrie (1980). "On the interaction of spherical double layers." The Journal of chemical physics **73**(12): 6279-6283.

Meijer, E. J. and D. Frenkel (1991). "Computer simulation of polymer-induced clustering of colloids." Physical review letters **67**(9): 1110.

Mendjan, S., M. Taipale, J. Kind, H. Holz, P. Gebhardt, M. Schelder, M. Vermeulen, A. Buscaino, K. Duncan and J. Mueller (2006). "Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila." Molecular cell **21**(6): 811-823.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller (1953). "Equation of state calculations by fast computing machines." The journal of chemical physics **21**(6): 1087-1092.

Mezzasalma, S. A. (2008). Macromolecules in solution and Brownian relativity, Academic Press.

Miao, L. and K. Schulten (2009a). "Transport-related structures and processes of the nuclear pore complex studied through molecular dynamics." Structure **17**(3): 449-459.

Miao, L. and K. Schulten (2009b). "Transport-related structures and processes of the nuclear pore complex studied through molecular dynamics." Structure **17**(3): 449-459.

Miao, L. and K. Schulten (2010a). "Probing a structural model of the nuclear pore complex channel through molecular dynamics." Biophysical journal **98**(8): 1658-1667.

Miao, L. and K. Schulten (2010b). "Probing a structural model of the nuclear pore complex channel through molecular dynamics." Biophys J **98**(8): 1658-1667.

Milchev, A., W. Paul and K. Binder (1993). "Off-lattice Monte Carlo simulation of dilute and concentrated polymer solutions under theta conditions." The Journal of chemical physics **99**(6): 4786-4798.

Milles, S., D. Mercadante, I. V. Aramburu, M. R. Jensen, N. Banterle, C. Koehler, S. Tyagi, J. Clarke, S. L. Shammas and M. Blackledge (2015). "Plasticity of an Ultrafast Interaction between Nucleoporins and Nuclear Transport Receptors." Cell **163**(3): 734-745.

Mincer, J. S. and S. M. Simon (2011). "Simulations of nuclear pore transport yield mechanistic insights and quantitative predictions." PNAS **108**(31): E351-E358.

Mohr, D., S. Frey, T. Fischer, T. Guttler and D. Gorlich (2009). "Characterisation of the passive permeability barrier of nuclear pore complexes." EMBO J **28**(17): 2541-2553.

Monticelli, L., S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S.-J. Marrink (2008). "The MARTINI coarse-grained force field: extension to proteins." Journal of chemical theory and computation **4**(5): 819-834.

Moore, D. S. (1985). "Amino acid and peptide net charges: a simple calculational procedure." Biochemistry and Molecular Biology Education **13**(1): 10-11.

Mor, A., S. Suliman, R. Ben-Yishay, S. Yunger, Y. Brody and Y. Shav-Tal (2010). "Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells." Nature cell biology **12**(6): 543-552.

Morrison, F. (2001). Understanding Rheology Morrison, OXFORD UNIVERSITY PRESS.

Morse, D. C. (1998). "Viscoelasticity of concentrated isotropic solutions of semiflexible polymers. 2. Linear response." Macromolecules **31**(20): 7044-7067.

Moussavi-Baygi, R., Y. Jamali, R. Karimi and M. R. Mofrad (2011a). "Biophysical coarse-grained modeling provides insights into transport through the nuclear pore complex." Biophys J **100**(6): 1410-1419.

Moussavi-Baygi, R., Y. Jamali, R. Karimi and M. R. Mofrad (2011b). "Brownian dynamics simulation of nucleocytoplasmic transport: a coarse-grained model for the functional state of the nuclear pore complex." PLoS computational biology **7**(6): e1002049.

Moussavi-Baygi, R. and M. R. K. Mofrad (2016). "Rapid Brownian Motion Primes Ultrafast Reconstruction of Intrinsically Disordered Phe-Gly Repeats Inside the Nuclear Pore Complex." Scientific reports **6**.

Mücke, N., L. Kreplak, R. Kirmse, T. Wedig, H. Herrmann, U. Aebi and J. Langowski (2004). "Assessing the flexibility of intermediate filaments by atomic force microscopy." Journal of molecular biology **335**(5): 1241-1250.

Munson, B. R., T. H. Okiishi, A. P. Rothmayer and W. W. Huebsch (2014). Fundamentals of fluid mechanics, John Wiley & Sons.

Murase, Y., S. Maeda, S. Hashimoto and R. Yoshida (2008). "Design of a mass transport surface utilizing peristaltic motion of a self-oscillating gel." Langmuir **25**(1): 483-489.

Nap, R., P. Gong and I. Szleifer (2006). "Weak polyelectrolytes tethered to surfaces: effect of geometry, acid–base equilibrium and electrical permittivity." Journal of Polymer Science Part B: Polymer Physics **44**(18): 2638-2662.

Neave, H. R. (1973). "On using the Box-Muller transformation with multiplicative congruential pseudo-random number generators." Applied Statistics: 92-97.

Nelson, D. L. and M. M. Cox (2013a). Lehninger principles of biochemistry, Macmillan.

Nelson, E., E. Nelson, E. Nelson and E. Nelson (1967). Dynamical theories of Brownian motion, Princeton university press Princeton.

Nelson, P. (2013b). Biological Physics. New York, Freeman.

Neogi, P. (1996). Diffusion in polymers, CRC Press.

Netz, R. R. and D. Andelman (2003). "Neutral and charged polymers at interfaces." Physics Reports **380**(1): 1-95.

Neves-Petersen, M. T. and S. B. Petersen (2003). "Protein electrostatics: a review of the equations and methods used to model electrostatic equations in biomolecules–applications in biotechnology." Biotechnol Annu Rev **9**(315): 95.

Nielsen, B., C. Jeppesen and J. H. Ipsen (2006). "Managing free-energy barriers in nuclear pore transport." Journal of biological physics **32**(5): 465-472.

Nishinari, K. (2009). Some thoughts on the definition of a gel. Gels: Structures, Properties, and Functions, Springer**:** 87-94.

Oberhauser, A. F., P. E. Marszalek, H. P. Erickson and J. M. Fernandez (1998). "The molecular elasticity of the extracellular matrix protein tenascin." Nature **393**(6681): 181-185.

Obradovic, Z., K. Peng, S. Vucetic, P. Radivojac, C. J. Brown and A. K. Dunker (2003). "Predicting intrinsic disorder from amino acid sequence." Proteins: Structure, Function, and Bioinformatics **53**(S6): 566-572.

Obradovic, Z., K. Peng, S. Vucetic, P. Radivojac and A. K. Dunker (2005). "Exploiting heterogeneous sequence properties improves prediction of protein disorder." Proteins: Structure, Function, and Bioinformatics **61**(S7): 176-182.

Ogden, R. W., G. Saccomandi and I. Sgura (2006). "On worm-like chain models within the three-dimensional continuum mechanics framework." Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences **462**(2067): 749-768.

Ohashi, T., S. D. Galiacy, G. Briscoe and H. P. Erickson (2007). "An experimental study of GFP-based FRET, with application to intrinsically unstructured proteins." Protein science : a publication of the Protein Society **16**(7): 1429-1438.

Ohashi, T., D. P. Kiehart and H. P. Erickson (1999). "Dynamics and elasticity of the fibronectin matrix in living cell culture visualized by fibronectin–green fluorescent protein." Proceedings of the National Academy of Sciences **96**(5): 2153-2158.

Ori, A., N. Banterle, M. Iskar, A. Andrés-Pons, C. Escher, H. K. Bui, L. Sparks, V. Solis-Mezarino, O. Rinner and P. Bork (2013). "Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines." Molecular systems biology **9**(1): 648.

Osada, Y. and K. Kajiwara, Eds. (1997). Ch. 2, Sec. 6 in "Gels Handbook".

Osada, Y., R. Kawamura and K.-I. Sano (2016). Hydrogels of Cytoskeletal Proteins.

Osmanovic, D., J. Bailey, A. H. Harker, A. Fassati, B. W. Hoogenboom and I. J. Ford (2012). "Bistable collective behavior of polymers tethered in a nanopore." Physical Review E **85**(6): 061917.

Osmanović, D., I. J. Ford and B. W. Hoogenboom (2013). "Model inspired by nuclear pore complex suggests possible roles for nuclear transport receptors in determining its structure." Biophysical journal **105**(12): 2781-2789.

Osmanović, D., M. Kerr-Winter, R. Eccleston, B. Hoogenboom and I. Ford (2015). "Effects of rotational symmetry breaking in polymer-coated nanopores." The Journal of chemical physics **142**(3): 034901.

