

**Mechanisms of diffusion in nuclear pore complex mimics**

by

**Laura K. Maguire**

B.S., Harvey Mudd College, 2013

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Physics

2019

This thesis entitled:  
Mechanisms of diffusion in nuclear pore complex mimics  
written by Laura K. Maguire  
has been approved for the Department of Physics

---

Assistant Professor Loren Hough

---

Professor Meredith Betterton

Date \_\_\_\_\_

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Maguire, Laura K. (Ph.D., Physics)

Mechanisms of diffusion in nuclear pore complex mimics

Thesis directed by Assistant Professor Loren Hough

Placeholder text for abstract.

## Dedication

Nice dedication goes here.

## Acknowledgements

There are lots of people to acknowledge.

## Contents

### Chapter

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Intrinsically disordered proteins are essential to cellular function . . . . .	1
1.2	Major components of nucleocytoplasmic transport . . . . .	2
1.2.1	Scaffold nucleoporins are ordered and form ringlike complexes. . . . .	3
1.2.2	FG nucleoporins are disordered and fill the central channel of the pore. . . .	3
1.2.3	Energy for selective transport is provided by the Ran cycle. . . . .	4
1.2.4	Transport factors . . . . .	5
1.3	Observed features of nucleocytoplasmic transport . . . . .	6
1.3.1	Flux through the NPC is very high. . . . .	6
1.3.2	TFs show millisecond dwell times. . . . .	7
1.3.3	TF-FG Nup binding kinetics are ultrafast. . . . .	7
1.3.4	Passive permeability barrier . . . . .	7
1.3.5	Effect of transport factors . . . . .	8
1.3.6	Permeability barrier/flux studies . . . . .	8
1.4	Theoretical models of the NPC . . . . .	8
1.4.1	Entropic barrier model . . . . .	8
1.4.2	Hydrogel model . . . . .	9
1.4.3	“Forest”/intermediate models . . . . .	9
1.4.4	Effect of crowding . . . . .	9

1.4.5	Effect of transport factors . . . . .	9
1.4.6	Energy landscape models . . . . .	9
1.5	Synthetic NPCs have seen limited success. . . . .	10
1.5.1	Gold nanopores (Rout group) . . . . .	10
1.5.2	Hydrogels (Gorlich) . . . . .	10
1.5.3	Other models (peptide hydrogels, DNA origami, etc) . . . . .	10
1.6	Conclusions of introduction? . . . . .	11
<b>2</b>	<b>Modeling</b>	<b>12</b>
2.1	Biophysical model of transport through the NPC . . . . .	13
2.1.1	No selective transport occurs if bound TFs are immobile . . . . .	15
2.1.2	Bound-state diffusion allows selective transport . . . . .	16
2.2	Mechanisms of bound transport factor mobility . . . . .	17
2.2.1	FG Nup flexibility allows tethered diffusion . . . . .	18
2.2.2	Inter-chain hopping increases selectivity . . . . .	20
2.2.3	Detailed explanation of hopping simulation . . . . .	21
2.2.4	Hopping simulation results . . . . .	23
2.2.5	Fickian diffusion . . . . .	23
2.3	Discussion . . . . .	24
2.3.1	Overcoming the limitations of binding . . . . .	25
2.3.2	Design principles of selective transport by binding . . . . .	26
<b>3</b>	<b>Hydrogels yay</b>	<b>27</b>
3.1	Pore size . . . . .	27
3.2	Porogens . . . . .	27
3.2.1	Alginate nanospheres . . . . .	27
3.2.2	Dextran/dextranase system . . . . .	27
3.3	Polymerization using confocal microscope . . . . .	27

3.4	Nup and transport factor constructs . . . . .	28
3.5	Dye-labeling and free dye . . . . .	28
<b>4</b>	<b>Actual real results</b>	<b>29</b>
<b>5</b>	<b>FG Nup aggregation under crowded conditions</b>	<b>30</b>
5.1	Results . . . . .	30
5.1.1	Thioflavin timecourses . . . . .	30
5.1.2	Fluorimetry . . . . .	35
5.2	Discussion . . . . .	37
5.3	Materials and Methods . . . . .	37
5.3.1	Buffers . . . . .	37
5.3.2	FG124 preparation . . . . .	37
5.3.3	Timecourse preparation . . . . .	38
5.3.4	NMR sample preparation . . . . .	39
5.3.5	NMR experiments . . . . .	39
5.3.6	Fluorimetry . . . . .	39

## Appendix



## Tables

### Table

2.1	Comparison between experimental results for NTF2 and GFP (a similarly-sized non-binding protein) and model predictions. Flux measured in units of molecules per pore per second. . . . .	17
5.1	FG124 aggregation lifetime and lag time with varying pH. Each condition run with 6 replicates in PTB buffer. One-way ANOVAs show no statistically significant differences between conditions. Standard errors are shown. . . . .	33

## Figures

### Figure

- 2.1 Schematics of the nuclear-pore complex and model. (A) The nuclear pore complex (gray) is filled with FG Nups (green polymers) that selectively passage transport factors that bind to FG Nups (blue) while blocking non-binding proteins (red). The central channel of the pore has length  $L$ . Protein concentration is high on the left (inlet) and low on the right (outlet). (B) Selectivity quantifies the degree of selective transport through the pore. A non-selective pore with  $S = 1$  has the same flux for a transport factor as for a non-binding protein (top). A selective pore with  $S > 1$  has a larger flux for a transport factor than a non-binding protein (lower). (C) The bound diffusion coefficient quantifies the mobility of a bound transport factor. A transport factor may be immobile (top) or mobile (lower) when bound. . . . . 12
- 2.2 Flux through the pore and selectivity for TFs with varying bound mobility. (A) Flux as a function of time when TFs are immobile while bound, with varying binding affinity as in (B). (B) Flux as a function of time when TFs are mobile while bound with  $D_B = D_F$ , with varying binding affinity. (C) Selectivity as a function of dissociation constant with varying bound diffusion coefficient. . . . . 14

2.3	(A) Schematic of the flexible tether model of bound-state diffusion. FG Nups are treated as entropic springs that constrain the motion of TFs more (top and center left, longer FG Nup) or less (top and center right, shorter Nup), which corresponds to changing width of the harmonic potential well (lower). (B) Ratio of bound to free diffusion coefficient as a function of dissociation constant, with varying polymer length in the tethered-diffusion model. (C) Selectivity as a function of $K_D$ , with varying polymer length in the tethered-diffusion model. . . . .	16
2.4	(A) Schematic of the inter-chain hopping model of bound-state diffusion. FG Nups are treated as entropic springs that constrain the motion of TFs, and inter-chain hopping allows a TF to move from one FG Nup (top and center left, green Nup) to another (top and center right, red Nup) without unbinding, which corresponds to switching from one harmonic potential well to another (lower). (B) Ratio of bound to free diffusion coefficient as a function of dissociation constant, with varying hopping rate in the inter-chain hopping model. (C) Selectivity as a function of $K_D$ with varying hopping rate. FG Nup contour length $L_c = 40$ nm in (B, C). . . . .	20
2.5	(A) Examples of mean-squared displacement (MSD) of a simulated TF in the inter-chain hopping model, with varying hopping rate. (B) Examples of MSD distributions $\rho_{\text{MSD}}(t)$ used in estimating the diffusion coefficient, with varying unbinding rate. Tethers have 40 nm contour length; other parameters are as discussed in the text. .	23
2.6	Bound diffusion and selectivity as a function of dissociation constant, with varying hopping rate for FG Nups with $L_c = 4$ nm. . . . .	24
2.7	Bound diffusion and selectivity as a function of dissociation constant, with varying hopping rate for FG Nups with $L_c = 12$ nm. . . . .	25
2.8	Bound diffusion and selectivity as a function of dissociation constant, with varying hopping rate for FG Nups with $L_c = 120$ nm. . . . .	26
5.1	test . . . . .	32

5.2	test . . . . .	34
5.3	Emission scan of fresh and aggregated FG124 in crowded conditions. Data normalized by subtracting blank sample. . . . .	35
5.4	Emission scan of phenylalanine, FSFG, and FG124. Excited at 240 nm. Phe data is not mine, need to check reference. Data normalized to a maximum amplitude of 1. .	36
5.5	Emission scan of phenylalanine, FSFG, and FG124. Excited at 240 nm. Phe data is not mine, need to check reference. Data normalized to a maximum amplitude of 1. .	36
5.6	Absorbance of 5% PEG and PVP solutions in PTB. Normalized by subtracting PTB absorbance. . . . .	37

## Chapter 1

### Introduction

#### 1.1 Intrinsically disordered proteins are essential to cellular function

For decades, it was conventional wisdom among biologists that a protein's folded shape determined its function. Most enzymes and other proteins that were studied had a stable folded configuration, the lowest point on a well-defined folding energy landscape. A protein's conformation provided specific docking points through which it could interact with ligands or other proteins in a "lock-and-key" model.

However, a few decades ago, it began to become clear that not all proteins have a well-defined tertiary or even secondary structure, but rather exist as extended polymer chains. These intrinsically disordered proteins (IDPs) were initially dismissed as nonfunctional, but evidence began to accumulate that they were in fact essential for cellular function, overturning the structure-function paradigm [?]. Their roles and importance are still being understood, as are the unusual mechanisms by which they accomplish their functions without a well-defined structure.

Today, it is estimated that 30% of eukaryotic proteins are disordered or contain significant disordered regions [?]. While there is significant sequence heterogeneity among IDPs, they tend to contain a large proportion of hydrophilic residues, and often have long stretches of low-complexity regions where only a few amino acids are represented. They also often have high net charge.

Some IDPs fold (or partially fold) upon binding with an ordered partner, while others form

“fuzzy” complex that remains disordered. Their advantages over folded proteins may include their plasticity, which enables them to bind many different binding partners. Multivalency, either as one-to-many or many-to-one binding, may also play a role. They may act as hubs that bring together larger complexes. Similarly, IDPs are often known for having high specificity at relatively weak binding strengths [?].

While the normal functioning of IDPs is very important to the cell, IDPs are also prone to aggregation and are at the root of pathologies such as Alzheimer’s disease, Parkinsons, and prion diseases. Often, normally-disordered proteins aggregate into amyloid fibrils, a stable structure based on parallel beta-sheets.

IDPs are commonly involved in cell signaling and regulation [?]. Their disordered nature makes them useful as hubs that bring together many other proteins, and as scaffolds that many proteins can bind to at once. IDPs appear to be prevalent in transcriptional regulation, and they are playing increasingly apparent roles in liquid-liquid phase separation within cells. One of the most fascinating examples of IDP function is in the nuclear pore complex (NPC), a unique selective barrier that regulates all transport between the nucleus and the cytoplasm. The link between disorder and selectivity is not well understood in this case.

