ROLES FOR UBIQUITIN AND DIMENSIONAL DEPENDENCE IN PROTEIN REGULATION

by

Megan Gorringe Dixon

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

The University of Utah

August 2015

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of   **Megan Gorringe Dixon**

has been approved by the following supervisory committee members:

James Keener	, Chair	**5/27/2015**	Date Approved
Markus Babst	, Member	**5/27/2015**	Date Approved
Frederick Adler	, Member	**5/27/2015**	Date Approved
Aaron Fogelson	, Member	**5/27/2015**	Date Approved
Alla Borisyuk	, Member		Date Approved

and by   **Peter Trapa**   , Chair/Dean of

the Department/College/School of   **Mathematics**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Ubiquitin is a small protein which interacts with other proteins as a post-translational modification and as a binding partner for proteins which contain a ubiquitin binding domain (UBD). Proteins modified with ubiquitin are often targeted for degradation. Ubiquitin regulates both soluble and membrane-bound proteins in cells of nearly all tissues. Here we use mathematical models to study three distinct regulatory systems involving ubiquitin: regulation of the yeast uracil transporter, Fur4, protein sorting mediated by the endosomal sorting complexes required for transport (ESCRTs), and regulation of Rad18 in the DNA damage tolerance pathway. Using a differential equation model of Fur4 regulation, we demonstrate that deubiquitination and retention are essential roles of the Rsp5/Ubp2 complex localized to the endosome. We also predict a nearly constant pool of endosomal Fur4 independent of extracellular conditions. ESCRTs are responsible for sorting ubiquitinated proteins (cargo) on the endosomal membrane prior to formation of intralumenal vesicles. However, the mechanisms of sorting remain unclear. Motivated by recent experimental data, we present a cellular automata model of ESCRT sorting which demonstrates that a flexible network of ESCRTs and cargoes is sufficient for high efficiency sorting under specified rules. ESCRT-cargo networks exist on membranes while all ESCRT binding studies consider ESCRT interactions in solution. We present novel results on the dimensional dependence of dissociation constants for general protein-protein interactions using stochastic methods. We present a conversion for transforming three-dimensional dissociation constants to two-dimensional dissociation constants and demonstrate that ESCRT-cargo interactions are more stable on membranes than in solution. Using our computed two-dimensional reaction rates, we present an ODE model for the evolution of the size of ESCRT-cargo networks. Our results suggest that ESCRT-mediated sorting can be achieved on the order of seconds. Lastly, we examine ubiquitin-dependent regulation of Rad18 in the DNA damage tolerance pathway, a system of strictly soluble proteins which does not rely on ubiquitin-dependent degradation. Results of ODE models suggest that the dissociation constants for Rad18 binding events must be measured in order to better understand the mechanisms behind damage-specific responses.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Since first being discovered in 1975 by Goldstein et al. in the thymus [15], ubiquitin is now known to be present in nearly all tissues and crucial in cellular protein regulation. Ubiquitin is a small protein composed of 76 amino acid residues [47]. Ubiquitin interacts with both soluble and membrane-bound proteins as a post-translational modification and through weaker bonds mediated by ubiquitin binding domains. In 2004, Aaron Ciechaover, Avram Hershko, and Irwin Rose were awarded the Nobel Prize in chemistry for their identification of ubiquitin as a protein degradation signal [13]. The process of ubiquitin tagging, known as ubiquitination, is carried out by three types of enzymes: ubiquitin activating enzymes, ubiquitin-conugating enzymes, and ubiquitin ligases named E1, E2, and E3 enzymes, respectively. These enzymes are responsible for the covalent attachment of the c-terminal glycine residue of ubiquitin to a lysine residue on a substrate protein. This process is reviewed in [35]. Deubiquitinating enzymes (DUBs) are required to cleave the ubiquitin from the ubiquitin-tagged proteins. While it is important to understand the ubiquitin tagging process, the aim of this work is to understand the larger scale effects of ubiquitination and deubiquitination in different settings. Ubiquitin also interacts with proteins via ubiquitin binding domains. More than 16 different ubiquitin binding domains have been identified, most with very low affinities for ubiquitin ($Kd \sim 100 \mu M$). (Reviewed in [18]). Through its interactions with various proteins, ubiquitin plays a critical role in a vast number of cellular regulatory systems, including DNA damage response, DNA repair, cell cycle regulation, and protein degradation. This work focuses on modeling three systems with specific regulatory roles for ubiquitin.

- Ubiquitin-dependent regulation of the yeast uracil transporter, Fur4 (Chapter 2)

- Ubiquitin-dependent endosomal protein sorting mediated by ESCRTs (Chapters 3 and 5)

- Ubiquitin-depedent regulation of Rad18 in the DNA damage tolerance pathway (Chapter 6)

The first two systems involve both cytosolic and membrane-bound proteins. Our work on ESCRT protein sorting revealed a need for improved mathematical understanding of how protein interactions depend on dimensionality. Others have similarly noticed this need [50, 46]. Chapter 4 presents our work on understanding the role dimensionality plays in protein-protein interactions. This mathematical theory is employed in Chapter 5 to better understand ESCRT-mediated protein sorting. As a whole, this dissertation provides insight into three distinct regulatory roles for ubiquitin and suggests that dimensionality has a significant effect on protein interactions and, therefore, protein regulation.