Osmanovic, D. and Y. Rabin (2017). "Effect of Grafting on Aggregation of Intrinsically Disordered Proteins." arXiv preprint arXiv:1706.03337.

Öttinger, H. C. (1996). Stochastic Processes in Polymeric Fluids. Berlin, Springer.

Palmer, A., T. Mason, J. Xu, S. Kuo and D. Wirtz (1999). "Diffusing wave spectroscopy microrheology of actin filament networks." Biophysical journal **76**(2): 1063-1071.

Pante, N. and M. Kann (2002). "Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm." Mol Biol Cell **13**(2): 425-434.

Park, S. K. and K. W. Miller (1988). "Random number generators: good ones are hard to find." Communications of the ACM **31**(10): 1192-1201.

Pascual-Garcia, P., B. Debo, J. R. Aleman, J. A. Talamas, Y. Lan, N. H. Nguyen, K. J. Won and M. Capelson (2017). "Metazoan nuclear pores provide a scaffold for poised genes and mediate induced enhancer-promoter contacts." Molecular Cell **66**(1): 63-76. e66.

Patel, S. S., B. J. Belmont, J. M. Sante and M. F. Rexach (2007). "Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex." Cell **129**(1): 83-96.

Pathria, R. and P. Beale (2011). Statistical mechanics, Amsterdam.

Pei, H., S. Allison, B. M. Haynes and D. Augustin (2009). "Brownian dynamics simulation of the diffusion of rods and wormlike chains in a gel modeled as a cubic lattice: application to DNA." J Phys Chem B **113**(9): 2564-2571.

Peleg, O., M. Tagliazucchi, M. Kröger, Y. Rabin and I. Szleifer (2011). "Morphology control of hairy nanopores." ACS nano **5**(6): 4737-4747.

Perez-Terzic, C., A. M. Gacy, R. Bortolon, P. P. Dzeja, M. Puceat, M. Jaconi, F. G. Prendergast and A. Terzic (1999). "Structural plasticity of the cardiac nuclear pore complex in response to regulators of nuclear import." Circulation research **84**(11): 1292-1301.

Peters, R. (2005). "Translocation through the nuclear pore complex: selectivity and speed by reduction-of-dimensionality." Traffic **6**(5): 421-427.

Peters, R. (2006). Introduction to nucleocytoplasmic transport. Xenopus Protocols, Springer**:** 235-258.

Peters, R. (2009). "Translocation through the nuclear pore: Kaps pave the way." Bioessays **31**(4): 466-477.

Petros, R. A. and J. M. DeSimone (2010). "Strategies in the design of nanoparticles for therapeutic applications." Nature reviews Drug discovery **9**(8): 615-627.

Peyro, M., M. Soheilypour, A. Ghavami and M. R. Mofrad (2015a). "Nucleoporin's like charge regions are major regulators of FG coverage and dynamics inside the nuclear pore complex." PloS one **10**(12): e0143745.

Peyro, M., M. Soheilypour, A. Ghavami and M. R. K. Mofrad (2015b). "Nucleoporin's Like Charge Regions Are Major Regulators of FG Coverage and Dynamics Inside the Nuclear Pore Complex." PLoS One **10**(12).

Peyro, M., M. Soheilypour, B. Lee and M. Mofrad (2015c). "Evolutionarily Conserved Sequence Features Regulate the Formation of the FG Network at the Center of the Nuclear Pore Complex." Scientific reports **5**.

Phillips, R., J. Kondev, J. Theriot and H. Garcia (2012). Physical biology of the cell, Garland Science.

Pierleoni, C. and J. P. Ryckaert (2000). "Excluded volume effects on the structure of a linear polymer under shear flow." Journal of Chemical Physics **113**(13): 5545-5558.

Popescu, G., A. Dogariu and R. Rajagopalan (2002). "Spatially resolved microrheology using localized coherence volumes." Physical Review E **65**(4): 041504.

Popken, P., A. Ghavami, P. R. Onck, B. Poolman and L. M. Veenhoff (2015a). "Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex." Molecular Biology of the Cell **26**(7): 1386-1394.

Popken, P., A. Ghavami, P. R. Onck, B. Poolman and L. M. Veenhoff (2015b). "Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex." Molecular biology of the cell **26**(7): 1386-1394.

Prabhakar, R. and J. R. Prakash (2004a). "Multiplicative separation of the influences of excluded volume, hydrodynamic interactions and finite extensibility on the rheological properties of dilute polymer solutions." Journal of non-newtonian fluid mechanics **116**(2): 163-182.

Prabhakar, R. and J. R. Prakash (2004b). "Multiplicative separation of the influences of excluded volume, hydrodynamic interactions and finite extensibility on the rheological properties of dilute polymer solutions." Journal of Non-Newtonian Fluid Mechanics **116**(2-3): 163-182.

Pryamitsyn, V. and V. Ganesan (2008). "Screening of hydrodynamic interactions in Brownian rod suspensions." The Journal of chemical physics **128**(13): 134901.

Pulupa, J., M. Rachh, M. D. Tomasini, J. S. Mincer and S. M. Simon (2017). "A coarse-grained computational model of the nuclear pore complex predicts Phe-Gly nucleoporin dynamics." The Journal of General Physiology: jgp. 201711769.

Purcell, E. M. (1977). "Life at low Reynolds number." American journal of physics **45**(1): 3-11.

Pyhtila, B. and M. Rexach (2003a). "A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex." J Biol Chem **278**(43): 42699-42709.

Pyhtila, B. and M. Rexach (2003b). "A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex." Journal of Biological Chemistry **278**(43): 42699-42709.

Quiroz, F. G. and A. Chilkoti (2015). "Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers." Nature Materials **14**: 1164–1171.

Rabut, G., V. Doye and J. Ellenberg (2004). "Mapping the dynamic organization of the nuclear pore complex inside single living cells." Nat Cell Biol **6**(11): 1114-1121.

Radu, A., G. Blobel and M. S. Moore (1995). "Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins." Proceedings of the National Academy of Sciences **92**(5): 1769-1773.

Radzicka, A. and R. Wolfenden (1988). "Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution." Biochemistry **27**(5): 1664-1670.

Redondo, A. and R. LeSar (2004). "Modeling and simulation of biomaterials." Annu. Rev. Mater. Res. **34**: 279-314.

Reichelt, R., A. Holzenburg, E. L. Buhle, Jr., M. Jarnik, A. Engel and U. Aebi (1990). "Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components." J Cell Biol **110**(4): 883-894.

Reichl, L. E. (2016). A modern course in statistical physics, John Wiley & Sons.

Reiner, M. (1964). "The deborah number." Physics today **17**(1): 62.

Rexach, M. and G. Blobel (1995). "Protein Import into Nuclei: Association and Dissociation Reactions Involving Transport Substrate, Transport Factors, and Nucleoporins." Cell **63**: 683-692.

Rey, A., J. J. Freire and J. de la Torre (1992). "Brownian dynamics simulation of flexible polymer chains with excluded volume and hydrodynamic interactions. A comparison with Monte Carlo and theoretical results." Polymer **33**(16): 3477-3481.

Ribbeck, K. and D. Gorlich (2001). "Kinetic analysis of translocation through nuclear pore complexes." EMBO J **20**(6): 1320-1330.

Ribbeck, K. and D. Gorlich (2002). "The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion." EMBO J **21**(11): 2664-2671.

Riddick, G. and I. G. Macara (2005). "A systems analysis of importin-α–β mediated nuclear protein import." The Journal of cell biology **168**(7): 1027-1038.

Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez and H. E. Gaub (1997). "Reversible unfolding of individual titin immunoglobulin domains by AFM." science **276**(5315): 1109-1112.

Rodriguez, A. G., S. Han, M. Regnier and N. Sniadecki (2010). "The Effect of Stiffness and Beta-Adrenergic Stimulation on Neonatal Cardiomyocyte Calcium-Mediated Contractile Force Dynamics." Biophysical Journal **98**(3): 757a.

Rogers, S. S., T. A. Waigh and J. R. Lu (2008). "Intracellular microrheology of motile Amoeba proteus." Biophysical Journal **94**(8): 3313-3322.

Romero, P., Z. Obradovic, X. Li, E. C. Garner, C. J. Brown and A. K. Dunker (2001). "Sequence complexity of disordered protein." Proteins: Structure, Function, and Bioinformatics **42**(1): 38-48.

Rose, G. D., A. R. Geselowitz, G. J. Lesser, R. H. Lee and M. H. Zehfus (1985). "Hydrophobicity of amino acid residues in globular proteins." Science **229**(4716): 834-838.