## 1.2 Major components of nucleocytoplasmic transport

The nuclear pore complex (NPC) resides in the nuclear envelope of eukaryotes and regulates all macromolecular traffic between the nucleus and cytoplasm. The NPC is one of the largest protein complexes in the cell, at about 60 MDa in yeast and 120 MDa in humans [?]. As the regulator of nucleocytoplasmic transport, the NPC must rapidly and specifically pass a wide array of macromolecules: transcription factors into the nucleus, and RNA into the cytoplasm. Moreover, it must be robust to problems and able to accomodate mechanical strain as the nuclear envelope changes shape, and to passage large cargo.

A typical yeast cell has xx nuclear pores, each with a dimension of xx. Human cells have about xx pores with xx dimensions.

There are about 30 different types of Nups, all present in multiple copy numbers.

These functions are accomplished through a structure with two main parts, both made of proteins known as nucleoporins, or Nups: the scaffold Nups, which form a ringlike complex, and the FG Nups, which are disordered and fill the central channel created by the scaffold Nups. Aside from the NPC itself, transport factors (TFs) are a class of proteins essential for selective transport. The energetic cost of selectivity is captured in the Ran GTP/GDP cycle.

### **1.2.1 Scaffold nucleoporins are ordered and form ringlike complexes.**

The nuclear pore itself is formed of scaffold Nups, which are ordered proteins that form ringlike complexes with eightfold symmetry [?]. There is an inner ring and two outer rings, the nuclear and cytoplasmic rings. The outer rings are slightly larger. The nuclear ring is on the side of the nucleus and includes the nuclear basket. The cytoplasmic ring includes the cytoplasmic filaments, which are (probably?) disordered proteins extending out into the cytoplasm.

The eightfold symmetry of the pore arises from its modular nature. Scaffold Nups form various stable subcomplexes, of which one of the most important is the Y-complex. The Y-complex forms the inner ring; there are 32 copies of the complex per pore [?]. The rings themselves are relatively flexible, as they need to be in order to accomodate deformations of the nuclear envelope. This flexibility is achieved in part by through short linear motifs (SLMs) which connect the subcomplexes to each other.

Recent cryo-EM studies have achieved unprecedented resolution of the scaffold Nups [?, ?].

### **1.2.2 FG nucleoporins are disordered and fill the central channel of the pore.**

The central channel of the pore is filled with disordered FG nucleoporins (FG Nups). FG Nups typically consist of an ordered domain that anchors them to the wall of the channel, and an entirely disordered domain that extends into the channel. As with all Nups, FG Nups have eightfold symmetry in the pore, and some of them are present in much higher copy number.

The disordered portion of every FG Nup contains phenylalanine- glycine (FG) motifs which

bind to the hydrophobic binding pockets of transport factors. While there are multiple binding motifs, all are short sequences which incorporate an FG repeat; for instance, FSFG, GLFG, and others. [?]. Each FG Nup contains x-x FG repeats, leading to an extremely high density of FG repeats within the pore.

Since the FG Nups are disordered, most conventional visualization techniques do not work. When imaged over time or when several pores are imaged, the averaged results do not show the disordered portion of the FG Nups. Cryo-EM and x-ray crystallography don't work. Techniques such as NMR and very fast AFM can help gain insight into their conformational ensembles [?, ?]. Early research suggested that the FG Nups formed a central plug or "transporter", but more recent work suggests that there is no central structure, just disordered proteins (the AFM study from Lim or Lemke group). There is some evidence from simulations that the density of the FG Nups, as well as their charge density and hydrophobic properties, are not uniform along either the radial or axial directions [?, ?]. This may contribute to selective transport, although the pore still functioned with all of the asymmetric FG Nups removed. Indeed, the NPC is remarkably robust to FG Nup deletion. Over half of the mass of FG Nups can be removed without eliminating the selectivity barrier [?].

### **1.2.3 Energy for selective transport is provided by the Ran cycle.**

Selective transport requires an energy source, which in the case of the NPC is provided by the Ran cycle. When a TF-cargo complex passes from the cytoplasm into the nucleus (nuclear import), it encounters a RanGTP on the nuclear side which binds to the TF and displaces the cargo. Then the TF-RanGTP complex can collect a cargo destined for nuclear export, and this ternary complex can diffuse back through the NPC to the cytoplasm. The protein RanGAP then hydrolyzes the RanGTP to RanGDP, disrupting the complex into its three original pieces. Ultimately, the energy source for selective nuclear transport comes from the RanGTP-RanGDP gradient from the cytoplasm to the nucleus, a gradient which is maintained partially by NTF2, which carries RanGDP through the pore [?].



From the perspective of transport, this means that the process of passing through the pore is itself passive and does not consume energy. The selectivity arises from concentration gradients maintained by the Ran cycle.

#### 1.2.4 Transport factors

Transport factors (TFs) are ordered proteins that carry cargo through the NPC. While there are various types, they share several features in common, most notably the fact that all known transport factors have more than one hydrophobic binding pocket which binds to FG repeats. In fact, many TFs have several binding pockets. Likewise, the binding affinity between TFs and FG Nups remains unknown for most TFs. Estimates of dissociation constant  $K_D$  vary from nanomolar to millimolar, depending on the environment (cellular, buffer, etc.) in which the measurement is made [?]. There are many types of TF, of which some of the most important are the importins and exportins (karyopherins), NTF2, (and mRNA exporters? CRM? mex67?).

The karyopherins (Kaps) are the most-studied family of TFs. They are also known (in human cells?) as importins and exportins, or collectively as the importin  $\beta$  superfamily [?]. The twenty or so different Kaps are responsible for most nucleocytoplasmic transport [?]. Kaps typically consist of multiple HEAT repeats, a helical motif which conveys structural flexibility [?]. Most Kaps bind their cargo directly via a nuclear localization signal (NLS, for nuclear import) or nuclear export signal (NES, for nuclear export). NLS and NES are 5-7 amino acid tags found on cargo []. However, Kap95? (importin  $\beta$ ) uses the adaptor protein Kap60? (importin  $\alpha$ ) to bind its cargo. In general, Kaps are on the order of 100 kDa in size, well above the passive permeability limit []. Kaps may contribute to the selectivity barrier.

Unlike the karyopherins, nuclear transport factor 2 (NTF2) does not transport a wide variety of cargo across the NPC. Instead, NTF2 is focused on maintaining the Ran gradient needed for transport. It transports RanGDP across the pore - why does this help maintain a gradient? If it transports in both directions, wouldn't it help wash out the gradient? NTF2 is a homodimer whose components are 14 kDa and have one FG binding site. Although its small size of 28 kDa is near

the 30 kDa cutoff for passive transit through the pore, its flux through the pore is still 30-150 that of similarly-sized proteins that do not bind to FG Nups.

There are other TFs such as Mex67 and CRM, but I don't really know what they do. Is RNA exported using Kaps or are there other TFs for that?

### **1.3 Observed features of nucleocytoplasmic transport**

The nuclear pore has been the subject of investigation since the 1950s [?]. During that time, many different experimental techniques have been used, from in vivo studies to single molecule fluorescence to SPR to NMR. Likewise, simulations have attempted to explain the selectivity properties of the NPC using a broad array of techniques, such as more or less coarse-grained molecular dynamics simulations. These studies together shed light on the mechanisms of selective transport, but are also all hindered by the small size and rapid timescale of transport. Many of the studies produce results that seem, at least at first glance, to contradict those of other studies. A comprehensive picture of the facts of nuclear transport has yet to emerge, although there are some universally accepted features.

#### **1.3.1 Flux through the NPC is very high.**

The high macromolecular traffic between nucleus and cytoplasm requires high flux through each NPC. Experiments with permeabilized cells estimate that the total molecular flow through the NPC could be as high as 10-20 MDa per pore per second, corresponding to roughly 1000 transport events per pore per second [?]. Experiments focusing particularly on NTF2 report fluxes between 50 and 250 molecules per pore per second [?, ?, ?]. Fluxes this high mean a continuously high occupancy of the NPC, estimated at around 100 karyopherins at once [?]. It is unclear whether the effect of crowding with TFs in the pore has an impact on selectivity.

### 1.3.2 TFs show millisecond dwell times.

One reason that individual NPCs can accomodate such high flux is the rapidity with which molecules transit the pore. A wide range of molecules, such as NTF2, Importin  $\beta$ , and GFP-NLS cargoes, have a dwell time of less than 10 ms in the pore [?, ?, ?, ?]. Typically, this is determined using single-molecule tracking with superresolution microscopy [?]. There is some evidence that a subpopulation of karyopherins resides in the pore over long timescales, potentially contributing to the selectivity barrier [?].

The flux through an NPC is determined not only by the transit time, but also by the success rate of transit attempts. Single-molecule microscopy suggests that the nuclear import efficiency of Importin  $\beta$  ranges from 50% to 80%, depending on concentration [?]. Modeling supports these numbers [?].

### 1.3.3 TF-FG Nup binding kinetics are ultrafast.

The short dwell time of transport factors and cargo at the pore contributes to a source of confusion in the NPC community: the binding kinetics of TFs and FG Nups. Measurements of dissociation constant  $K_D$  and its components  $k_{\text{on}}$  and  $k_{\text{off}}$  ( $K_D = k_{\text{off}}/k_{\text{on}}$ ) span several orders of magnitude, depending on the measurement conditions. Original in vitro measurements, performed in buffer, show low  $K_D$  values in the nanomolar range, indicating tight binding between TFs and FG Nups. However, these values were not compatible with a transit time around 10 ms, as a typical binding event as calculated from the off-rate would last much longer than 10 ms.

More recent measurements are converging on a consensus that TF-FG Nup binding kinetics are in fact much faster than previously estimated. Diffusion-limited on-rates and  $K_D$  values around 1 mM would permit rapid transit through the pore.

### 1.3.4 Passive permeability barrier

There is a cutoff around 30 kDa where molecules stop being able to passively transit the pore at any appreciable rate. Work from the Timney lab suggests that the passive permeability barrier

is not a sharp cutoff, but broad.

Dwell times for various TFs and TF-cargo complexes has been studied in many cases, giving a typical dwell time of 5-10 ms, with a small minority of complexes spending much longer in the pore [?].

### **1.3.5 Effect of transport factors**

The presence of transport factors may make the selectivity barrier more robust. SPR measurements as discussed above give estimates of layer compaction or extension, and populations of tight- or weak-binding TFs. Other studies (?) suggest that crowding with TFs might help reduce non-specific interactions and increase the selectivity of transport.

### **1.3.6 Permeability barrier/flux studies**

People have measured the flux through the pore in vitro and in vivo in various ways. One group developed OSTR, where they seal an NPC to the pore of a membrane and measure fluorescence flux of TFs through that pore. The Gorlich group (Ribbeck) have permeabilized cells and injected fluorescent transport factors and watched how long it took them to localize to the nucleus. These measurements show selectivities of 20-150 (ish) fold flux of TFs as compared to inert proteins.