CHAPTER 2

FUR4 REGULATION

2.1   MVB Pathway Overview

Transmembrane proteins, including signaling receptors and transporters, often need to be deleted due to overproduction, damage, or in response to current cellular conditions. A ubiquitin tag acts as a deletion signal for a wide number of transmembrane proteins targeting them for degradation via the lysosome. The uracil permease, Fur4, is one example of a protein which is degraded in this ubiquitin-dependent manner [11, 22]. Vesicles containing ubiquitinated proteins are endocytosed and targeted to early endosomes. Early endosomes are heavily involved in protein trafficking. While ubiquitin is a necessary tag for proteins to be degraded in this pathway, proteins lacking ubiquitin tags are also trafficked through early endosomes and may be recycled to the plasma membrane, transported to the golgi, or targeted to the membrane of the lysosome. GLUT4 is an example of a protein which is trafficked through endosomes while lacking a ubiquitin tag [17]. As the endosome matures, ubiquitin-tagged proteins are packaged into intraluminal vesicles (ILVs) forming mature endosomes or multivesicular bodies (MVBs). As part of the packaging process, a deubiquitinating enzyme, Doa4, removes ubiquitin tags so that ubiquitin is recycled back to the cytoplasm rather than degraded. The endosomal sorting complexes required for transport (ESCRTs) are implicated in the sorting and packaging of cargoes into ILVs. We discuss their role in this system in Chapter 3. MVBs ultimately fuse with the lysosome, emptying the ILVs into the lumen of the lysosome where they are degraded. A simple overview of this system is shown in Figure 2.1.

The ubiquitin-tagged state of endosomal transmembrane proteins is dynamic as both ubiquitinating and deubiquitinating enzymes are localized to early endosomes. Since ubiquitin tagging is necessary for protein degradation, the interplay between these enzymes is important for regulation. A protein complex has been described on the early endosome which includes both the E3 ubiquitin ligase, Rsp5, and the deubiquitinating enzyme, Ubp2 [21, 36, 25]. Naturally, questions have arisen as to why a complex exists with two enzymes
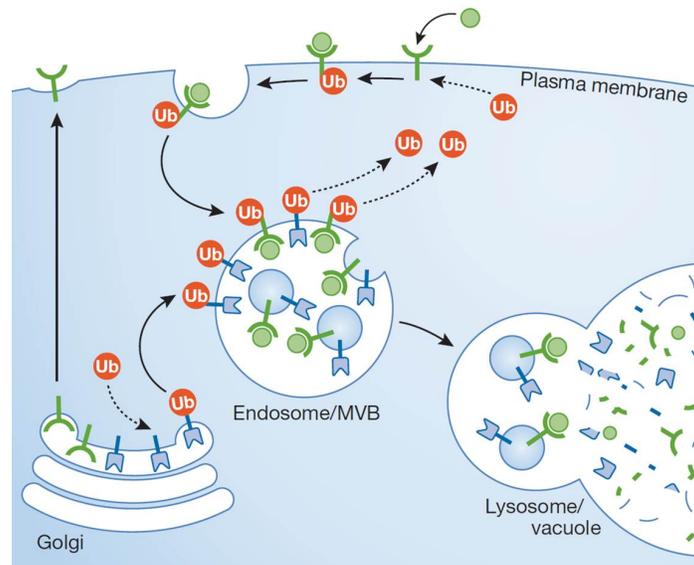
Figure 2.1: Overview of the MVB pathway. Reprinted with permission from [34].

performing opposing functions. Experimental observations suggest a more complex role for the Rsp5/Ubp2 complex than that of merely ubiquitin tagging and untagging. It has been observed that deletion of Ubp2 leads to increased protein recycling to the plasma membrane [25]. This opposes intuition which suggests that deleting the deubiquitinating enzyme would increase the amount of ubiquitinated cargo, leading to increased protein degradation. In contrast to these observations, our collaborators have recently observed that mutating only the active site of Ubp2, rather than deleting Ubp2 entirely, led to increased protein degradation (unpublished data). This suggests a further regulatory role for Ubp2. We hypothesize that the Rsp5/Ubp2 complex serves also as a retention complex, binding to ubiquitin of tagged proteins and retaining them on the endosome, leading to inclusion in ILVs. Under this hypothesis, deubiquitination is impaired under both deletion and mutation of Ubp2. However, deletion of Ubp2 also disrupts the function of the entire Rsp5/Ubp2 complex, leading to less retention of cargo on the endosome and more recycling to the plasma membrane while mutation of only the Ubp2 active site allows for the complex to remain intact thereby preserving its ability to retain cargo, leading to an increase in protein degradation. This hypothesis also suggests that recycling to the plasma membrane is the default pathway for all proteins on the early endosome and it is only through interaction with the Rsp5/Ubp2 complex that proteins are retained during endosomal maturation. This complex includes more players that could help facilitate endosomal retention. Rup1 interacts with both Rsp5 and Ubp2 in this complex and contains a ubiquitin associating