Ross-Murphy, S. B. (1998). "Reversible and irreversible biopolymer gels—structure and mechanical properties." Berichte der Bunsengesellschaft für physikalische Chemie **102**(11): 1534-1539.

Rouse Jr, P. E. (1953). "A theory of the linear viscoelastic properties of dilute solutions of coiling polymers." The Journal of Chemical Physics **21**(7): 1272-1280.

Rout, M. P., J. D. Aitchison, M. O. Magnasco and B. T. Chait (2003). "Virtual gating and nuclear transport: the hole picture." Trends Cell Biol **13**(12): 622-628.

Rowat, A. C., L. Foster, M. Nielsen, M. Weiss and J. H. Ipsen (2005). "Characterization of the elastic properties of the nuclear envelope." Journal of The Royal Society Interface **2**(2): 63-69.

Rubinstein, M. and R. Colby (2003). Polymer Physics. Oxford, Oxford University Press.

Ryan, G. L., H. M. Petroccia, N. Watanabe and D. Vavylonis (2012). "Excitable actin dynamics in lamellipodial protrusion and retraction." Biophys J **102**(7): 1493-1502.

Satterly, N., P. L. Tsai, J. van Deursen, D. R. Nussenzveig, Y. Wang, P. A. Faria, A. Levay, D. E. Levy and B. M. Fontoura (2007). "Influenza virus targets the mRNA export machinery and the nuclear pore complex." Proc Natl Acad Sci U S A **104**(6): 1853-1858.

Saxton, M. J. and K. Jacobson (1997). "Single-particle tracking: applications to membrane dynamics." Annual review of biophysics and biomolecular structure **26**(1): 373-399.

Schellman, J. A. (1974). "Flexibility of DNA." Biopolymers **13**(1): 217-226.

Schlick, T. (1989). "A Recipe for Evaluating and Differentiating Cos-Phi Expressions." Journal of Computational Chemistry **10**(7): 951-956.

Schlick, T. (2010). Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide, Springer Science & Business Media.

Schnurr, B., F. Gittes, F. MacKintosh and C. Schmidt (1997). "Determining microscopic viscoelasticity in flexible and semiflexible polymer networks from thermal fluctuations." Macromolecules **30**(25): 7781-7792.

Schoch, R. L., L. E. Kapinos and R. Y. Lim (2012). "Nuclear transport receptor binding avidity triggers a self-healing collapse transition in FG-nucleoporin molecular brushes." PNAS **109**(42): 16911-16916.

Scott, W. R., P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger and W. F. van Gunsteren (1999). "The GROMOS biomolecular simulation program package." The Journal of Physical Chemistry A **103**(19): 3596-3607.

Seedorf, M., M. Damelin, J. Kahana, T. Taura and P. A. Silver (1999). "Interactions between a nuclear transporter and a subset of nuclear pore complex proteins depend on Ran GTPase." Molecular and Cellular Biology **19**(2): 1547-1557.

Sellés, J., M. Penrad-Mobayed, C. Guillaume, A. Fuger, L. Auvray, O. Faklaris and F. Montel (2017). "Nuclear pore complex plasticity during developmental process as revealed by super-resolution microscopy." Scientific Reports **7**(1): 14732.

Sexton, T., H. Schober, P. Fraser and S. M. Gasser (2007). "Gene regulation through nuclear organization." Nature structural & molecular biology **14**(11): 1049-1055.

Shah, S., S. Tugendreich and D. Forbes (1998). "Major binding sites for the nuclear import receptor are the internal nucleoporin Nup153 and the adjacent nuclear filament protein Tpr." The Journal of cell biology **141**(1): 31-49.

Shenoy, A. V. (2013). Rheology of filled polymer systems, Springer Science & Business Media.

Shibata, S., Y. Matsuoka and Y. Yoneda (2002). "Nucleocytoplasmic transport of proteins and poly (A)+ RNA in reconstituted Tpr-less nuclei in living mammalian cells." Genes to Cells **7**(4): 421-434.

Shinohara, S.-i., T. Seki, T. Sakai, R. Yoshida and Y. Takeoka (2008). "Chemical and optical control of peristaltic actuator based on self-oscillating porous gel." Chem. Commun.(39): 4735-4737.

Shulga, N., N. Mosammaparast, R. Wozniak and D. S. Goldfarb (2000). "Yeast nucleoporins involved in passive nuclear envelope permeability." J Cell Biol **149**(5): 1027-1038.

Siebrasse, J. P., T. Kaminski and U. Kubitscheck (2012). "Nuclear export of single native mRNA molecules observed by light sheet fluorescence microscopy." Proceedings of the National Academy of Sciences of the United States of America **109**(24): 9426-9431.

Siginer, D. A., D. De Kee and R. P. Chhabra (1999). Advances in the flow and rheology of non-Newtonian fluids, Elsevier.

Simon, S. M., C. S. Peskin and G. F. Oster (1992). "What drives the translocation of proteins?" Proceedings of the National Academy of Sciences **89**(9): 3770-3774.

Smit, B. (1995). "Grand canonical Monte Carlo simulations of chain molecules: adsorption isotherms of alkanes in zeolites." Molecular Physics **85**(1): 153-172.

Smith, A. E., B. M. Slepchenko, J. C. Schaff, L. M. Loew and I. G. Macara (2002). "Systems analysis of Ran transport." Science **295**(5554): 488-491.

Smith, S. B., Y. Cui and C. Bustamante (1996). "Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules." Science: 795-799.

Smith, S. B., L. Finzi and C. Bustamante (1992). "Direct Mechanical Measurements of the Elasticity of Single DNA-Molecules by Using Magnetic Beads." Science **258**(5085): 1122-1126.

Snook, I. (2006). The Langevin and generalised Langevin approach to the dynamics of atomic, polymeric and colloidal systems, Elsevier.

Soheilypour, M. and M. Mofrad (2016). "Regulation of RNA-binding proteins affinity to export receptors enables the nuclear basket proteins to distinguish and retain aberrant mRNAs." Scientific reports **6**: 35380.

Spring, J., K. Beck and R. Chiquet-Ehrismann (1989). "Two contrary functions of tenascin: dissection of the active sites by recombinant tenascin fragments." Cell **59**(2): 325-334.

Squires, T. M. and T. G. Mason (2010). "Fluid mechanics of microrheology." Annual review of fluid mechanics **42**.

Stachel, J. and V. V. Raman, Eds. (1990). Eidtorial note: "Einstein on Brownian motion", The Collected Papers of Albert Einstein. Vol. 2. The Swiss Years: Writings, 1900– 1909. American Journal of Physics.

Stachel, J. J. (2005). "Einstein's miraculous year: five papers that changed the face of physics."

Starr, B. D., B. C. Hoopes and D. K. Hawley (1995). "DNA bending is an important component of site-specific recognition by the TATA binding protein." Journal of molecular biology **250**(4): 434-446.

Starrs, L. and P. Bartlett (2002). "One-and two-point micro-rheology of viscoelastic media." Journal of Physics: Condensed Matter **15**(1): S251.

Stauffer, D. (1976). "Gelation in concentrated critically branched polymer solutions. Percolation scaling theory of intramolecular bond cycles." Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics **72**: 1354-1364.

Stellman, S. D. and P. J. Gans (1972). "Efficient computer simulation of polymer conformation. I. Geometric properties of the hard-sphere model." Macromolecules **5**(4): 516-526.

Stewart, M. (2007). "Molecular mechanism of the nuclear protein import cycle." Nature Reviews Mol Cell Biol **8**(3): 195-208.

Still, G. K. (1993). "New computer system can predict human behavioural response to building fires." Fire **85**(1051): 40.

Stockmayer, W. H. (1943). "Theory of molecular size distribution and gel formation in branched-chain polymers." The Journal of chemical physics **11**(2): 45-55.

Stoffler, D., B. Feja, B. Fahrenkrog, J. Walz, D. Typke and U. Aebi (2003). "Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport." J Mol Biol **328**(1): 119-130.

Stoffler, D., K. N. Goldie, B. Feja and U. Aebi (1999). "Calcium-mediated structural changes of native nuclear pore complexes monitored by time-lapse atomic force microscopy." J Mol Biol **287**(4): 741-752.

Storm, C., J. J. Pastore, F. C. MacKintosh, T. C. Lubensky and P. A. Janmey (2005). "Nonlinear elasticity in biological gels." Nature **435**(7039): 191-194.

Strambio-De-Castillia, C., M. Niepel and M. P. Rout (2010a). "The nuclear pore complex: bridging nuclear transport and gene regulation." Nature Reviews Mol Cell Biol **11**(7): 490-501.