## **1.4 Theoretical models of the NPC**

Many theoretical models exist, both qualitative and quantitative.

### **1.4.1 Entropic barrier model**

The entropic barrier model postulates that Nups remain disordered within the pore (should talk about experimental evidence for and against). In this model, inert proteins are kept out of the pore through an entropic barrier, because their entry into the pore would restrict the possible conformations of the disordered Nups. TFs can get into the pore, in contrast, because the binding

energy offsets the entropic penalty. This is one of the two main qualitative models of the pore, the other being the hydrogel or selective phase model. The Rout lab supports this model. I think most labs other than the Gorlich lab are beginning to support this model, but I'm not sure.

#### **1.4.2 Hydrogel model**

The hydrogel model is the other main qualitative NPC model. The Gorlich group supports this model. This model postulates that the Nups interact via their FG motifs (discuss evidence in previous section, maybe?) and form a hydrogel. A dynamic hydrogel. Inert proteins are kept from passing through because they can't get through the gel mesh, but transport factors disrupt the crosslinks by binding to FGs and "melt" through.

#### **1.4.3 "Forest"/intermediate models**

This model proposes that a mix of the entropic barrier model and selective phase model are at play.

#### **1.4.4 Effect of crowding**

Some people (Zilman paper, etc) have modeled the effect of crowding on the selectivity barrier. A lot of overlap with effect of transport factors.

#### **1.4.5 Effect of transport factors**

Same as previous section sort of. Kap-centric models of the NPC (short for karyopherin-centric, for the class of TFs called karyopherins) propose that a permanent population of Kaps lives within the NPC. This strengthens the selectivity barrier.

#### **1.4.6 Energy landscape models**

Most quantitative models of the nuclear pore are energy landscape models. These are typically computational studies that require a fair amount of detail and assumptions about the pore. They

incorporate effects of charge, hydrophobicity, specific binding interactions, etc. Generally the result is a picture of the free energy landscape encountered by TFs and by inert proteins as they travel along the axis of the pore.

## **1.5 Synthetic NPCs have seen limited success.**

Many different groups have attempted to make synthetic nuclear pore complexes, but they are generally not very successful.

### **1.5.1 Gold nanopores (Rout group)**

One of Loren's colleagues in New York grafted FG Nups onto a gold-coated nanopore and monitored flux through the pore. She saw low (less than 10-fold) selectivity. I'm not sure whether other nanopore-based approaches have been tried.

### **1.5.2 Hydrogels (Gorlich)**

The Gorlich group keeps making hydrogels out of Nups and testing the entry of various proteins. They take Nsp1 or fragments of it or other Nups that spontaneously form hydrogels in buffer, and let them form a gel. Then they introduce fluorescently-tagged TFs and inert proteins and monitor the progression of the fluorescent front into the gel. They see very high (100 or more) partition coefficients, indicating that the TFs really bind very strongly to the gel. They do not see rapid exit from the gel as would be required for rapid transport.

### **1.5.3 Other models (peptide hydrogels, DNA origami, etc)**

There are a grab-bag of other models that I need to learn more about. Some groups have done similar things to what we want to do, making hydrogels out of non-aggregating Nup peptides and a pentameric crosslinking domain. There was a group that made an NPC ring out of DNA origami and attached Nups at particular points, though I'm not sure how that can be used to test selectivity because I can't see how to anchor the rings into a membrane.

## 1.6 Conclusions of introduction?

NPCs are interesting and important.

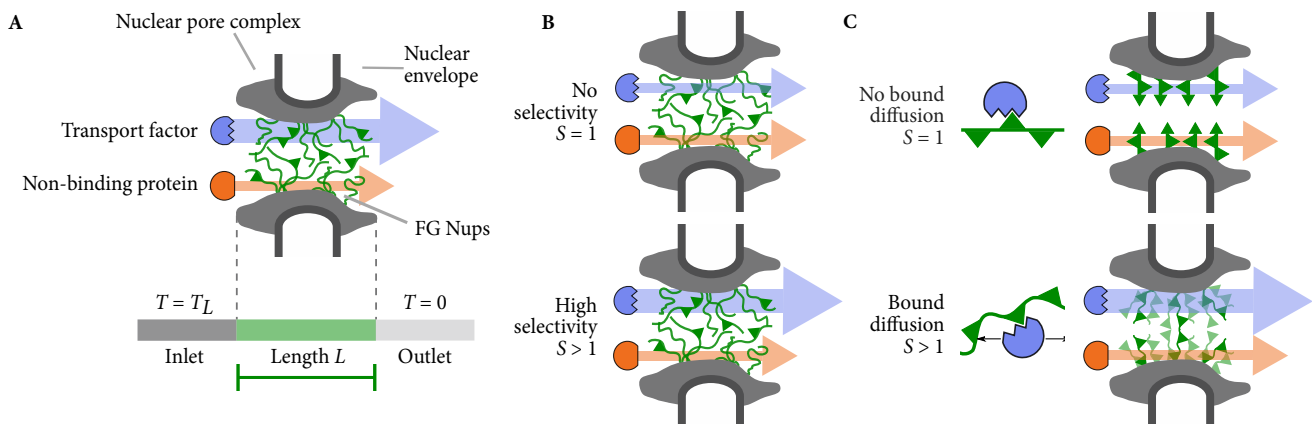


Figure 2.1: Schematics of the nuclear-pore complex and model. (A) The nuclear pore complex (gray) is filled with FG Nups (green polymers) that selectively passage transport factors that bind to FG Nups (blue) while blocking non-binding proteins (red). The central channel of the pore has length  $L$ . Protein concentration is high on the left (inlet) and low on the right (outlet). (B) Selectivity quantifies the degree of selective transport through the pore. A non-selective pore with  $S = 1$  has the same flux for a transport factor as for a non-binding protein (top). A selective pore with  $S > 1$  has a larger flux for a transport factor than a non-binding protein (lower). (C) The bound diffusion coefficient quantifies the mobility of a bound transport factor. A transport factor may be immobile (top) or mobile (lower) when bound.

## Chapter 2

### Modeling

Models of the NPC selective barrier have proposed that the FG Nups may form an entropic brush [?], a dynamic hydrogel [?,?], an intermediate state between a brush and gel [?], or liquid droplets [?]. These mechanisms may be modulated by spatial organization [?,?] and binding of TFs to multiple FG repeats [?,?]. Attempts to distinguish these models have been hindered by the pore's small size, the redundancy and multiple copies of FG Nups, and contradictory experimental results on FG Nups and TF binding [?]. Some FG Nup fragments form less-dynamic hydrogels *in*



*vitro* [?], but remain highly dynamic within cells [?]. Molecular dynamics simulations find highly dynamic FG Nups, though the degree and extent of motion depends on the affinity of FG repeats for each other and for TFs [?, ?]. Crowding and competition modulate affinity [?] and may contribute to selective transport [?]. However, the connection between the amino-acid level behavior of the FG-TF interaction and macroscopic transport selectivity remains unclear. Here we address the central contradiction of selective transport through the NPC: how does binding of TFs to FG Nups within the pore increase the flux rather than decreasing it [?, ?]?

Using a biophysical model, we demonstrate that TF diffusion and binding are sufficient for selective transport, as long as binding only partially immobilizes TFs. Binding increases the local concentration, and these molecules contribute to the flux if mobile. Thermally-driven diffusion of TFs bound to flexible tethers gives sufficient particle mobility to produce selectivity similar to experimental measurements. Tether flexibility also allows bound TFs to hop between tethers, further enhancing selectivity.

## 2.1 Biophysical model of transport through the NPC

We consider a minimal model of the central channel of the NPC containing FG Nups homogeneously anchored Figure 2.1. This model is sufficiently general to describe the common features of a range of biopolymer filters. The NPC, unlike most other biopolymer filters, has a wide capture area that may increase transport rates [?]. In order to focus on basic principles of transport, we neglect this effect. A varying free energy landscape along the axis of the NPC may play a role in selective transport [?, ?, ?, ?]. However, the NPC is robust to deletion of all asymmetric Nups and many Nup combinations, indicating that spatial variation in pore properties is not necessary [?, ?]. Experiments *in vitro* with simplified, homogeneous Nup composition produced selective transport [?, ?].

Rapid transport requires TF-FG Nup binding, while a protein similar to a TF but unable to bind FGs is excluded. Therefore, in our model we compare two proteins that are identical, except that one binds FG Nups and the other does not. As a model TF, we consider nuclear transport factor

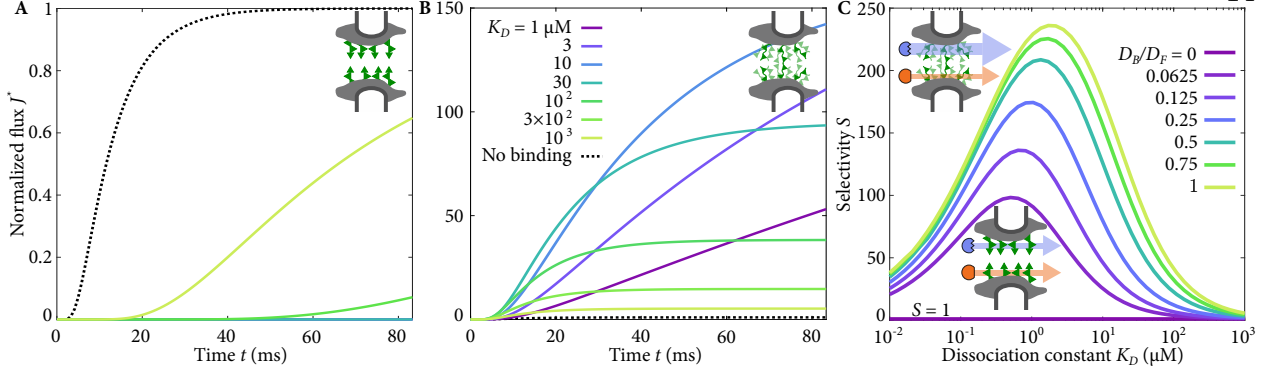


Figure 2.2: Flux through the pore and selectivity for TFs with varying bound mobility. (A) Flux as a function of time when TFs are immobile while bound, with varying binding affinity as in (B). (B) Flux as a function of time when TFs are mobile while bound with  $D_B = D_F$ , with varying binding affinity. (C) Selectivity as a function of dissociation constant with varying bound diffusion coefficient.