domain (UBA) which may facilitate the retention of cargo [21]. Rsp5 and Ubp2 also possess the ability to interact with ubiquitin which is unsurprising given their enzymatic roles. Hua1 also associates with this complex and has the ability to bind Hse1, a domain in ESCRT-0. [36] This gives the Rsp5/Ubp2 complex the ability to interact with the ESCRTs upon endosomal maturation. ESCRT interaction with cargo is necessary to achieve proper degradation of cargo.

Another potential function of the Rsp5/Ubp2 complex is quality control of transmembrane proteins. It is of upmost importance that cells maintain the integrity of their receptors and transporters at the plasma membrane. Further, the quantity of receptors and transporters at the plasma membrane must be tightly controlled. The plasma membrane appears to be deficient of deubiquitinating enzymes but rich in ubiquitinating enzymes. Therefore, once a protein is tagged with ubiquitin it will certainly be internalized and targeted to endosomes. This seems appropriate in order to efficiently remove any protein from the plasma membrane which is incorrectly folded or not needed. Upon arrival at the endosomme, a tagged protein is able to interact with the Rsp5/Ubp2 complex. There is then a competition between ubiquitination and deubiqutination at the endosome. As ubiquitin tags are removed by Ubp2, the cell is provided with a chance to "double check" whether or not a protein should be degraded. Misfolded proteins will quickly be tagged again by Rsp5 while proteins that may have been improperly tagged will not receive another tag and will be recycled back to the plasma membrane. In this way, cells are able to avoid over-degradation and expending unnecessary energy to produce unneeded proteins. In this functional model of the Rsp5/Ubp2 complex, the relative rates of ubiquitination and deubiquitination become important as do the mechanisms by which individual proteins are able to be tagged. While not considered in our present model, polyubiquitination is likely to prove important in this process. With the proposed roles of the Rsp5/Ubp2 complex in mind, we now refer to the Rsp5/Ubp2 complex as the endosomal quality control and retention system (EQRS). A schematic of the proposed functional model of the EQRS is shown in Figure 2.2. We now explore the consequences of this complex further in the context of Fur4 regulation.

2.2   Fur4 Specific Regulation

The yeast uracil permease, Fur4, is one of many proteins which is degraded via the MVB pathway in a ubiquitin-dependent manner. Here we seek a better understanding of the mechanisms underlying Fur4 regulation and the function of the EQRS in relation to Fur4. Fur4 is a bidirectional cotransporter of both uracil and protons on the plasma membrane.
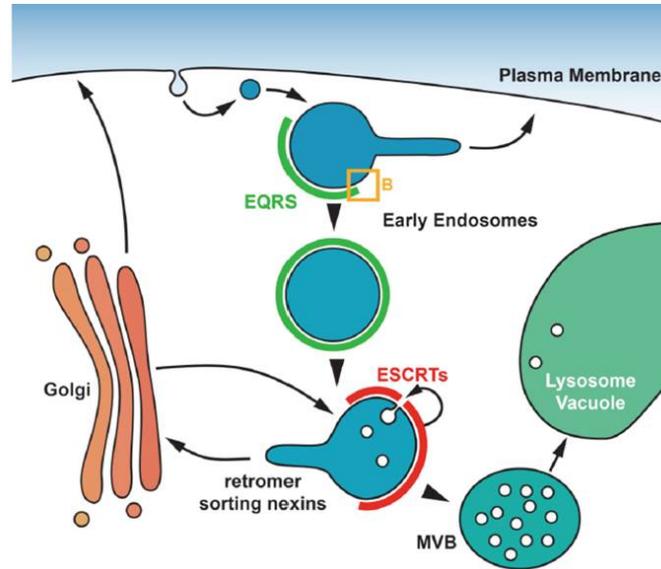
Figure 2.2: Endosomal Quality Control and Retention System (EQRS). Figure courtesy of Markus Babst.

The difference in intracellular pH and extracellular pH account for an approximately 32-fold increase in internal uracil uptake over external uracil excretion. As uracil is an RNA nucleobase crucial for cell function, it is important that the cell uptakes available uracil without reaching toxic levels. Uracil is fairly scarce in typical cellular environments; however, it is very beneficial for a cell to be able to use available uracil from its surroundings when possible. When extracellular uracil is not available, cells possess pathways for synthesizing uridine monophosphate (UMP) without uracil uptake [39]. It has been shown that Fur4 is downregulated due to increases in both extracellular and intracellular uracil [6], thereby keeping the intracellular uracil in check even in an excess of uracil in the environment. Similarly, increasing amounts of environmental stress also downregulates Fur4 in the cell [8]. We seek to further understand the mechanisms underlying this regulation. Our collaborators recently elucidated the mechanism by which Fur4 is able to be ubiquitinated [22]. A LID (loop-interacting-domain) was shown to be present on the cytosolic side of Fur4. Upon damage or binding to uracil, the LID undergoes a conformational change exposing a lysine residue whereby Fur4 may be ubiquitinated at the plasma membrane by Rsp5. This process is depicted in Figure 2.3. Ubiquitin tagged Fur4 in the plasma membrane is then transported to the endosome as previously described.