Strambio-De-Castillia, C., M. Niepel and M. P. Rout (2010b). "The nuclear pore complex: bridging nuclear transport and gene regulation." Nature reviews Molecular cell biology **11**(7): 490-501.

Strawn, L. A., T. Shen, N. Shulga, D. S. Goldfarb and S. R. Wente (2004). "Minimal nuclear pore complexes define FG repeat domains essential for transport." Nature cell biology **6**(3): 197-206.

Suntharalingam, M. and S. R. Wente (2003). "Peering through the pore: Nuclear pore complex structure, assembly, and function." Developmental Cell **4**(6): 775-789.

Sutherland, W. (1905). "LXXV. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin." The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **9**(54): 781-785.

Swaminathan, R., C. P. Hoang and A. Verkman (1997). "Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion." Biophysical journal **72**(4): 1900-1907.

Szleifer, I. and M. Carignano (2000). "Tethered polymer layers: phase transitions and reduction of protein adsorption." Macromolecular rapid communications **21**(8): 423-448.

Tagliazucchi, M., O. Peleg, M. Kröger, Y. Rabin and I. Szleifer (2013). "Effect of charge, hydrophobicity, and sequence of nucleoporins on the translocation of model particles through the nuclear pore complex." PNAS **110**(9): 3363-3368.

Takizawa, T., K. J. Meaburn and T. Misteli (2008). "The meaning of gene positioning." Cell **135**(1): 9-13.

Tanford, C. (1980). The hydrophobic effect: formation of micelles and biological membranes 2d ed, J. Wiley.

te Nijenhuis, K. (1997). Gelatin, Springer.

Teraoka, I. (2002a). POLYMER SOLUTIONS. New York, John Wiley & Sons, Inc.

Teraoka, I. (2002b). Polymer Solutions: An Introduction to Physical Properties, John Wiley & Sons, Inc.

Terry, L. J. and S. R. Wente (2009). "Flexible gates: dynamic topologies and functions for FG nucleoporins in nucleocytoplasmic transport." Eukaryotic Cell **8**(12): 1814-1827.

Tetenbaum-Novatt, J., L. E. Hough, R. Mironska, A. S. McKenney and M. P. Rout (2012). "Nucleocytoplasmic transport: a role for nonspecific competition in karyopherin-nucleoporin interactions." Molecular & Cellular Proteomics **11**(5): 31-46.

Thomas, D. B., W. Luk, P. H. W. Leong and J. D. Villasenor (2007). "Gaussian Random Number Generators." Association for Computing Machinery **39**(4): 11:11-11:38.

Timney, B. L., J. Tetenbaum-Novatt, D. S. Agate, R. Williams, W. Zhang, B. T. Chait and M. P. Rout (2006). "Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo." J Cell Biol **175**(4): 579-593.

Tirado, M. M. and J. G. de la Torre (1979). "Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders." The Journal of chemical physics **71**(6): 2581-2587.

Tompa, P. (2002). "Intrinsically unstructured proteins." TRENDS in Biochemical Sciences **27**: 527-533.

Underhill, P. T. and P. S. Doyle (2004a). "On the coarse-graining of polymers into bead-spring chains." Journal of Non-Newtonian Fluid Mechanics **122**(1-3): 3-31.

Underhill, P. T. and P. S. Doyle (2004b). "On the coarse-graining of polymers into bead-spring chains." Journal of non-newtonian fluid mechanics **122**(1): 3-31.

Underhill, P. T. and P. S. Doyle (2007). "Accuracy of bead-spring chains in strong flows." Journal of Non-Newtonian Fluid Mechanics **145**(2): 109-123.

Uversky, V. N. (2008). "Amyloidogenesis of natively unfolded proteins." Current Alzheimer research **5**(3): 260-287.

Uversky, V. N. (2011). "Intrinsically disordered proteins from A to Z." The international journal of biochemistry & cell biology **43**(8): 1090-1103.

Valentine, M., Z. Perlman, M. Gardel, J. Shin, P. Matsudaira, T. Mitchison and D. Weitz (2004). "Colloid surface chemistry critically affects multiple particle tracking measurements of biomaterials." Biophysical journal **86**(6): 4004-4014.

Valentine, M. T., P. D. Kaplan, D. Thota, J. C. Crocker, T. Gisler, R. K. Prud'homme, M. Beck and D. A. Weitz (2001). "Investigating the microenvironments of inhomogeneous soft materials with multiple particle tracking." Physical Review E **64**(6): 061506.

Van den Brule, B. (1993). "Browian dynamics simulation of finitely extensible bead-spring chains." Journal of non-newtonian fluid mechanics **47**: 357-378.

Van den Brule, B. and P. Hoogerbrugge (1995). "Brownian dynamics simulation of reversible polymeric networks." Journal of non-newtonian fluid mechanics **60**(2-3): 303-334.

van der Aa, M. A., E. Mastrobattista, R. S. Oosting, W. E. Hennink, G. A. Koning and D. J. Crommelin (2006). "The nuclear pore complex: the gateway to successful nonviral gene delivery." Pharm Res **23**(3): 447-459.

Vanzi, F., Y. Takagi, S. , H, B. Cooperman and Y. Goldman (2005). "Mechanical studies of single ribosome/mRNA complexes." Biophys J **89**: 1909–1919.

Vaquerizas, J. M., R. Suyama, J. Kind, K. Miura, N. M. Luscombe and A. Akhtar (2010). "Nuclear pore proteins nup153 and megator define transcriptionally active regions in the Drosophila genome." PLoS genetics **6**(2): e1000846.

Vaziri, A. and M. R. K. Mofrad (2007). "Mechanics and deformation of the nucleus in micropipette aspiration experiment." Journal of Biomechanics **40**(9): 2053-2062.

Vázquez-Quesada, A., M. Ellero and P. Español (2012). "A SPH-based particle model for computational microrheology." Microfluidics and nanofluidics **13**(2): 249-260.

von Appen, A., J. Kosinski, L. Sparks, A. Ori, A. L. DiGuilio, B. Vollmer, M.-T. Mackmull, N. Banterle, L. Parca and P. Kastritis (2015). "In situ structural analysis of the human nuclear pore complex." Nature **526**: 140-143.

Von Smoluchowski, M. (1906). "Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen." Annalen der physik **326**(14): 756-780.

Waigh, T. A. (2005). "Microrheology of complex fluids." Reports on Progress in Physics **68**(3): 685-742.

Wang, M. D., H. Yin, R. Landick, J. Gelles and S. M. Block (1997). "Stretching DNA with optical tweezers." Biophysical journal **72**(3): 1335-1346.

Wang, W., O. Donini, C. M. Reyes and P. A. Kollman (2001). "Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions." Annual review of biophysics and biomolecular structure **30**(1): 211-243.

Warner Jr, H. R. (1972). "Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells." Industrial & Engineering Chemistry Fundamentals **11**(3): 379-387.

Watson, M. L. (1955). "The nuclear envelope: Its structure and relation to cytoplasmic membranes." The Journal of biophysical and biochemical cytology **1**(3): 257.

Weis, K. (2002). "Nucleocytoplasmic transport: cargo trafficking across the border." Curr Opin Cell Biol **14**(3): 328-335.

Weiss, M., M. Elsner, F. Kartberg and T. Nilsson (2004). "Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells." Biophysical journal **87**(5): 3518-3524.

Wente, S. R. and M. P. Rout (2010). "The nuclear pore complex and nuclear transport." Cold Spring Harbor perspectives in biology **2**(10): a000562.

Whittaker, G. R., M. Kann and A. Helenius (2000). "Viral entry into the nucleus." Annu Rev Cell Dev Biol **16**: 627-651.

Wiest, J. and R. Tanner (1989). "Rheology of Bead-NonLinear Spring Chain Macromolecules." Journal of Rheology **33**(2): 281-316.

Wolf, C. and M. R. K. Mofrad (2008). "On the Octagonal Structure of the Nuclear Pore Complex: Insights from Coarse-Grained Models." Biophys J **95**: 2073–2085.

Xie, Y. J., H. Y. Yang, H. T. Yu, Q. W. Shi, X. P. Wang and J. Chen (2006). "Excluded volume effect on confined polymer translocation through a short nanochannel." Journal of Chemical Physics **124**(17).

Xu, J., A. Palmer and D. Wirtz (1998). "Rheology and microrheology of semiflexible polymer solutions: actin filament networks." Macromolecules **31**(19): 6486-6492.

Xu, K., M. G. Forest and I. Klapper (2007). "On the correspondence between creeping flows of viscous and viscoelastic fluids." Journal of Non-Newtonian Fluid Mechanics **145**(2): 150-172.