2 (NTF2) [?]. NTF2 is small ( $\sim 5$  nm) relative to the diameter ( $\sim 50$  nm) and length of the pore ( $\sim 100$  nm), suggesting that passage of NTF2 does not require large-scale molecular rearrangements that have been proposed for larger molecules [?, ?]. Because of the small size of NTF2 we neglect effects of steric crowding, which can enhance selectivity in a transport model [?]. NTF2 appears not to be actively released from the pore, suggesting that selective transport is an intrinsic property of the NPC [?, ?], and in contrast to actively released karyopherins [?, ?, ?, ?].

Transport through the NPC requires entry into the pore, passage, and exit. In single-molecule measurements, most of the transport time is spent in a random walk within the central channel [?, ?]. We therefore assume that entry and exit rates are determined by binding kinetics (see Supporting Information, section ?? for the model when entry and exit are rate-limiting.) The directional bias in TF transport is controlled outside the NPC through a concentration difference between the nucleus and cytoplasm generated by the Ran-GTP system [?]. In our model, we impose a fixed concentration difference across the pore.

We consider a channel of length  $L$  filled homogeneously with Nups that separates two reservoirs Figure 2.1A. Within the channel are free transport factor (concentration  $T$ ), free FG Nups ( $N$ ), and bound TF-FG complex ( $C$ ), with total Nup concentration  $N_t = N + C$ . TF diffusion

within the channel ( $0 < x < L$ ) is described by the reaction-diffusion equations

$$\frac{\partial T}{\partial t} = -k_{\text{on}}TN + k_{\text{off}}C + D_F \frac{\partial^2 T}{\partial x^2}, \quad (2.1)$$

$$\frac{\partial C}{\partial t} = k_{\text{on}}TN - k_{\text{off}}C + D_B \frac{\partial^2 C}{\partial x^2}. \quad (2.2)$$

TF-FG interaction has on-rate constant  $k_{\text{on}}$ , off-rate  $k_{\text{off}}$ , and dissociation constant  $K_D = k_{\text{off}}/k_{\text{on}}$ . We include competition between TFs for FG binding sites [?]. The diffusion constants of free ( $D_F$ ) and bound ( $D_B$ ) TFs are spatially constant. The fixed reservoir TF concentrations are  $T_L$  (inlet, left) and 0 (outlet, right).

The flux of transport factor out of the pore  $J = -D_F \partial T / \partial x|_{x=L}$ . We numerically integrated the full equations. Because flux measured in experiments is typically linearly proportional to TF concentration [?, ?], TF concentration likely remains below binding saturation in the NPC. Therefore, we also solved eqns. (2.1, 2.2) analytically in the low binding limit. We define the transport selectivity  $S$  as the ratio of steady-state flux of a binding versus a non-binding species Figure 2.1B

$$S = \frac{J_{\text{binding}}(t \rightarrow \infty)}{J_{\text{non-binding}}(t \rightarrow \infty)}. \quad (2.3)$$

### 2.1.1 No selective transport occurs if bound TFs are immobile

If TF-FG Nup binding immobilizes the TF, the bound-state diffusion coefficient  $D_B = 0$ . For immobile bound TFs, transport is not selective: the steady-state flux  $J = D_F T_L / L$  for both binding and non-binding proteins, so  $S = 1$  (figs. 2.2, ??). The binding TF accumulates within the pore, but its immobility means it does not enhance transport compared to the non-binding case. Notably, this effect is independent of binding kinetics. Prior to steady state, binding slows transport Figure 2.2A. In systems such as airway mucus, immobilization may increase the time available for degradation or active clearance, consistent with the observation that binding tends to inhibit selective transport in those systems [?, ?, ?]. This effect is related to the binding-site barrier seen in antibody delivery to tumors [?], and observations that non-binding nanoparticles are often more effective in drug delivery to tumors than binding particles [?].

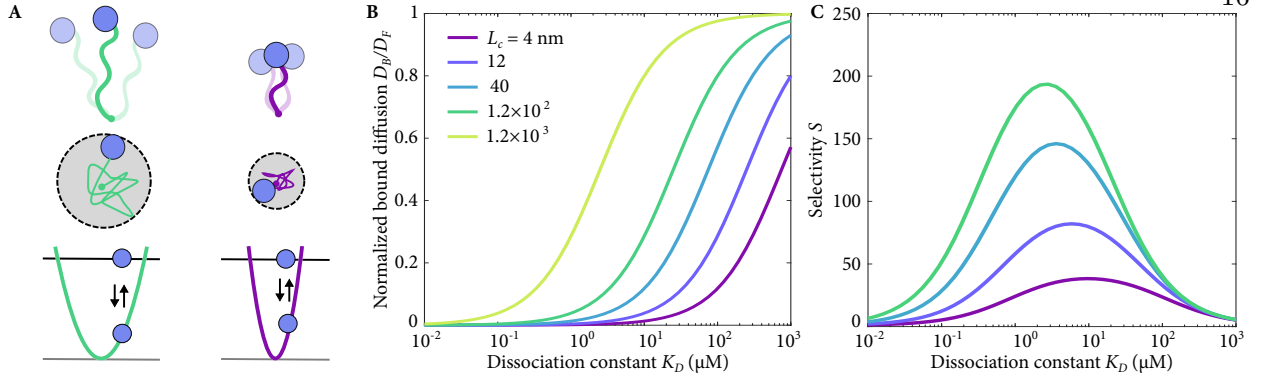


Figure 2.3: (A) Schematic of the flexible tether model of bound-state diffusion. FG Nups are treated as entropic springs that constrain the motion of TFs more (top and center left, longer FG Nup) or less (top and center right, shorter Nup), which corresponds to changing width of the harmonic potential well (lower). (B) Ratio of bound to free diffusion coefficient as a function of dissociation constant, with varying polymer length in the tethered-diffusion model. (C) Selectivity as a function of  $K_D$ , with varying polymer length in the tethered-diffusion model.

Our model is related to the classic problem of molecular transport through an oil membrane separating two aqueous reservoirs [?]. The relative concentration of a species just inside the oil barrier to the concentration in water is called the partition coefficient. The steady-state flux through the membrane is directly proportional to the partition coefficient (Supporting Information, section ??, fig. ??). By analogy, one might expect the TF-FG binding affinity to determine the flux across the pore. However, binding is different from partitioning. In systems where the increase in intra-pore concentration arises from binding, the effective diffusion coefficient is typically inversely proportional to the partition coefficient, making the flux independent of binding affinity [?]. This result led us to consider whether TFs may be mobile while bound to FG Nups.

### 2.1.2 Bound-state diffusion allows selective transport

When bound TFs are mobile, selective transport occurs with a selectivity up to 240 for a conservative set of parameters (figs. 2.2B,C, ??, Supporting Information, section ??). Remarkably, this selectivity is comparable to experimental measurements of NTF2 versus GFP flux (Table 2.1). The interplay between binding kinetics and diffusion leads to an optimal dissociation constant  $\sim 1 \mu\text{M}$  for maximum selectivity Figure 2.2C. Selectivity decreases for high  $K_D$  because binding is too

weak to significantly increase TF concentration in the pore. For low  $K_D$ , tight binding causes the concentration of bound complexes to become approximately constant across the pore. Because diffusive flux is driven by a concentration gradient, this washing out of the gradient by tight binding decreases flux and selectivity.

Our model predicts that selectivity is increased by increasing binding on-rate constant  $k_{\text{on}}$  Figure ???. Consistent with this, the on-rate constants of TF-FG Nup interactions have been measured to be diffusion limited [?, ?]. Large  $k_{\text{on}}$  makes transport more selective because fast binding kinetics relative to diffusive motion are necessary to maintain steep concentration gradients within the pore. High FG Nup concentration (as measured experimentally) leads to large  $N_t$  and low  $D_F$ , both of which increase selectivity. Decreasing  $D_F$  or increasing the length of the pore both reduce the magnitude of the flux and increase selectivity (figs. ??, ??). Therefore, varying TF free diffusion coefficient and pore length involves a trade-off between transit time and selectivity.

## 2.2 Mechanisms of bound transport factor mobility

Our result that bound-state diffusion is required for selective transport raises a mechanistic question: how can TFs move while bound to FG Nups? Here we consider two experimentally based mechanisms: movement of the bound TF due to the intrinsic flexibility of the FG Nups [?] and

Table 2.1: Comparison between experimental results for NTF2 and GFP (a similarly-sized non-binding protein) and model predictions. Flux measured in units of molecules per pore per second.

Method	Cell type	Species	Flux	Selectivity	Notes
OSTR	<i>Xenopus</i>	NTF2 GFP	91–123 3.3–3.8	24–37	[?]
OSTR	<i>Xenopus</i>	NTF2 GFP	47.3 1.1	43	[?]
Permeabilized cells	HeLa	NTF2 GFP	250 2	125	[?]
Model	–	Binding Non-binding	2–480 2	1–240	This work

multivalent binding that allows hopping of TFs between neighboring Nups [?].

### 2.2.1 FG Nup flexibility allows tethered diffusion

Previous measurements have found that FG Nups are flexible and dynamic [?,?,?]. One end of an FG Nup is attached to the inner ring of the NPC scaffold. However, the Nup's flexibility means that a TF which binds far from the scaffold-attachment point is mobile. Tethered diffusion therefore provides a mechanism of bound-state diffusion and selectivity.

#### 2.2.1.1 Bound-state diffusion depends on polymer and kinetic properties.

Flexible polymers behave as entropic springs [?] if they are not highly stretched. Therefore, a bound TF diffuses while attached to a spring-like tether, which can be represented as diffusion in a harmonic potential well Figure 2.3A. The width of the harmonic well is related to the length of the flexible domain. The effective length is the full FG Nup length if the FG Nups are not crosslinked, while the effective length is reduced if they are crosslinked or entangled [?].

In order to calculate the bound diffusion coefficient of the TFs, an averaging procedure is followed. First, the diffusion is assumed to be Fickian, which is a reasonably good though not perfect assumption (see section on Fickian diffusion). In the Fickian diffusion case, the diffusion coefficient is proportional to a mean-squared displacement (MSD) divided by time. We calculate the mean binding lifetime  $\tau$  and the MSD corresponding to this “typical” binding event and divide them.

To begin, note that the duration of a binding event follows the exponential distribution

$$\rho(t) = \exp(-t/\tau)/\tau, \quad (2.4)$$

where  $\tau = 1/k_{\text{off}}$  is the mean binding lifetime.

Next, the positional probability density of a bound TF is

$$P(x, t) = e^{-\frac{x^2}{2\alpha(t)}} / \sqrt{2\pi\alpha(t)}, \quad (2.5)$$

$$\alpha(t) = (1 - e^{-2kD_F\beta t}) / (k\beta) \quad (2.6)$$

where  $k$  is the spring constant of FG Nup tethering and  $1/\beta = k_B T$  is the thermal energy [?]. The center of the well is set at  $x = 0$ .