Once ubiquitinated Fur4 reaches the endosome, it is able to interact with the EQRS likely leading to deubiquitination by Ubp2. If intracellular uracil levels are high, the LID will remain open as uracil readily binds to Fur4. Due to the internal pH of endosomes, it is
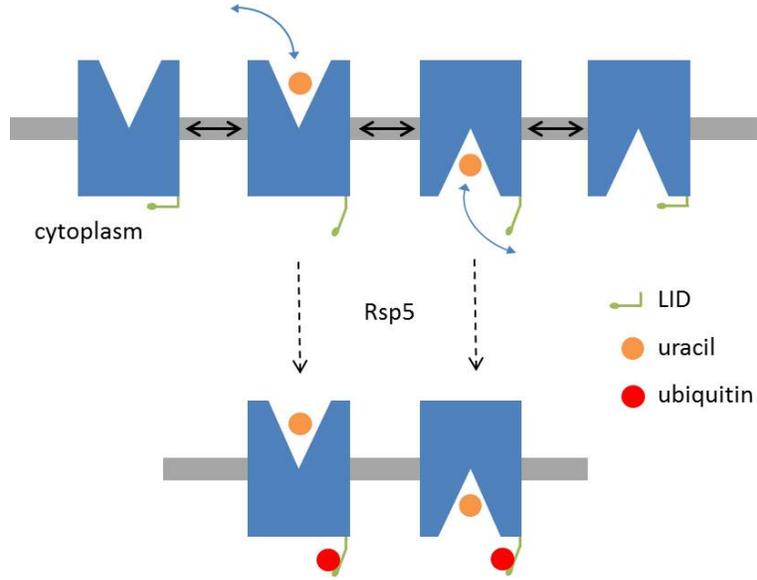
Figure 2.3: Mechanism of Fur4 uracil induced ubiquitination. Adapted from [22].

highly unlikely that any uracil will be transported into the endosome. The pH discrepancy does not prevent uracil from binding to the cytoplasmic side of Fur4 which leads to the opening of the LID. With the LID open, Fur4 is able to be ubiquitinated by Rsp5. On the other hand, if intracellular uracil is low, Fur4 on the endosome will have less uracil to bind to leading to the LID being closed more often. This will prevent ubiquitination of Fur4 by Rsp5. As a feature of the proposed EQRS, a ubiquitin tag is necessary for retention on the endosome. Therefore, if there is less ubiquitinated Fur4 on the endosome, there will be less retention, leading to more recycling and less degradation. We propose that the EQRS plays a vital role in Fur4 regulation by allowing the cell to respond to intracellular levels of uracil and prevent over-degradation of Fur4. To better understand the effects of the EQRS and the general behavior of this system, we analyze the following ODE model.

2.3   ODE Model

We model the regulation of Fur4 using a system of 15 differential equations. Our model includes 14 different states for Fur4. We need to keep track of where Fur4 is located in the cell, whether or not it is bound to uracil, and if it is tagged with ubiquitin. We refer to the uracil binding state of Fur4 as bound or unbound while we call the ubiquitination state of Fur4 tagged or untagged. For the plasma membrane states, we also specify the orientation of Fur4, opened inward (toward the cytoplasm) or outward. The endosomal Fur4 states do not include orientation as Fur4 is highly unlikely to be opened toward the lumen of the

endosome due to pH differences. The states for Fur4 are given in Table 2.1. The final variable in our model is intracellular uracil, S. We assume a fixed amount of extracellular uracil, S_e.

In yeast cells, there are approximately 5-10 early endosomes at one time. We model the endosome as one compartment, rather than modeling individual endosomes.

2.3.1 Fur4 State Transitions

We now discuss the different events that determine the transitions between Fur4 states in our model.