Xu, S. and M. A. Powers (2013). "In vivo analysis of human nucleoporin repeat domain interactions." Molecular biology of the cell **24**(8): 1222-1231.

Yamada, J., J. L. Phillips, S. Patel, G. Goldfien, A. Calestagne-Morelli, H. Huang, R. Reza, J. Acheson, V. V. Krishnan, S. Newsam, A. Gopinathan, E. Y. Lau, M. E. Colvin, V. N. Uversky and M. F. Rexach (2010). "A bimodal distribution of two distinct categories of intrinsically-disordered structures with separate functions in FG nucleoporins." Mol Cell Proteomics **9**: 2205-2224.

Yamakawa, H. (1971). Modern theory of polymer solutions, Harper & Row.

Yamakawa, H. and M. Fujii (1974). "Intrinsic viscosity of wormlike chains. Determination of the shift factor." Macromolecules **7**(1): 128-135.

Yang, G., C. Cecconi, W. A. Baase, I. R. Vetter, W. A. Breyer, J. A. Haack, B. W. Matthews, F. W. Dahlquist and C. Bustamante (2000). "Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme." Proceedings of the National Academy of Sciences **97**(1): 139-144.

Yang, Q., M. P. Rout and C. W. Akey (1998). "Three-Dimensional Architecture of the Isolated Yeast Nuclear Pore Complex: Functional and Evolutionary Implications." Molecular Cell **1**: 223–234.

Yang, W. and S. M. Musser (2006). "Nuclear import time and transport efficiency depend on importin beta concentration." J Cell Biol **174**(7): 951-961.

Yang, W. D., J. Gelles and S. M. Musser (2004). "Imaging of single-molecule translocation through nuclear pore complexes." Proceedings of the National Academy of Sciences of the United States of America **101**(35): 12887-12892.

Yokokawa, M., K. Takeyasu and S. H. Yoshimura (2008). "Mechanical properties of plasma membrane and nuclear envelope measured by scanning probe microscope." Journal of Microscopy-Oxford **232**(1): 82-90.

Yoon, R.-H., D. H. Flinn and Y. I. Rabinovich (1997). "Hydrophobic interactions between dissimilar surfaces." Journal of colloid and interface science **185**(2): 363-370.

Zachariae, U. and H. Grubmüller (2006). "A highly strained nuclear conformation of the exportin Cse1p revealed by molecular dynamics simulations." Structure **14**(9): 1469-1478.

Zaner, K. S. and P. A. Valberg (1989). "Viscoelasticity of F-actin measured with magnetic microparticles." The Journal of cell biology **109**(5): 2233-2243.

Zhao, C. L., S. H. Mahboobi, R. Moussavi-Baygi and M. R. Mofrad (2014). "The interaction of CRM1 and the nuclear pore protein Tpr." PLoS One **9**(4): e93709.

Zilman, A., S. Di Talia, B. T. Chait, M. P. Rout and M. O. Magnasco (2007). "Efficiency, selectivity, and robustness of nucleocytoplasmic transport." PLoS computational biology **3**(7): e125.

Zilman, A., S. Di Talia, T. Jovanovic-Talisman, B. T. Chait, M. P. Rout and M. O. Magnasco (2010). "Enhancement of transport selectivity through nano-channels by non-specific competition." PLoS computational biology **6**(6): e1000804.

Zimowska, G., J. Aris and M. Paddy (1997). "A Drosophila Tpr protein homolog is localized both in the extrachromosomal channel network and to nuclear pore complexes." Journal of cell science **110**(8): 927-944.

Appendix A. Random numbers in the Brownian dynamic simulation

A.1 Random number generator

The displacement equation is calculated in every timestep for every particle in all three directions. For an N-particle system, it thus requires 3N independent normal random numbers in every timestep. Considering an ordinary simulation with 10^9 timesteps, a total of $3N \times 10^9$ random numbers are required to be generated to fulfill this particular simulation. Therefore, from a computational efficiency viewpoint, a fast RNG is essential for facilitating the simulation. Incidentally, this argument also makes it clear why an RNG must have a long period—the number of random numbers before the RNG repeats itself (Carter1994).

However, a crucially way more significant aspect of an RNG for a scientific simulation is its role in shaping the correct physical behavior of the model, which is the basis for any measurement, observation, and prediction made by that model. Considering $3N \times 10^9$ random numbers in each simulation, even an infinitesimal statistical deviation in random numbers from a true normal distribution is accumulated during the simulation, leading to nonphysical behavior of the model that is unlikely to be detected. It becomes more significant when force-fields with singularity are used in the simulation. The Lennard Jones potential and wormlike chain (WLC; see below) model are examples of such force-fields. In the latter caser, for instance, when the end-to-end length of a polymer approaches the contour length, an infinite retracting force would be applied to both ends of the polymer. An inappropriate RNG can lead to accumulation of nonphysical displacements, potentially leading to a big random jump and thus infinite force, which in turn would result in unrealistic behavior in the simulation and spurious results.

In some cases, like infinite force on the polymer segment, the nonphysical behavior can be easily detected by visualization or measurement. This is not, unfortunately, always the case and the effects of inappropriate RNG, on nonphysical displacement for example, might not be detected readily. It thus can invisibly affect the dynamics of the system, obscuring the analysis and leaving the results unreliable. It is, therefore, of extreme importance to ensure that the RNG used in the simulation is mathematically *and* statistically accurate, and fully consistent with the distribution of interest.

There are usually built-in RNGs for uniformly distributed random numbers over (0,1) in different languages, which then can be transformed into normal distribution. Probably one of the most important transformations was developed by Box and Muller (Box1958), converting a pair of independent uniform random numbers (U_1, U_2) to a pair of independent Gaussian random numbers (X_1, X_2) as:

$$(X_1, X_2) = \sqrt{-2\ln U_1}\left(\cos 2\pi U_2, \sin 2\pi U_2\right) \qquad \text{(A.1)}$$

While this approach is to some extent common among programmers, the current dissertation strongly recommends against it for scientific simulations for two reasons. First, the compilers' built-in RNGs might be useful in different computer applications and games, but they are not mathematically of interest in numerically analyzing stochastic processes. Secondly, many of transformations suffer from some kind of instability, which becomes prominent when they are

being used numerous times during the simulation. For example, if U_1 in eq. (A.1) approaches 0, the Box-Muller transformation becomes unstable, which leads to unsmooth distribution, particularly in the tails (Neave1973). In addition, it has been shown that the transformation can generate nonsymmetric distribution (Neave1973).

To the best of my experience, the safest, and scientifically most reliable, way of generating normally distributed random numbers for a stochastic simulation is to use mathematically rigorous algorithms that have been developed by mathematicians, rather than computer engineers, specifically for normal distribution without any external transformation, i.e. Gaussian RNG (GRNG). The caveat is that when it comes to the computational efficiency these algorithms might not be faster than, or even as fast as, algorithms developed by computer engineers. This is indeed of secondary importance relative to predicting the correct physics of the system, and thus should not be a concern.

In the current work, I used Ziggurat algorithm (Marsaglia2000) as a GRNG to produce normally distributed random numbers. The algorithm is mathematically rigorous and statistically stable. It is also computationally optimized and fast. Using Fortran 90 implementation of the Ziggurat algorithm, it takes about $5.17s$ to generate 10^9 random numbers, i.e. 2×10^8 random numbers per second, on a single core of the Stampede machine (Intel Xeon Phi processors with the clock speed of 2.7GHz) at the Texas Advanced Computing Center (TACC).

As illustrated in Fig. A.1, the histogram of generated random numbers is perfectly symmetric around zero. The real values of the distribution's mean and SD calculated in R (-0.0001978965 and 1.0001311152, respectively) show less 0.01% deviation from expected values (0 and 1).

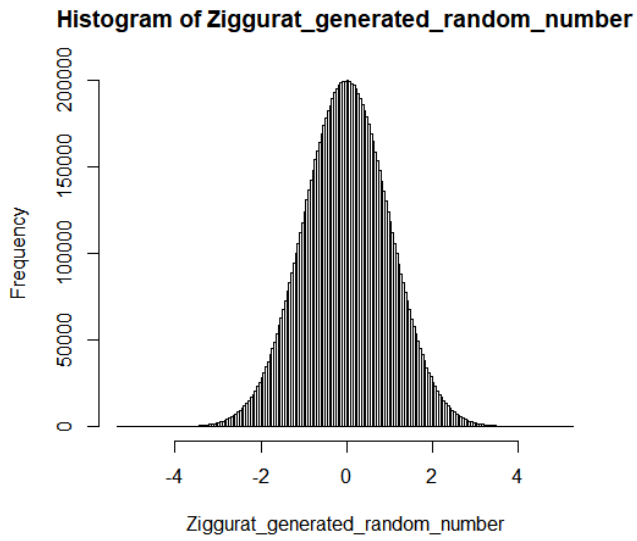


Figure A.1. The histogram of 10^9 random numbers generated using the Ziggurat algorithm. The random numbers were generated in Fortran 90 on TACC supercomputers. The histogram was plotted in R. The real mean and SD of these numbers are -0.0001978965 and 1.0001311152, while their expected values are 0.0 and 1.0, respectively.