The mean-squared displacement (MSD) of the TF as a function of time is calculated, as any expected value, with the integral

$$\langle x^2(t) \rangle = \int_{-\infty}^{\infty} P(x, t) x^2 dx = \alpha(t). \quad (2.7)$$

Finally, the typical TF MSD during a binding event can be determined by evaluating

$$\overline{\langle x^2 \rangle} = \int_0^{\infty} \rho(t') \langle x^2(t') \rangle dt' = \frac{2D_F L_c \ell_p}{L_c \ell_p k_{\text{off}} + 3D_F}. \quad (2.8)$$

Here we assume that the spring constant is that of a worm-like chain polymer  $k = 3/(2\beta L_c \ell_p)$ , where  $L_c$  is the contour length and  $\ell_p$  the persistence length [?].

Combining these results, the one-dimensional bound diffusion coefficient is

$$D_B \approx \frac{\overline{\langle x^2 \rangle}}{2\tau} = \frac{D_F L_c \ell_p k_{\text{off}}}{L_c \ell_p k_{\text{off}} + 3D_F} = \frac{D_F}{1 + 3\frac{D_F}{D_P}}. \quad (2.9)$$

Here  $D_P = L_c \ell_p k_{\text{off}}$  controls the bound-state diffusion coefficient: higher  $D_P$  corresponds to a lower constraint of the TF by the tether and greater bound mobility. Bound mobility increases with increasing chain length and persistence length, or decreasing binding lifetime. When  $D_P$  is large ( $D_F/D_P \ll 1$ ),  $D_B$  approaches  $D_F$ , since the long chains barely affect TF motion during the short binding event. For small  $D_P$  ( $D_F/D_P \gg 1$ ), TF motion is inhibited by a short tether, giving  $D_B \approx D_P/3 \ll D_F$ . This result highlights that the kinetics of TF-FG Nup interaction are a primary determinant of the bound mobility: the faster the binding kinetics, the higher the bound diffusion constant.

Flexible disordered proteins typically have low persistence lengths [?], so we estimate  $\ell_p \approx 1$  nm. If the on-rate constant is diffusion limited,  $k_{\text{on}} = 10^{-3} \mu\text{M}^{-1} \mu\text{s}^{-1}$  [?, ?], the binding affinity determines the off rate. Disordered FG Nups have  $L_c \approx 100\text{--}280$  nm (250–700 amino acids long [?] with a contour length per amino acid  $\approx 0.4$  nm). For our conservative parameters, tethered diffusion alone predicts selectivity  $S \approx 200$ , a large flux enhancement for TFs over nonbinding proteins.

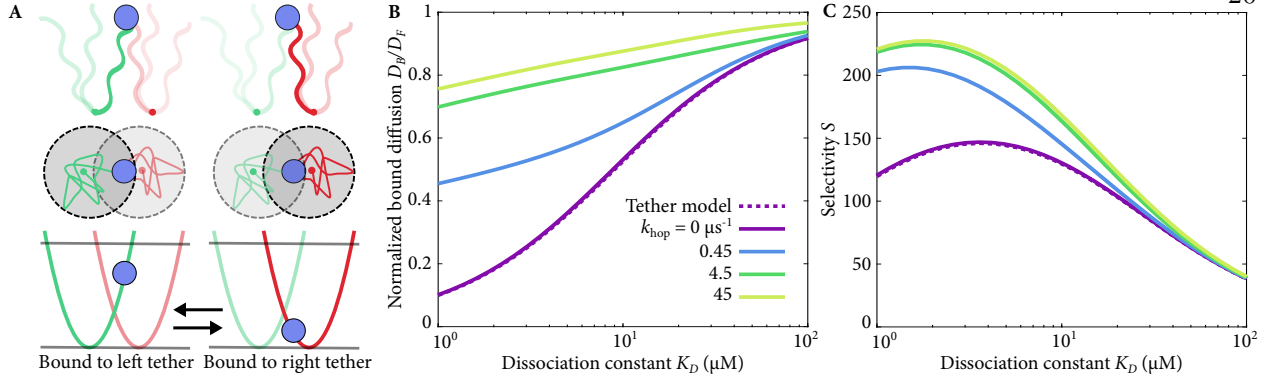


Figure 2.4: (A) Schematic of the inter-chain hopping model of bound-state diffusion. FG Nups are treated as entropic springs that constrain the motion of TFs, and inter-chain hopping allows a TF to move from one FG Nup (top and center left, green Nup) to another (top and center right, red Nup) without unbinding, which corresponds to switching from one harmonic potential well to another (lower). (B) Ratio of bound to free diffusion coefficient as a function of dissociation constant, with varying hopping rate in the inter-chain hopping model. (C) Selectivity as a function of  $K_D$  with varying hopping rate. FG Nup contour length  $L_c = 40$  nm in (B, C).

### 2.2.2 Inter-chain hopping increases selectivity

The tethered diffusion mechanism is constrained by a trade-off: tighter binding increases the TF concentration in the pore, but hinders motion. Multivalent TF-FG interactions can relax this constraint, because a TF can bind simultaneously to more than one FG Nup, moving hand-over-hand while remaining bound [?]. Consistent with this, TFs may slide between nearby FG sites rather than fully unbinding and re-binding [?]. If the newly-bound FG repeat is on a neighboring chain, the FG tether site that constrains TF motion moves while the TF remains bound. We model inter-chain hopping with a TF that undergoes tethered diffusion when bound to an FG Nup and hops between neighboring, randomly distributed tethers Figure 2.4. Numerical simulations of this model determined the bound diffusion coefficient (fig. 2.5, Supporting Information, section ??). We note that intra-chain hopping does not change the flux, since the anchor point of the tethering chain is not changed; therefore we neglect it.



### 2.2.3 Detailed explanation of hopping simulation

In our simulation of TF motion with hopping between FG Nups while bound, we represented each FG Nup as an entropic spring (i.e. as a harmonic potential well). Well positions were randomly chosen from a uniform distribution, with the exception that we always placed one well at the starting position of the TF. The particle (the TF) started the simulation bound to this FG Nup, and remained bound throughout the simulation. While bound to one FG Nup, the TF diffused within the harmonic well representing that FG Nup. We recorded the position and mean-squared displacement of the TF from its starting location, which we then used to determine a bound diffusion coefficient, as described in more detail below. The TF could hop between tethers by changing which well it moved in.

#### 2.2.3.1 Diffusion in a potential well

The TF moved in the harmonic potential of the FG Nup according to Brownian dynamics. At each timestep, the TF position was updated using a force-dependent diffusive step [?].

$$x(t + \delta t) = x(t) + \frac{F}{\Gamma} \delta t + \delta x, \quad (2.10)$$

where  $F$  is the force acting on the particle,  $\Gamma$  is the drag coefficient,  $\delta t$  is the timestep, and  $\delta x$  is a random Brownian step drawn from a Gaussian distribution with variance  $\sigma^2 = 2D\delta t$ . The drag coefficient of a spherical particle at low Reynolds number is given by Stokes' Law as  $\Gamma = 6\pi\eta r$ , where  $\eta$  is the fluid's viscosity and  $r$  is the sphere's radius. This result can be combined with the Einstein relation  $D = k_B T / (6\pi\eta r)$  to give

$$\Gamma = \frac{k_B T}{D}. \quad (2.11)$$

The force  $F = -k\Delta x$ , where  $k$  is the spring constant of the FG Nup and  $\Delta x$  is the displacement of the particle from the Nup attachment point. We model the FG Nup as a worm-like-chain at small extension, so that  $k = 3k_B T / (2\ell_p L_c)$ , where  $\ell_p$  is the tether persistence length and  $L_c$  is

the contour length. Then

$$x(t + \delta t) = x(t) - \frac{3D\Delta x\delta t}{2\ell_p L_c} + \delta x = x(t) - DK\Delta x\delta t + \delta x, \quad (2.12)$$

where  $K$  is the normalized spring constant  $K = k/k_B T = 3/(2\ell_p L_c)$ .

### 2.2.3.2 Hopping probability

We designed the hopping probability  $P_{\text{hop}}$  in order to satisfy the principle of detailed balance. During every iteration of the simulation, we picked an FG Nup at random from a list of the  $M$  Nups near enough to have a reasonable probability of hopping. TF hopping to the new FG Nup was attempted with success probability

$$P_{\text{hop}} = r_{\text{hop}} M \delta t e^{-\Delta G/2}. \quad (2.13)$$

Here the base hopping rate  $r_{\text{hop}}$  is a dimensionless input parameter, and the change in free energy (in units of  $k_B T$ ) between the current Nup and the proposed new Nup is

$$\Delta G = \frac{1}{2} K (x - x_{\text{new}})^2 - \frac{1}{2} K (x - x_{\text{cur}})^2, \quad (2.14)$$

where  $K$  is the normalized spring constant,  $x$  is the particle's current position,  $x_{\text{cur}}$  is the anchor location of the Nup to which the particle is currently bound, and  $x_{\text{new}}$  is the anchor location of the proposed new Nup. Note that when a hop succeeds, the energy landscape changes to that of the new Nup, but the TF's position does not change during the hop. There is no upper bound on  $P_{\text{hop}}$ , but we adjusted the timestep to ensure that  $P_{\text{hop}}$  was greater than unity no more than 0.5% of the time that a hop was attempted.

### 2.2.3.3 Mean-squared displacement and diffusion coefficient calculation

We ran each simulation for  $10^7$  time steps with  $\delta t = 0.01 \mu\text{s}$ , and recorded the particle's position every 100 time steps. We calculated the mean-squared displacement  $\langle x^2 \rangle$  (MSD) of the TF and averaged it over 100 runs Figure 2.5A. We then computed

$$\rho_{\text{MSD}}(t) = \langle x^2(t) \rangle \rho(k_{\text{off}}, t) = k_{\text{off}} \langle x^2(t) \rangle e^{-k_{\text{off}} t}, \quad (2.15)$$

as shown in fig. 2.5B, and numerically integrated the distribution in time. We determined the bound diffusion coefficient from the typical MSD-per-binding-event  $\overline{\langle x^2 \rangle}$  using

$$D_B = \frac{k_{\text{off}} \overline{\langle x^2 \rangle}}{2}. \quad (2.16)$$

Here, the factor of 1/2 is appropriate because we consider a one-dimensional random walk.