- **Uracil transport**

  In the plasma membrane, Fur4 exists in two orientations either facing the cytoplasm or facing the extracellular matrix. We assume that regardless of binding or tagging states, Fur4 flips between these two positions at a constant rate, r. Uracil is cotransported with protons and the transport is bidirectional. However, the pH difference between the inside and the outside of the cell account for a 32-fold difference in proton concentration. We account for this using an orientation-dependent scaling of the on-rate of uracil binding to Fur4, w. Therefore, in our model, uracil is 32 times as likely to bind to Fur4 from the outside of the cell than it is to bind from the inside of the cell. This results in an on-rate of k for extracellular uracil binding and $\frac{k}{w}$ for intracellular binding. We assume a constant off-rate of bound Fur4, j, regardless of orientation. Experimental measurements suggest the K_d for Fur4 is $\sim .74\mu$M and the

Table 2.1: Fur4 states

P^e	plasma membrane, unbound, untagged, outward-facing
Pb^e	plasma membrane, bound, untagged, outward-facing
P^i	plasma membrane, unbound, untagged, inward-facing
Pb^i	plasma membrane, bound, untagged, inward-facing
Pb^e_u	plasma membrane, bound, tagged, outward-facing
Pu^e	plasma membrane, unbound, tagged, outward-facing
Pb^i_u	plasma membrane, bound, tagged, inward-facing
Pu^i	plasma membrane, unbound, tagged, inward-facing
E	endosomal membrane, unbound, untagged
Eb	endosomal membrane, bound, untagged
Eb_u	endosomal membrane, bound, tagged
E_u	endosomal membrane, unbound, tagged
X	endosomal membrane, unbound, tagged, retained by EQRS
Xb	endosomal membrane, bound, tagged, retained by EQRS

V_{max} for uracil uptake is $\sim 8.8(\text{nmol}\cdot\text{min}^{-1}10^7\text{cells}^{-1})$ [45]. We therefore take $k = \frac{j}{K_d}$ and compute the fliprate r from the observed V_{max} as shown in Appendix A.

- **Ubiquitination and Deubiquitination**

  As depicted in Figure 2.3, ubiquitination of Fur4 by Rsp5 is possible if the LID is open. This happens when Fur4 is bound to uracil. Therefore, bound Fur4 is ubiquitinated at a rate, a. We allow for a small amount of uracil-independent ubiquitination with rate, q. This accounts for a small amount of damaged transporters or for thermal fluctuations in Fur4 leading to faulty opening of the LID. Deubiquitination by Ubp2 occurs at the endosome at a rate, b.

- **Transport**

  Vesicle transport is responsible for transitions between plasma membrane and endosomal membrane states. The rate g represents the rate of transport from the plasma membrane to the endosome and the rate f accounts for the rate of recycling. As the plasma membrane and the endosomal membrane have different surface areas, we must scale the transport terms appropriately in our equations to obtain appropriate concentrations. The yeast cells studied by our collaborators have an approximate diameter of 5μm leading to a surface area $A_p = 314\mu\text{m}^2$. Early endosomes have an approximate average diameter of 0.5μm. With approximately 10-20 early endosomes per cell, we estimate the total surface area of early endosomes to be $A_e = 47\mu\text{m}^2$.

- **Retention**

  We model the retention mediated by interaction with the EQRS as two separate Fur4 states, X and Xb. We do not track the amount of EQRS components in our model as presumably there is an abundance of Rsp5, Ubp2, Rup1, and Hua1 to form the EQRS. We assume that ubiquitinated Fur4 on the endosome bind to the EQRS at a rate, c, and unbind at a rate, d.

- **Production and Degradation**

  We assume that Fur4 is constitutively produced at a rate, y. Our collaborators have estimated that the rate of production of Fur4 is ~ 5 Fur4 per minute. As early endosomes are constantly maturing into MVBs leading to degradation, we include a constant rate of matruation/degradation, z.

The above transitions in our model are depicted in Figure 2.4.
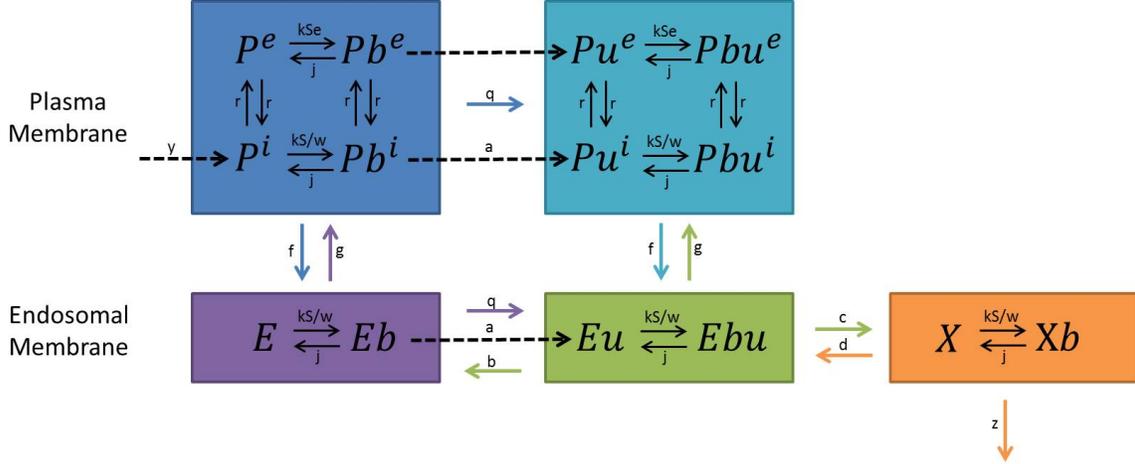
Figure 2.4: Fur4 state diagram