Table A.1. The GRNG algorithms' speeds relative to the Polar Rejection, and the arithmetic operation counts per generated random number. U: the number of input uniform; C: number of counts used in the implantation of algorithm. Calculations are done in single precision. Adopted from (Thomas2007).

No.	GRNG Algorithm [published year]	Speed	*U*	+	×	÷	Cmp	$\sqrt{x}$	Ln, Exp, Trig	C
1	Wallace (qual = 1) [1996]	6.41	0.001	10.02	1.50	$\mathcal{E}$	1.51	$\mathcal{E}$		9
2	**Ziggurat [2000]**	**4.29**	**1.04**	**1.10**	**1.07**		**2.07**		**0.001, 0.03, 0**	**388**
3	Wallace (qual = 4) [1996]	2.48	0.003	37.07	3.01	$\mathcal{E}$	3.04	$\mathcal{E}$		9
4	Monty Python [1998]	1.61	1.30	0.88	1.96		2.57		0.03, 0, 0	16
5	PPND7 (ICDF) [1988]	1.16	1	8.15	7.40	1	1.45	0.15	0.15, 0, 0	26
6	Mixture-of-Triangles [2000]	1.14	3	3	2	1	1			122
7	Polar Rejection [1969]	1.00	1.27	1.91	3.30	1	1.27	1	1, 0, 0	4
8	Leva (Ratio) [1992b]	0.98	2.74	6.84	6.89	1	3.12		0.01, 0, 0	9
9	Marsaglia-Bray [1964]	0.94	3.92	3.22	1.36	0.01	1.42	0.006	0.01, 0.05, 0	15
10	GRAND [1974]	0.92	1.38	8.65	6.49	1.16	4.88			27
11	Box-Muller [1958b]	0.81	1		2		0	0.5	0.5, 0, 1	2
12	Ahrens-Dieter [1988]	0.78	1.02	4.55	4.04	1.5	4.51	0.5	0, 0.01, 0	20
13	Kinderman (Ratio) [1977]	0.76	2.74	3.20	4.34	1.84	3.44		0.23, 0, 0	6
14	Hastings (ICDF) [1959]	0.62	1	8	7	2	1	1	1, 0, 0	7
15	PPND16 (ICDF) [1988]	0.55	1	14.45	14.85	1	1.45	0.15	0.15, 0, 0	52
16	Central-Limit (n = 12)	0.39	12	12						1
17	CLT-Stretched [1959]	0.35	12	17	8					5

In a comprehensive review by Thomas et. al. (Thomas2007), authors compared 17 mathematical GRNG algorithms in terms of computational efficiency and statistical properties. They used χ^2 goodness of fit and high sigma-multiple to test statistical properties of GRNG algorithms.

To determine the computational efficiency of each of the GRNG algorithms, the operation counts per generated random number and the relative speed were measured. The speed was measured relative to the Polar Rejection algorithm, as this algorithm is simple and commonly used as a baseline for performance. Table A.1 shows the computational efficiency of these 17 GRNG algorithms, while Table A.2 lists the statistical quality of generators. It appears that the Ziggurat algorithm has the best statistical properties (Table A.2) and is the second fastest algorithm (Table A.1), rendering the algorithm the best available GRNG candidate for stochastic simulations and scientific simulations of random processes.

Table A.2 Statistical quality of different GRNG algorithms as measured by χ^2 test and high sigma test in single precision. The "+" sign indicates GRNGs passing the χ^2 test for more than 2^{36} samples. The high sigma test standard shows the highest score after which the test become computationally infeasible for the generator. Adopted from (Thomas2007).

GRNG Algorithm [published year]	χ^2 Test ($\log_2(n)$)		High Sigma Test	
	Standard	Full-Fraction	Standard	Full-Fraction
Ziggurat [2000]	+	+	**8.15**	**17.4**
Wallace (qual = 1) [1996]	+	n/a	6+	n/a
Wallace (qual = 4) [1996]	+	n/a	n/a	n/a
Monty Python [1998]	34	n/a	8.27	14.88

PPND7 (ICDF) [1988]	34	34	4.11	12.64
Ahrens-Dieter [1988]	15	+	17.3	17.3
Mixture-of-Triangles [2000]	26	n/a	n/a	n/a
Polar [1969]	36	+	8.09	11.59
GRAND [1974]	36	+	9.2	17+
Hastings (ICDF) [1959]	29	30	5.25	12.64
Leva (Ratio) [1992b]	+	+	7.91	17+
PPND16 (ICDF) [1988]	35	+	4.11	13.7
Marsagla-Bray [1964]	35	+	8.35	15.78
Box-Muller [1958b]	26	35	5.57	13.96
Kinderman (Ratio) [1977]	+	+	7.91	17+
Central-Limit (n = 12)	20	n/a	0.99	n/a
CLT-Stretched [1959]	28	n/a	2.84	n/a

A.2 Random seed

Importantly, in every RNG the sequence of random numbers is dictated by a seed (starting point; usually an integer), as this is a purely deterministic procedure (Park1988) (this, incidentally, explains why the generated numbers are not literally random, but pseudorandom). Thus, it is very important to start with a different seed for every simulation, ensuring disparity between sequences of random numbers per simulation. Different languages and also Linux have intrinsic functions producing a pseudorandom seed usually based on the current (CPU) time or frequency. While these functions are currently widely used at the beginning of simulations to feed the RNG, this approach is prone to flaws. A particular case is when the High Throughput Computing is desired. For example, if, for statistical analysis, one needs to execute a stochastic simulation with the same initial configuration many times simultaneously and *independently*, each copy of the simulation must have its unique random seed. If all of these simulations are executed at the same time, which is often the case in supercomputers with distributed memory model, it is very likely that all copies are assigned the same seed. This would cause all simulations producing exactly the same results that are statistically fruitless.

Perhaps one of the best and the safest way to feed the RNG with initial seed is to use true random numbers from reliable resource. An excellent online resource is the random.org that generates true integer random numbers from atmospheric noise (inclusive limits of ±1,000,000,000). It can be invoked by a simple Python script at beginning of the first set of simulations. The last random number generated in each simulation then can be used as the seed in the following simulation.

The rest of the current chapter is devoted to development of coarse-grained biophysical modeling of the NPC.

Appendix B. The prediction of the grafting locations and disordered domains of FG-Nups in the human NPC.

It is known that mass and dimensions of NPC are species-dependent (Peters2006, van der Aa2006, Maimon2012, Bui2013, Eibauer2015, Kosinski2016). The species-dependent mass of FG-Nups implies different lengths and compositions in FG-repeat domains than those in the yeast NPC. Unfortunately, however, there is not such a comprehensive data on the locations and disordered domains of FG-Nups in other species as there is for the yeast NPC. Thus, here by performing a comparative analysis on available data about the FG-Nups in the yeast, vertebrate, and human NPCs, I predicted the location of each FG-Nups across the central channel. The approximated locations of different Nups in a vertebrate NPC are superimposed on the scaffold as shown in the Fig. B1 (Rabut2004). To draw functional analogy between the FG-Nups in hNPC and the yNPC, the comparative properties of FG-Nups in human vs yeast were used (Tables B1 and 3.5) (Xu2013). This yielded the approximated locations of each FG-Nup in the human NPC as shown in the Fig. B2.