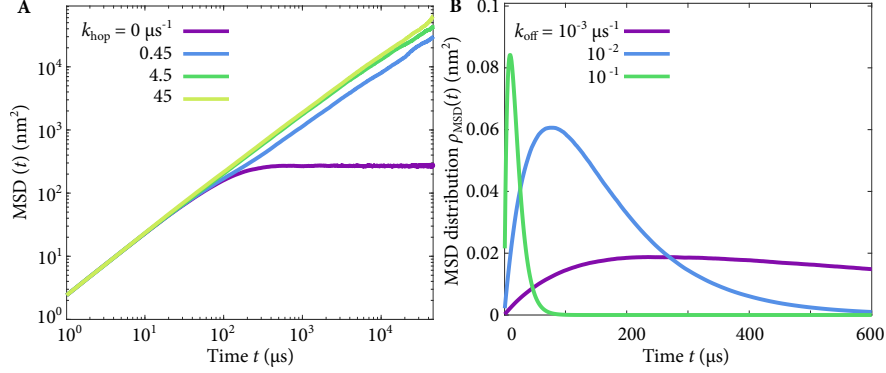


Figure 2.5: (A) Examples of mean-squared displacement (MSD) of a simulated TF in the inter-chain hopping model, with varying hopping rate. (B) Examples of MSD distributions  $\rho_{\text{MSD}}(t)$  used in estimating the diffusion coefficient, with varying unbinding rate. Tethers have 40 nm contour length; other parameters are as discussed in the text.

#### 2.2.4 Hopping simulation results

Inter-chain hopping increases selectivity most for tight binding and short chains, the parameter regime where tethered diffusion gives limited selectivity (figs. 2.4, 2.6, 2.7). Hopping may therefore be important for FGs that form transient crosslinks: if FG Nups are highly crosslinked, our model suggests that inter-chain hopping is the key mechanism of TF movement. For weaker binding and longer chains, inter-chain hopping leads to a modest increase in selectivity.

#### 2.2.5 Fickian diffusion

Talk about how diffusion is almost but not quite Fickian, and show the anomalous diffusion plots (made with Mike’s script) for various parameters. Talk about timescales and how this might affect the results.

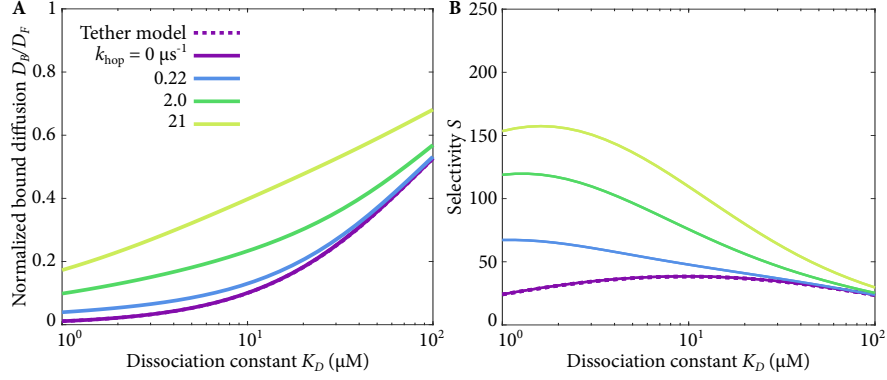


Figure 2.6: Bound diffusion and selectivity as a function of dissociation constant, with varying hopping rate for FG Nups with  $L_c = 4$  nm.

### 2.3 Discussion

A key puzzle of the NPC is how transport-factor binding allows rapid transport through the pore. Binding typically immobilizes the bound particle, and so the increase in concentration resulting from binding does not, in general, result in increased flux. The biophysical theory we developed includes diffusion of TFs due to thermal fluctuations, binding to polymeric tethers, and the hopping of bound species between these tethers. Thus we identified principles of selective transport resulting from binding Figure 2.1, emphasizing that bound-state mobility is essential for selective transport Figure 2.2. Binding increases the local concentration, and any bound mobility increases the flux. We characterized two mechanisms to obtain bound-state mobility and found that thermally-driven diffusion of TFs bound to flexible tethers and rapid binding kinetics [?, ?] allow TF mobility, leading to selectivity similar to that observed experimentally Figure 2.3. In addition, tether flexibility enables multivalent bound particles to hop between binding regions Figure 2.4 [?, ?], further enhancing selectivity. Mobility of bound or partitioned molecules occurs in many biological contexts, suggesting that the mechanisms we study here may be broadly applicable [?, ?].

Our model for selective transport by tethered diffusion generalizes to a range of FG-FG interactions [?], if we decrease the effective chain length  $L_c$  for cohesive FG Nups. For short chains the selectivity simply due to chain flexibility is modest, suggesting that other mechanisms,

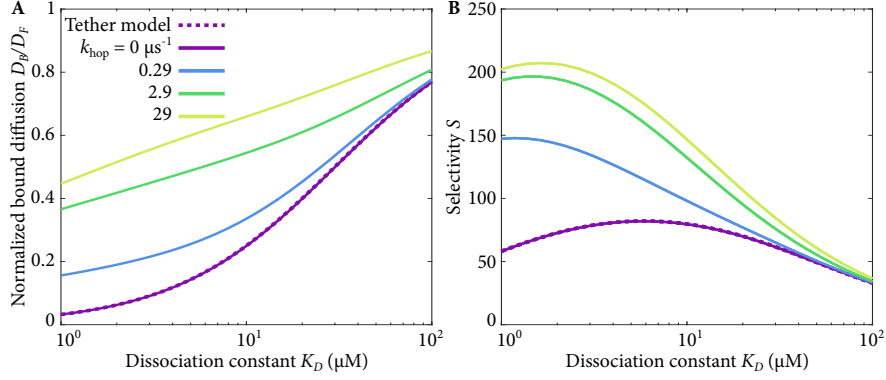


Figure 2.7: Bound diffusion and selectivity as a function of dissociation constant, with varying hopping rate for FG Nups with  $L_c = 12$  nm.

like hopping, may be important. Our model suggests that transient cross-linking of FG repeats proposed to occur within the pore may serve to increase the viscosity and therefore the selectivity. Crosslinks need not be actively melted by TFs to enhance selectivity Figure ??.

Our model provides a quantitative tool to evaluate selective transport. Materials formed *in vitro* by spontaneous self-assembly of FG Nups [?] or transient crosslinking by alpha-helical peptides [?] show strong selective *entry*. Using published data, we predicted whether these gels also showed selective *transport* (table ??). Most synthetic gels are predicted to have  $S < 10$ , less than the selectivity of NTF2 in cells (table 2.1). The predicted selectivity of one hydrogel is  $S \approx 200$ , apparently the most selective synthetic gel to date [?].

### 2.3.1 Overcoming the limitations of binding

Binding, even in the presence of bound-state motion, limits selectivity. Biological systems appear to have developed strategies to avoid this, for example, by using true partitioning. Lipid domains in complex membranes partition proteins [?]. Membraneless organelles spontaneously assembled from low-complexity proteins and nucleic acids can localize a molecule without immobilizing it [?]. Because membraneless organelles are fluid, the constraints imposed in our NPC model by binding are released. Our work thereby suggests a benefit of phase-separated droplets to cells: they provide significantly higher selectivity than can occur with immobilizing binding. This may

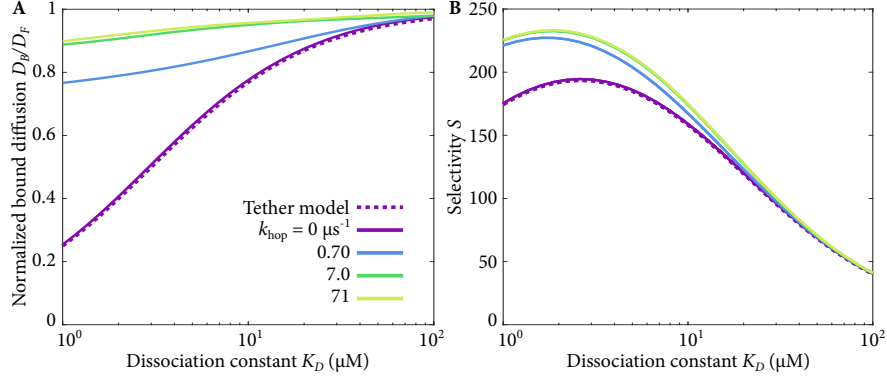


Figure 2.8: Bound diffusion and selectivity as a function of dissociation constant, with varying hopping rate for FG Nups with  $L_c = 120$  nm.

be especially important for spatially complex assemblies [?].

Though we show it is not necessary, the active dissolution of polymeric biomaterials has been proposed to occur in the NPC [?]. This strategy is used by *Helicobacter pylori* to penetrate the gastric mucus [?]. Because the particularly dense extracellular matrix of solid tumors blocks the motion of particles, especially larger nanoparticles, ECM dissolution has been used to enhance drug delivery [?]. Unfortunately, this approach may not be universally applicable: breaking down the ECM surrounding tumors may promote cancer metastasis [?].

### 2.3.2 Design principles of selective transport by binding

Filtering by polymeric biomaterials occurs in many systems for particles of different sizes: for example, nutrients reach our intestinal walls while larger molecules are excluded. However, controlling the selective transport of similarly-sized molecules by tuning specific interactions has proven elusive. In drug delivery applications, inert nanoparticles are typically more effective at penetrating extracellular spaces and reaching their cellular targets [?]. Because biopolymer filters are the first point of contact of nanoparticles used for drug delivery, specific targeting of transport through mucus may enhance the effectiveness of drug delivery. If NPC-like bound mobility as described in our model could be achieved in these systems, it would increase the rates of transport and drug delivery.

## Chapter 3

### Hydrogels yay

#### 3.1 Pore size

#### 3.2 Porogens

I tried to make gels with a larger pore size by using porogens. These are nanoscale particles that can be mixed into the gel precursor solution and then removed/dissolved after crosslinking. The idea is that then they would leave pores of a better size and also make it easier for the gel to swell to equilibrium. I tried alginate nanospheres (which could then be removed by chelating the calcium ions) and large molecular-weight dextrans that could be removed by digesting with dextranase. Neither of these options worked out. I couldn't get the nanospheres to reliably resuspend in the precursor solution. The dextran/dextranase system was more promising, but too much dextran remained in the hydrogels even after dextranase digestion. Also the dextranase chewed up the FSFG.

##### 3.2.1 Alginate nanospheres

Figures:

- (1) diagram of fabrication process
- (2) DLS results on size distribution

### 3.2.2 Dextran/dextranase system

Figures:

- (1) diagram of porogen process + molecular structure of dextran, dextranase?
- (2) removal of fluorescent dextran from gel
- (3) comparison of gels with and without 250 kD and 70kD dextran
- (4) weird mCherry suppression?

### 3.3 Polymerization using confocal microscope

5) Writing thin gels with confocal microscope a. 50 um minimum line thickness – good b. Wrote thin lines and donuts c. Tried using photoinhibitors to reduce stray crosslinking, but didn't work well enough d. Swelling/buckling was a problem

### 3.4 Nup and transport factor constructs

6) Constructs a. SSSG as a negative control for FSFG gels (this actually works) b. Covalently-tethered NTF2 (didn't work) c. FSFG cct2 (works) and cct3 (doesn't express) d. Kap121 with GFP-NLS (only tested once, didn't enter gel)

### 3.5 Dye-labeling and free dye

7) Fixed free dye problem a. Rigorous labeling protocol and testing, keep frozen until use b. Tried a few different labeling chemistries (maleimide, NHS and SDP esters) c. Made GFP-NTF2 (didn't bind to FSFG)



## **Chapter 4**

### **Actual real results**

I am going to have results from the hydrogel nuclear pore mimics.