2.3.2 Equations

We now provide the equations corresponding the above descriptions and Figure 2.4. The equations modeling Fur4 at the plasma membrane are

$$\frac{dP^e}{dt} = y - kS_eP^e + jPb^e - fP^e - rP^e + rP^i - qP^e, \tag{2.1}$$

$$\frac{dPb^e}{dt} = kS_eP^e - jPb^e - rPb^e + rPb^i - qPb^e, \tag{2.2}$$

$$\frac{dP^i}{dt} = -\frac{k}{w}SP^i + jPb^i - rP^i + rP^e + \frac{A_e}{A_p}gE - qP^i, \tag{2.3}$$

$$\frac{dPb^i}{dt} = \frac{k}{w}SP^i - jPb^i - rPb^i + rPb^e + \frac{A_e}{A_p}gEb - qPb^i - aPb^i, \tag{2.4}$$

$$\frac{dPb_u^e}{dt} = kS_eP_u^e - jPb_u^e - rPb_u^e + rPb_u^i + qPb^e - fPb_u^e, \tag{2.5}$$

$$\frac{dPu^e}{dt} = -kS_ePu^e + jPb_u^e - rPu^e + rPu^i + qP^e - fPu^e, \tag{2.6}$$

$$\frac{dPb_u^i}{dt} = \frac{k}{w}SP_u^i - jPb_u^i - rPb_u^i + rPb_u^e + qPb^i \tag{2.7}$$
$$+aPb^i - fPb_u^i + \frac{A_e}{A_p}gEb_u,$$

$$\frac{dP_u^i}{dt} = -\frac{k}{w}SP_u^i + jPb_u^i - rP_u^i + rP_u^e + qP^i - fP_u^i + \frac{A_e}{A_p}gE_u. \tag{2.8}$$

The equations modeling Fur4 at the endosomal membrane are

$$\frac{dE}{dt} = -\frac{k}{w}SE + jEb - gE - qE + bE_u, \tag{2.9}$$

$$\frac{dEb}{dt} = \frac{k}{w}SE - jEb - gEb - qEb - aEb + bEb_u, \tag{2.10}$$

$$\frac{dEb_u}{dt} = \frac{k}{w}SE_u - jEb_u - gEb_u + \frac{A_p}{A_e}fPb_u^i \tag{2.11}$$

$$+ \frac{A_p}{A_e} f P b_u^e + qEb + aEb - bEb_u - cEb_u + dXb,$$

$$\frac{dE_u}{dt} = \frac{k}{w} S E_u + j E b_u - g E_u + \frac{A_p}{A_e} f P_u^i + qE - bE_u - cE_u + dX, \qquad (2.12)$$

$$\frac{dX}{dt} = cE_u - dX - zX, \qquad (2.13)$$

$$\frac{dXb}{dt} = cEb_u - dXb - zXb. \qquad (2.14)$$

The final equation in our system tracks the concentration of intracellular uracil. To appropriately model uracil concentration, we must include surface area to cytoplasm ratios to preserve correct units. Assuming the cell has a radius of 5μm, we take the cytoplasmic volume to be $V = 523\mu$m^3. Uracil is metabolized in the cell by enzymes such as Fur1 which catalyzes its transformation into UMP [23]. We therefore model uracil metabolism using a Mechaelis-Menten reaction rate. As the coefficients for this term are not known, we use the reported values of $K_m \sim 2.5\mu$M and $V_{max} \sim 8.8$(nmol $\cdot$ min$^{-1}10^7$cells^{-1}) for uracil uptake by Fur4 [45]. The equation for intracellular uracil is

$$\frac{dS}{dt} = -\frac{k}{w} S \left(\frac{A_p}{V}(P^i + P_u^i) + \frac{A_e}{V}(E + Eu) \right) + j \left(\frac{A_p}{V}(Pb^i + Pb_u^i) \qquad (2.15)$$
$$+ \frac{A_e}{V}(Eb + Eb_u) \right) + z \frac{A_e}{V} Xb - \frac{V_{max}S}{V(K_m + S)}.$$

2.3.3 Parameters

For the parameters not explicitly mentioned in the above descriptions of state transitions, we use order of magnitude estimates determined with the help of our collaborators. The base set of parameters used is given in Table 2.2. We have scaled the parameters so that time is in units of milliseconds, uracil concentrations are in units of μM, and Fur4 concentrations are in areal concentrations (number per unit area).

2.3.4 Steady State Analysis

We wish to analyze this model at steady state and assess the affects of changing extracellular uracil. We are able to analytically solve for the steady states of our 14 Fur4 variables as functions of S as they are each linear functions of S. This leaves us with a single equation for the steady state of S which we are unable to solve analytically. We, therefore, use a bisection method to numerically determine the positive steady state of S between $S = 0$ and $S = wS_e$. Given the parameters in Table 2.2, we compute the steady state of our system for a range of S_e values. In all the results that follow, we plot Fur4 as total number of Fur4 in the cell rather than in units of areal concentration.