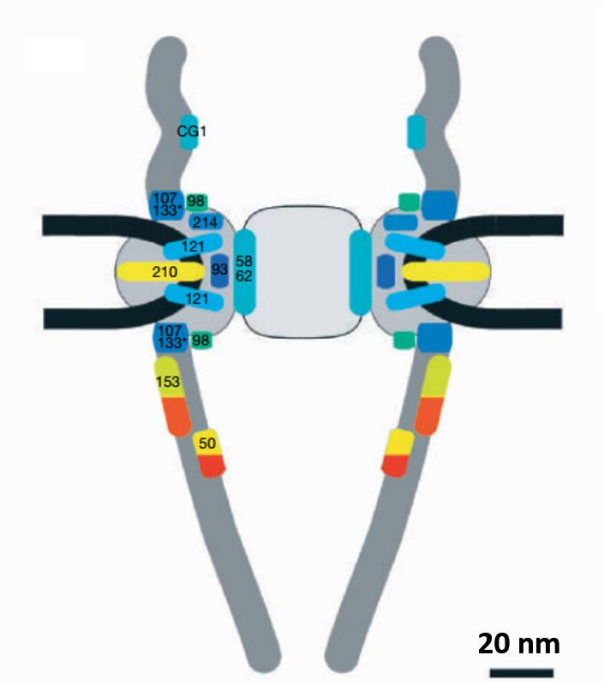


Figure B1. The approximate locations of Nups, including FG-Nups in a vertebrate NPC (Rabut2004).

Table B1. The propoerties of FG-Nups in the human NPC. Adopted from (Xu2013).

Properties of Human Nucleoporin FG Repeat domains

Nucleoporin:	Nup98	Nup214	Nup153	Nup62	Nup58	Nup54	Nup50	CG1	POM121
Repeat domain (AAs)	1-469	1632-2090	900-1419	1-143	1-69, 472-579	1-88	73-304	215-365	569-953
Total AA	468	458	520	143	177	88	232	150	385
MW (kDa)	46.04	43.15	52.85	13.7	16.41	8.01	24.09	14.26	36.72
Total (-) chgs	*7	3	34	0	0	0	20	3	3
Total (+) chgs	*16	8	32	2	0	1	21	5	6
pI	9.86	10.02	5.57	10.00	7.98	8.50	7.86	9.56	9.54
% charged (DERK)	4.9 (1.7)	2.3	12.7	1.4	0.6	1.1	17.7	5.4	2.4
% hydrophobic (AILVMFW)	26.2	29.8	27.1	32.9	27.1	30.6	30.2	35.4	35.3
Total FG[1]	38	37	24	6	14	8	5	9	23
(H)FG[2]	14	13	3	0	1	1	2	0	5
(H)xFG[3]	4	3	15	5	6	3	2	2	7
Pro-cohesive repeats[4]	18	14	5	5	7	4	0	2	10

*All (-) chgs and 8/16 (+) chgs are found within the Gle2/Rae1 binding site (GBD) of Nup98. With the GBD omitted from the calculations, charged residues = 1.7%.

[1]FGFGs were scored as 2 x FG rather than 1 FxFG.

[2](H)FG indicates L, V, I, M,or F immediately preceding FG, as in GLFG.

[3](H)xFG indicates L, V, I, M,or F at -2 position relative to FG, as in FxFG.

[4]Pro-cohesive repeats: 1) are found within a region of low charge density spacers, and 2) have sequence motifs (H)FG or (H)xFG.

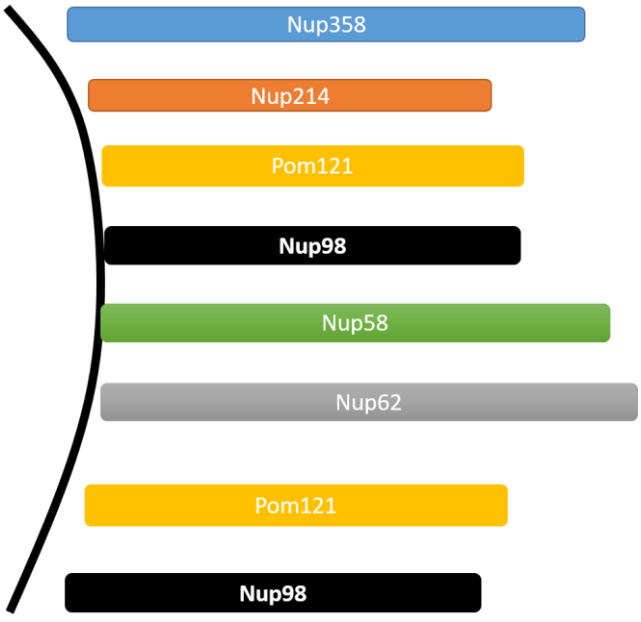


Figure B2. The approximated locations of FG-Nups (calculated in the current dissertation) in the human NPC based on their functional analogy with FG-Nups in the yeast NPC.

Next unknows are the disordered domains of FG-Nups in the hNPC. As there is no data available on the disordered domains, here I used the Predictor of Natural Disordered Regions (PONDR) (Romero2001) available at http://www.pondr.com/ to predict the disordered domains of FG-Nups in the human NPC. To do so, first, sequences of human FG-Nups were taken from the Unitprot.

Initially, only VL3 algorithm (Obradovic2003) was used to predict the the disordered domains of each FG-Nup in the hNPC (Fig. B3). Next, the results were compared with two other algorithms, P-FIT and VSL2B (Obradovic2005) to evaluate the confidence level (Fig. B4).

The disordered domains were then selected in accordance with the prediction of two algorithms VL3 and VSL2B (Obradovic2003, Obradovic2005). Two algorithms are in good agreement with each other.

The beginnings and start points of the sequences in disordered domains are shown in Table B2.

Table B2. Beginning and end of the disordered regions in the FG-Nups predicted based on PONDR

Nup	Beginning of the disordered region	End of the disordered region	Total number of residues in the disordered region.
Nup98	1	1100	1100
Pom121	300	1249	949
Nup62	1	300	300
Nup58	1	599	599
Nup358	600	3224	2624
Nup214	400	2090	1690

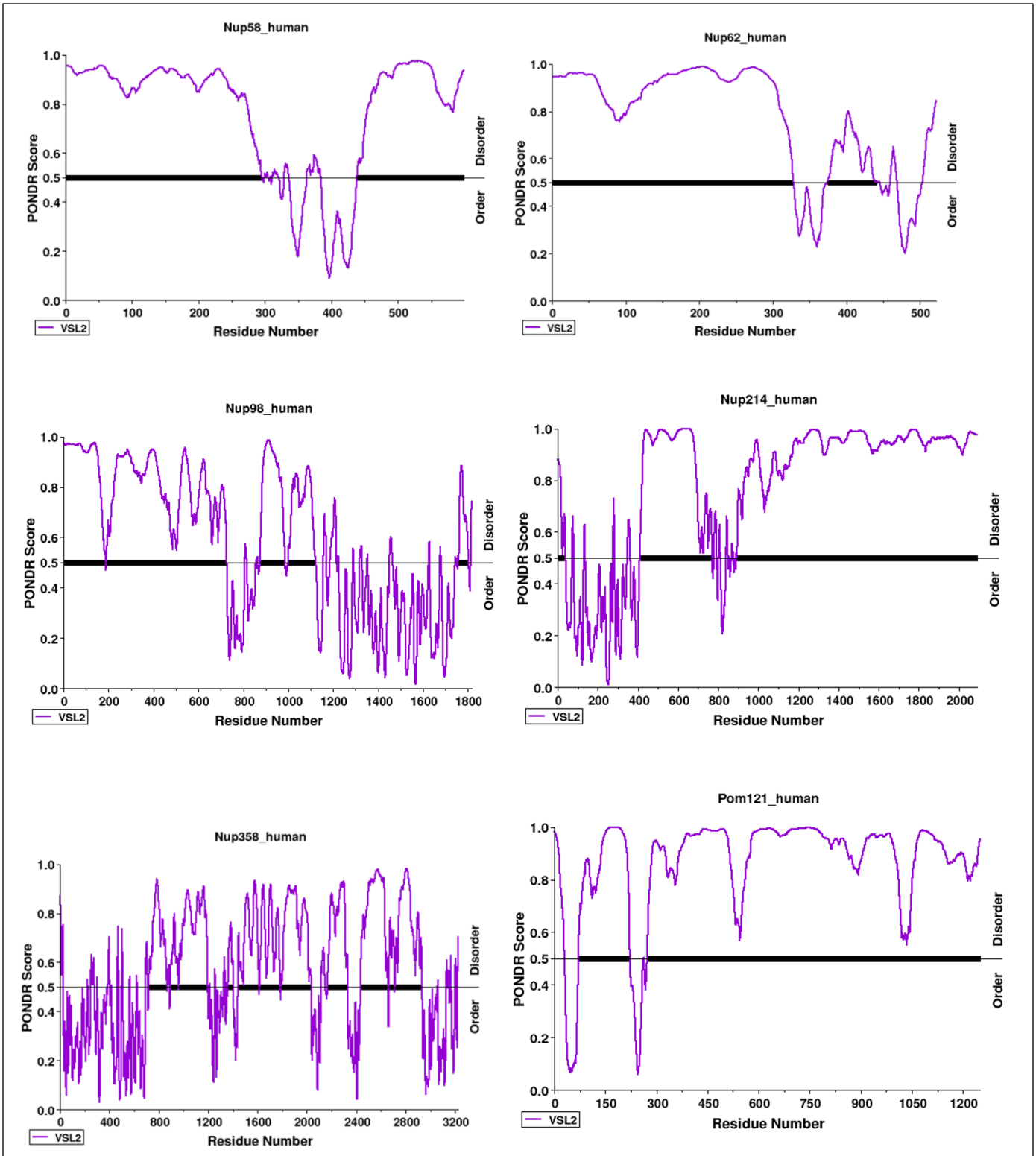


Figure B3. The predictions of PONDR for the disordered domains of FG-Nups in the human NPC.