## Chapter 5

### FG Nup aggregation under crowded conditions

#### 5.1 Results

##### 5.1.1 Thioflavin timecourses

We analyzed aggregation dynamics using thioflavin T, a fluorophore whose excitation and emission maxima shift and which grows much brighter in the presence of thioflavin. Samples containing FG124, crowding agent, and thioflavin T were incubated while shaking overnight in a 96-well plate and their fluorescence intensity recorded every ten minutes. The resulting traces show a typical aggregation pattern, with a lag phase, burst phase, and saturation phase.

The final concentration of crowding agent was 19% serine w/v, 13% PEG, and 13% PVP. Lysate concentration varied between time courses and was in the 1-10 mg/mL range. Two timecourses were run with varying PEG and PVP concentrations: 25%, 20%, 13%, and 5%. Samples with no crowding agent ('buffer' samples) were run as a reference, and samples in 7M guanidine hydrochloride (GuHCl) with no crowding agent were run as negative controls in each timecourse. No aggregation was observed in the GuHCl samples. In every case, blanks were run alongside the sample conditions. In the blanks, the FG124 was omitted.

First the data were normalized to the blanks. In nearly all cases, the blank intensity remained steady over time, as expected. In those cases, the mean blank intensity was subtracted from the corresponding data. (In cases where the blank intensity changed over time, it was subtracted pointwise from the data.)

Then the normalized curves were fit to a sigmoid given by

$$I(t) = C + \frac{A}{1 + \exp(-k(t - T_{1/2}))} \quad (5.1)$$

where  $I(t)$  is the normalized fluorescence intensity as a function of time. The useful physical parameters for our purposes are the time constant  $k$ , which gives a measurement of the steepness of the slope at the beginning of the burst phase, and the lag time  $T_l$ . The lag time is calculated as

$$T_l = T_{1/2} - \frac{2}{k} \quad (5.2)$$

and represents the duration of the lag phase  $\square$ .

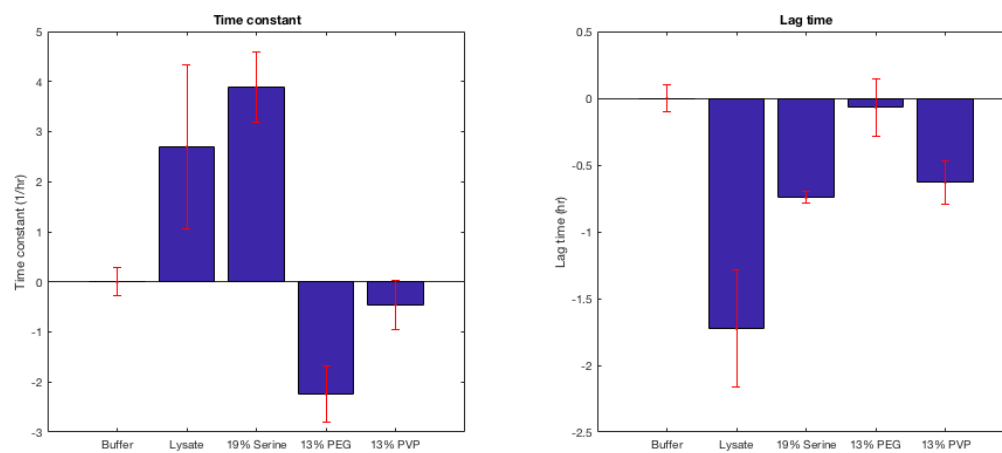
Ideally, we expect  $C = 1$ , as aggregation should not have begun at the start of the experiment, so there should be no increase in sample fluorescence over that of the blank. Experimentally,  $C$  usually ranged between 1 and 2, up to about 10 for some sample conditions. This might have been because aggregation had already begun, but this seems unlikely because the lag phase continued for some time. I don't know why  $C$  wasn't close to unity for all conditions and replicates.

The saturation phase asymptotes to an intensity given by  $I_{\text{sat}} = C + A$ . We found significant variation in  $I_{\text{sat}}$  for the same condition between timecourses, and the relative magnitudes of different conditions also varied between timecourses. I think this is a limitation of the thioflavin tests. Therefore, we do not consider  $I_{\text{sat}}$  in our analysis, leaving  $k$  and  $T_l$  as parameters of interest.

Next we normalized once again to account for differences between timecourses. It was impossible to hold the concentration exactly fixed between timecourses, and there were probably some other environmental variables we couldn't control perfectly, so it was important to normalize again. Within each timecourse, we averaged the fit parameters for all buffer replicates and subtracted that average from each other replicate.

We then combined replicates from all timecourses by condition and took the final average and standard error of the mean for both fit parameters. The results are shown in Fig. 5.1. I ran a one-factor ANOVA to reject the null hypothesis that all means were the same, and then two-sample t-tests on each pair of conditions to find which differences were statistically significant. For the

Figure 5.1: test



time constant, all pairs of conditions were significantly different from each other with  $p < 0.05$  **except** buffer-PVP. For the lag time, all pairs of conditions were significantly different from each other with  $p < 0.05$  **except** buffer-PEG and serine-PVP.

I followed the same process for the timecourses involving the concentration series for PEG and PVP, but there were fewer significant differences between concentrations, as seen in Fig. 5.2. The ANOVA for the PEG time constants gave  $p = 0.0024$  with the t-test showing differences between: buffer-25% PEG, buffer-13% PEG, and 25% PEG-20% PEG. The ANOVA for the PEG lag times was likewise significant with  $p = 0.0102$ . The significant pairs were: buffer-25% PEG, buffer-20% PEG, buffer-5% PEG, and 25% PEG-13% PEG.

For PVP, there were no significant differences in the time constant. For the lag time, after the ANOVA and t-tests, there were differences between: buffer and everything except 5% PVP, 25% PVP and everything but 20% PVP, and 20%PVP-13% PVP.

For both conditions, it looks roughly like the lag time increases with the crowder concentration, but given the large error bars, I wouldn't read too much into it. The time constants don't follow any particular trend and are mostly indistinguishable anyway.

We tested the effects of pH on FG124 aggregation by running several experiments in buffer at varying pH. We ran six replicates of each condition, for pH values between 5 and 8. No significant differences were found in the lifetimes or lag times of any pH condition. Results are summarized in Table 5.1. We concluded that aggregation is not affected by pH in the range 5-8.

Table 5.1: FG124 aggregation lifetime and lag time with varying pH. Each condition run with 6 replicates in PTB buffer. One-way ANOVAs show no statistically significant differences between conditions. Standard errors are shown.

pH	Lifetime $\tau$ (hr)	Lag time $T_{lag}$ (hr)
5	$0.34 \pm 0.09$	$6.8 \pm 0.1$
6	$0.35 \pm 0.12$	$6.2 \pm 0.2$
7	$0.38 \pm 0.14$	$6.8 \pm 0.4$
8	$0.50 \pm 0.14$	$6.6 \pm 0.2$

Figure 5.2: test

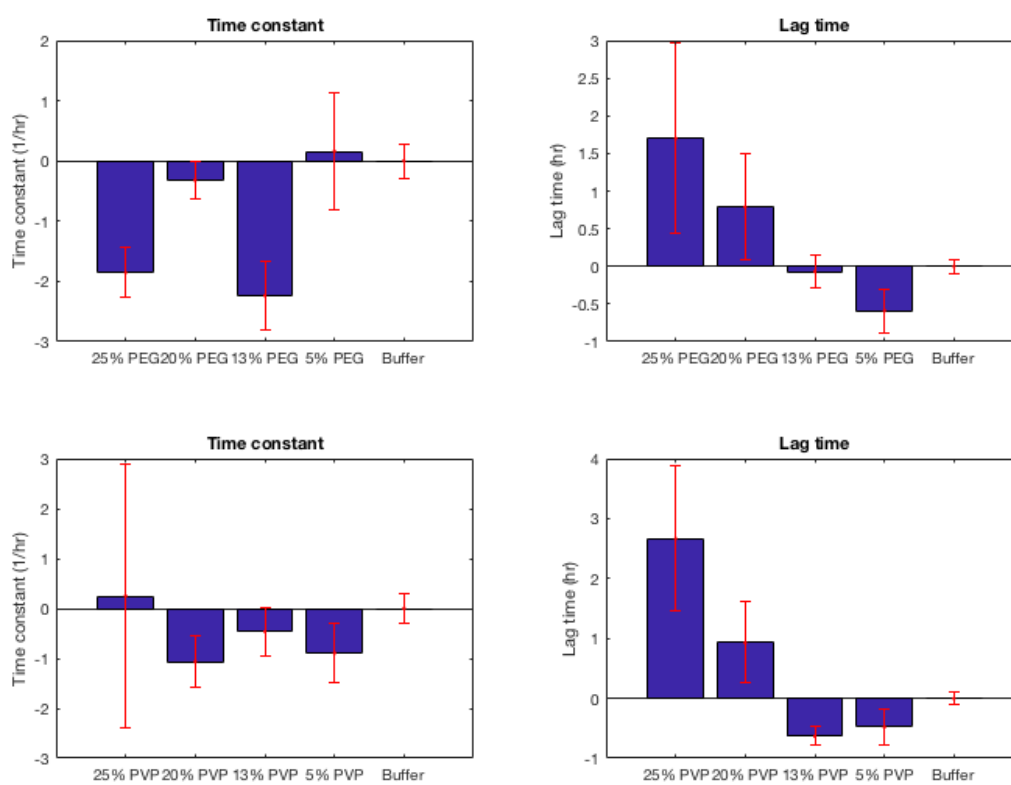
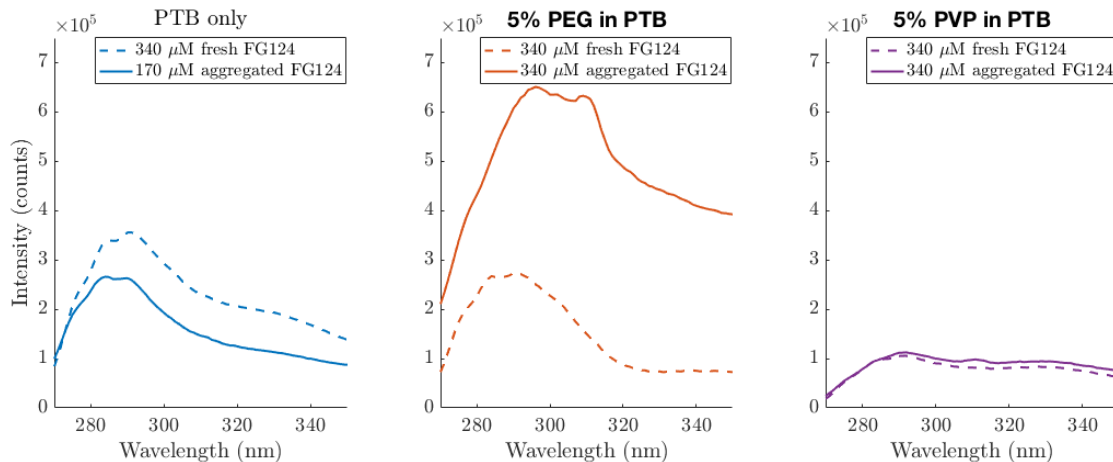


Figure 5.3: Emission scan of fresh and aggregated FG124 in crowded conditions. Data normalized by subtracting blank sample.