Table 2.2: Parameters

a	1	rate of ubiquitination
b	1	rate of deubiquitination
q	.01	basal rate of ubiquitination
K_d	.74	dissociation constant for uracil/Fur4
k	$\frac{j}{K_d}$	binding rate of uracil/Fur4
w	32	scale factor for pH
j	10^2	unbinding rate of uracil/Fur4
r	17.915	flip rate of Fur4
f	.25	transport rate of endocytosis
g	.1	transport rate of recycling
c	1	binding rate of Fur4/EQRS
d	.5	unbinding rate of Fur4/EQRS
y	.000083	production rate of Fur4
z	.002	degradation rate of Fur4
Ae	47	surface area of endosomes
Ap	314	surface area of plasma membrane
V	523	volume of cytoplasm
V_{max}	$8.8 \cdot 10^3$	maximal rate of uracil uptake
K_m	2.5	Michaelis-Menten constant for uracil uptake

2.4   Results

2.4.1   Effects of Changing Extracellular Uracil

Importantly, our system reproduces the well-documented behavior that Fur4 is down-regulated due to increasing amounts of extracellular uracil (Figure 2.5A) suggesting that the inclusion of LID-mediated ubiquitin tagging and the EQRS is sufficient to reproduce this behavior. Interestingly, our results suggest that most of the Fur4 downregulation results from a decrease in plasma membrane Fur4. Our model predicts that the pool of endosomal Fur4 remains almost constant as extracellular S_e varies (Figure 2.5B). With a nearly constant endosomal pool and plasma membrane Fur4 decreasing, we observe a switch in the localization of Fur4 in the cell as S_e increases. At low levels of S_e, there is not only

more Fur4 but the majority is located in the plasma membrane. As S_e increases, more of the total amount of Fur4 in the cell is located at the endosome. This behavior is observed in Figure 2.5C. The severity of this switch appears to result from the ratio of the transport parameters f and g. If recycling is far slower than endocytosis, one would expect far more Fur4 to localize to the endosome under conditions with high extracellular uracil. In the case where the rates are equal, this shift still occurs, however, it is slightly less dramatic.

As seen in Figure 2.6, our ODE model predicts that our functional model of Fur4 regulation as hypothesized here is indeed enough to prevent the cell from reaching toxic levels of intracellular uracil. Even under extremely high extracellular conditions, the cell maintains a much lower level of intracellular uracil despite having a 32-fold increase in transport from outside to in driven by the pH discrepancy. This qualitative behavior is seen over a wide range of parameter values. However, the quantitative levels change in response to a changes in the parameters j, r, V_{max}, and K_m representing uracil transport and metabolism. Obtaining better experimental measurements of these rates would prove useful in predicting the actual amount of intracellular uracil present as S_e increases.

2.4.2   Roles of the EQRS

One goal of our model is to better understand the role of the EQRS in Fur4 regulation. We propose three functions for this complex in our system: ubiquitination, deubiquitination, and retention. While the role of ubiquitination at the endosome seems intuitive to prevent misfolded or unneeded Fur4 from recycling, the other two proposed roles of the EQRS are not well understood. To address the role of deubiquitination, we compare the steady state behavior shown in Figure 2.5 to the steady states observed with the deubiquitination rate, b, set to 0 as shown in Figure 2.7.