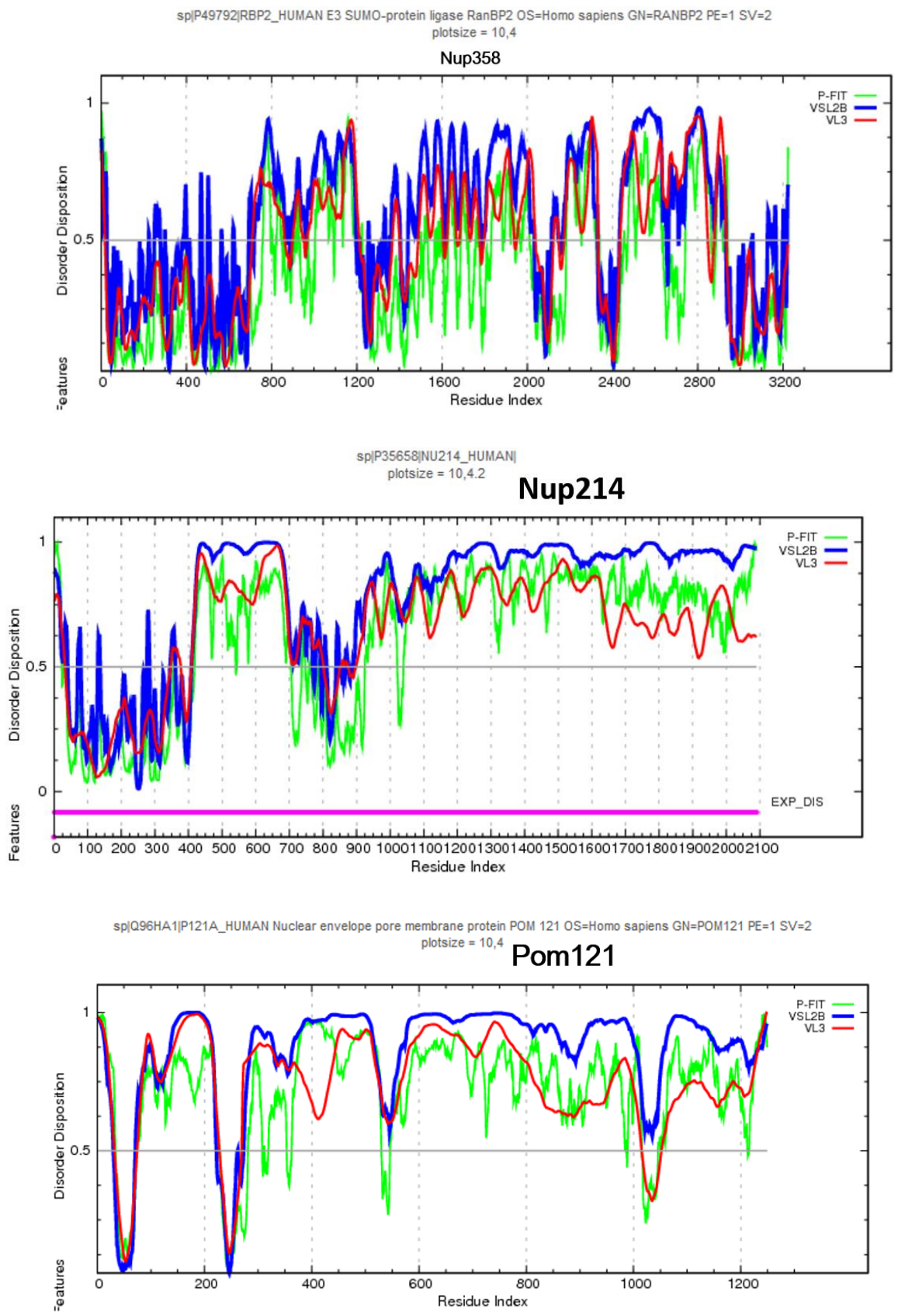


Figure B4 The comparative results of two algorithms VSL2B and VL3 along with P-FIT values for the prediction of disordered regions in FG-repeat domains in the hNPC.

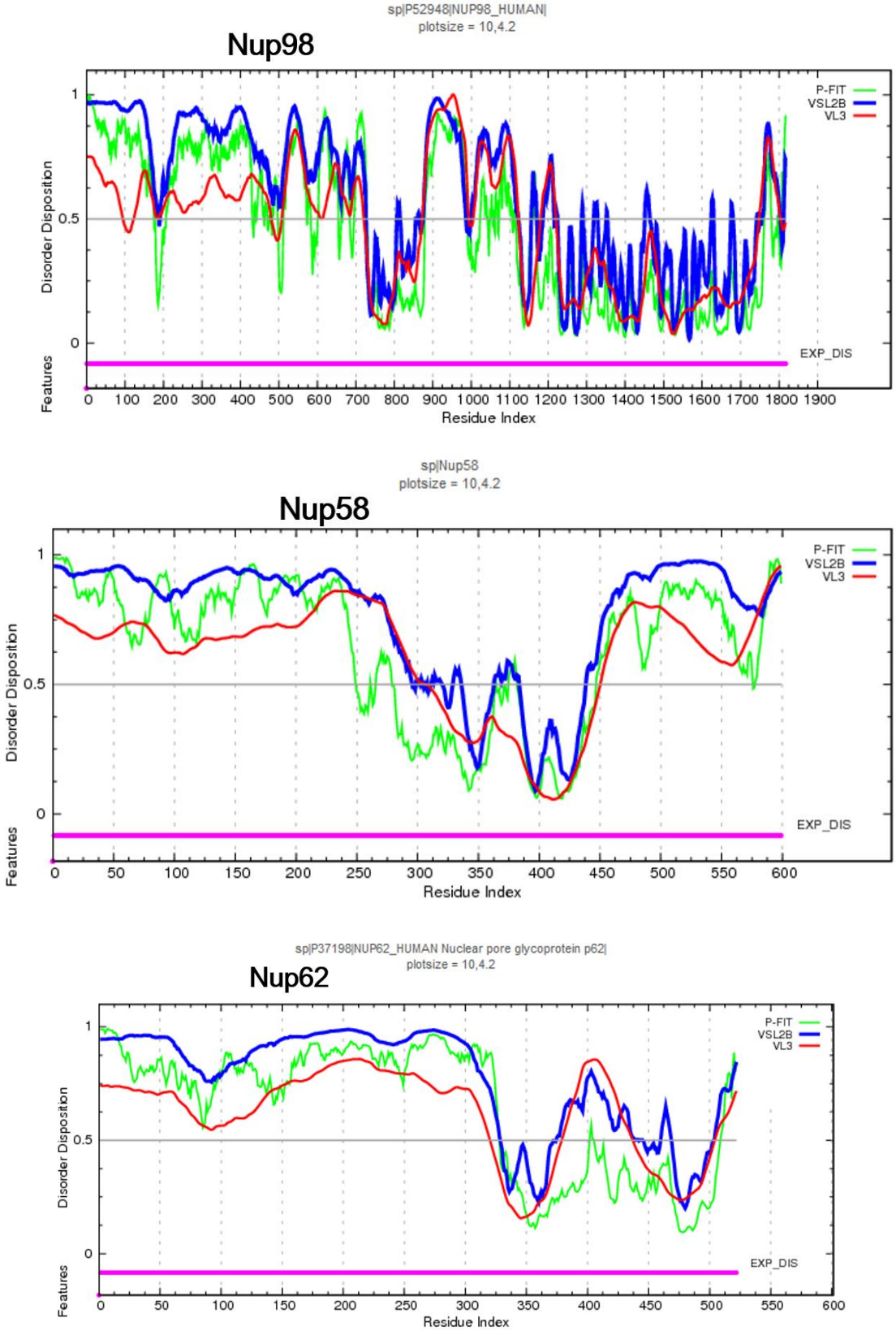


Figure B4 Continue

الحمدلله رب العالمين