### 5.1.2 Fluorimetry

We hypothesized that the crowder-dependent difference in aggregation might arise from the aromatic ring in PVP, which could interact with the phenylalanine in FG124 in a ring-stacking interaction. To test this idea, I collected emission spectra of fresh and aggregated FG124 in PEG and PVP crowders near the phenylalanine peak wavelength (see Fig. 5.3). Only 5% PEG and PVP in PTB were tested because PEG has a peak near the FG124 peak which dominates at higher PEG concentrations. Data were normalized by averaging over two runs and subtracting a blank run (containing crowder and buffer but no protein). The PEG sample showed the largest difference upon aggregation, both in peak height and location. Figure 5.4 shows the same data, normalized to a maximum amplitude of one and offset, in order to emphasize the changes in peak shape. In both crowder conditions, but not in the buffer condition, a small peak appears in the aggregated FG124 near 310 nm.

Additionally, I compared the phenylalanine peaks in FSFG and fresh and aggregated FG124 to those of pure phenylalanine [?], as seen in Fig. 5.5. Data are blanked and normalized to a maximum intensity of one. The peak slightly shifts toward longer wavelengths as the data progress from phenylalanine to FSFG to fresh and aggregated FG124. I also measured PEG and PVP

Figure 5.4: Emission scan of phenylalanine, FSFG, and FG124. Excited at 240 nm. Phe data is not mine, need to check reference. Data normalized to a maximum amplitude of 1.

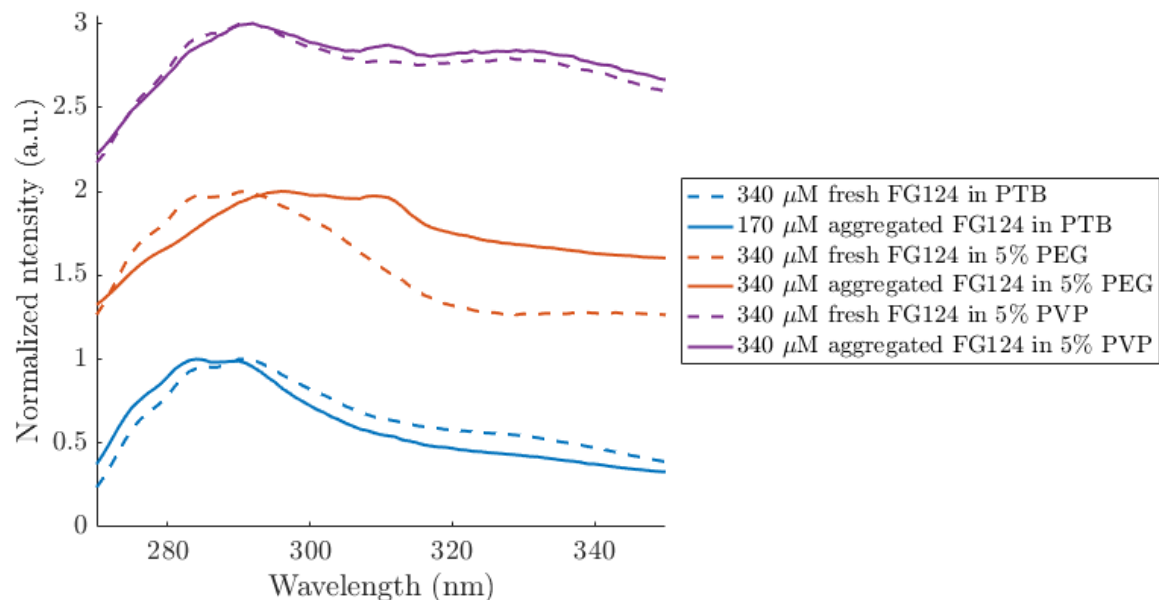


Figure 5.5: Emission scan of phenylalanine, FSFG, and FG124. Excited at 240 nm. Phe data is not mine, need to check reference. Data normalized to a maximum amplitude of 1.

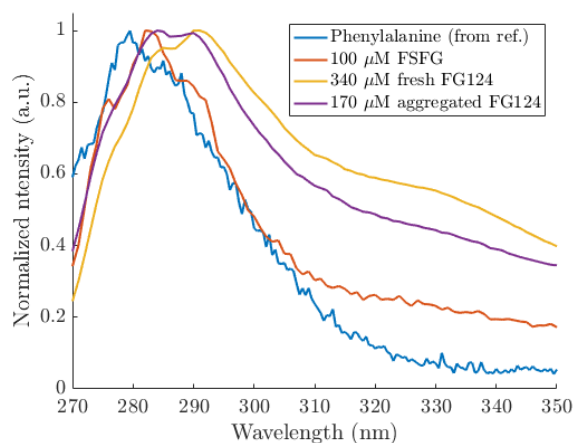
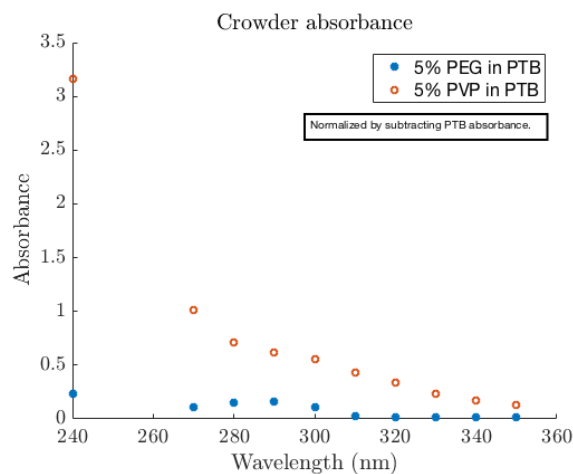




Figure 5.6: Absorbance of 5% PEG and PVP solutions in PTB. Normalized by subtracting PTB absorbance.



absorbance near 240 nm (see Fig. 5.6). PVP has a very high absorbance in that range, which might explain why the recorded counts were lowest for the PVP conditions.

## 5.2 Discussion

## 5.3 Materials and Methods

### 5.3.1 Buffers

Potassium transport buffer (PTB) (150 mM KCl, 20 mM HEPES, 2 mM  $\text{MgCl}_2$ ) was used for all timecourse and NMR samples.

### 5.3.2 FG124 preparation

His-tagged FG124 was expressed in *E. coli* in the plasmid pRSF. Cultures were grown in LB and induced at 37 degrees C for 2-4 hr with 1 mM IPTG at OD 0.6-0.8. Periplasmic matrix was removed prior to lysis. Cells were then lysed via sonication and FG124 purified using TALON cobalt resin. All purification buffers were PTB with 7M GuHCl and PIC. The elution buffer also contained 250 mM imidazole.

### 5.3.3 Timecourse preparation

Stocks of PEG and PVP in PTB were prepared at 20 or 40% w/v; serine stocks were prepared at 30% w/v. PEG and serine were at pH 7; PVP was pH 7 or pH 5. A pH series with no crowding agent showed no significant differences based on pH. Lyophilized lysate was prepared by homogenizing BL21 DE3 Gold cells and spinning them down. The supernatant was lyophilized in a decomposing ammonium bicarbonate buffer and resuspended in PTB to the desired concentration when needed. A 10 mM stock solution of ThT in PTB was prepared and filtered no more than a week before the timecourse, stored at room temperature and protected from light. Immediately prior to starting the timecourse, FG124 was desalted into PTB to remove the imidazole and GuHCl. Samples were promptly prepared containing the appropriate percentage of crowder, a final concentration of 1-2 mg/mL FG124, and 200  $\mu$ M thioflavin T. All samples in the same timecourse had the same concentration of FG124, including the buffer sample, which contained no crowding agent. Blanks were prepared with crowding agent and thioflavin T, but no FG124. Samples were pipetted into black, flat-bottomed, clear-bottomed 96-well plates with 150  $\mu$ L per replicate. Each sample yielded four to six replicates. Only one blank replicate was used per condition. One negative control and corresponding blank were prepared per timecourse containing 7M GuHCl and no crowding agent but using the same protein sample as all other conditions. Each well contained a 3mm-diameter glass or teflon bead. The plate was sealed with a PCR seal and taken to a Safire II plate reader. The fluorescence was measured from the bottom at 10-minute intervals with an excitation wavelength of 450 nm, emission wavelength of 482 nm, and 5 nm bandwidths. The plate shook orbitally at high speed between measurements and was held at a temperature of 30 degrees C. The time between desalting and beginning the plate reader measurements was typically about an hour; the time of desalting was taken as  $t = 0$  for the purposes of calculating lag time. In parallel with the sample preparation, the concentration of the desalted FG124 was measured with a BCA assay.

#### **5.3.4 NMR sample preparation**

#### **5.3.5 NMR experiments**

#### **5.3.6 Fluorimetry**

FG124 was purified as described above and stored in PTB with 7M GuHCl. Immediately before use, 130 uL of 520 uM FG124 was desalted with a Zeba spin desalting column to remove the GuHCl. The resulting stock was used in the crowder samples, which had a final concentration of 340 uM Fg124. I used the shared-instrumentation fluorimeter. I tested PTB, 5% PEG (MW, source, purity?) in PTB, and 5% PVP (same questions?) in PTB as blanks. I measured 340 uM FG124 in PTB and the two crowders as well. Between runs, I cleaned the cuvette (micro quartz cuvette from Kaar lab (sample volume?)) with ethanol 3x, then 5x with DI water, and gently blotted the outside with ethanol. Fluorimeter settings: 4 nm slits, 1000 V PMT, excitation wavelength of 240 nm. Step size = 1nm, average over 2 runs. After taking this data with fresh FG124, I let the samples sit at room temp overnight to aggregate (no shaking) and most were cloudy in the morning. I took similar data with the aggregated samples. I needed to rinse with 7M GuHCl, let soak in 7M GuHCl for 5 minutes, and then perform the same cleaning procedure as the previous day in order to remove the aggregates from the cuvette.