Even without deubiquitination, we observe a decrease in total Fur4 as S_e increases. We again observe a near constant pool of Fur4 on the endosome. However, we do not observe a switch in Fur4 localization as S_e increases. Most of the total Fur4 is located at the endosome for any given extracellular uracil concentration. The most drastic differences seen, however, are with the total numbers of Fur4 in the cell and the amount of downregulation as S_e increases. Without deubiquitination, our model predicts that there are very low levels of Fur4 in the cell, $\sim$ 31 Fur4. In addition, without deubiquitination, the model shows this number is only decreased to $\sim$ 28 Fur4 as S_e varies from 0 to 5 μM which gives approximately 10% downregulation. With such a small amount of Fur4, it is not likely that a 10% change will be easily observed experimentally. In contrast, when deubiquitination functions, we see
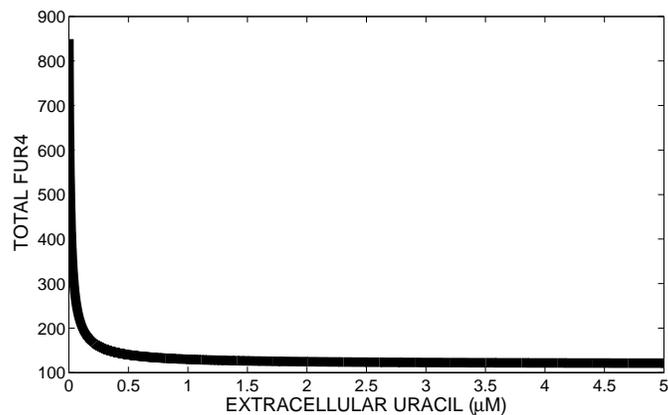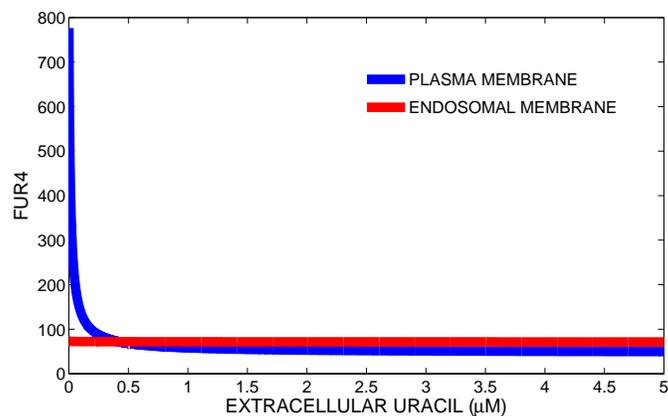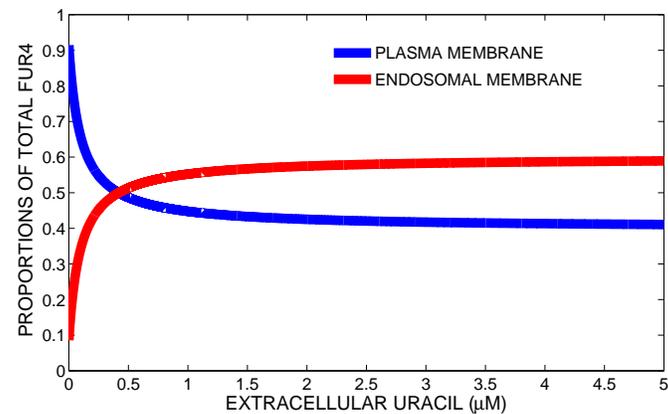
Figure 2.5: Steady states of Fur4 as S_e varies with parameters given in Table 2.2. (a) Total number of Fur4 in the cell. (b) Number of Fur4 localized to the plasma membrane and the endosomal membrane. (c) Proportion of total Fur4 localized to the plasma membrane and the endosomal membrane.
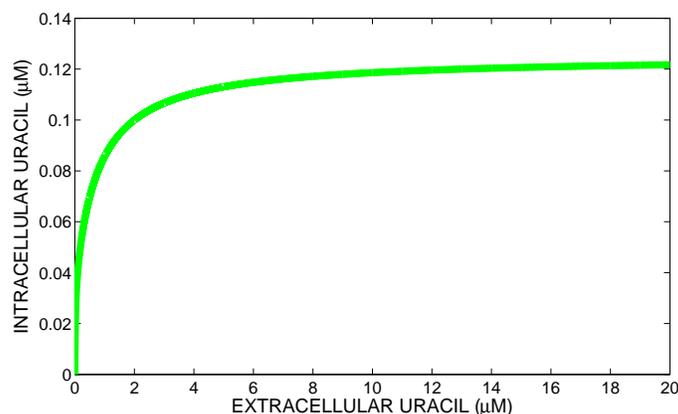
Figure 2.6: Steady state of intracellular uracil as S_e varies with parameters given in Table 2.2.

an almost 90% downregulation with ~ 850 Fur4 when no extracellular uracil is present to ~ 100 Fur4 with $S_e = 5\mu M$. The comparison of downregulation is shown in Figure 2.8A. These results suggest that without deubiquitination, the cell degrades almost all of its Fur4 even with low extracellular uracil. Without deubiquitination, intracellular uracil levels are kept extremely low over a wide range of extracellular uracil levels (data not shown), but this is due to the lack of transporter in the plasma membrane. Figure 2.8B shows that deubiquitination allows for a ten-fold increase in uracil uptake. From these data, we suggest that deubiquitination of Fur4 at the endosome is indeed necessary in this system. Without deubiquitination, the cell would essentially be producing Fur4 only to degrade it, which is highly inefficient, leads to poor uracil uptake, and is energetically unfavorable.

The results presented in Figures 2.7 and 2.8 mimic mutating the active site of Ubp2. Indeed we see much less Fur4 without deubiquitination, as was seen experimentally. The other, seemingly contradictory, experimental result occurred from Ubp2 deletion. In this case, experimental evidence suggests Fur4 is stabilized on the membrane. We hypothesize that this is due to the retention function of the EQRS complex. Under this hypothesis, deletion of Ubp2 would not only impair deubiquitination but would prevent ubiquitinated cargo from remaining on the endosome until it is included in ILVs. We consider the effects of Ubp2 deletion in our model by setting $b = 0$ and varying c, the rate of binding to the EQRS, for a fixed extracellular uracil concentration. These results are shown in Figure 2.9. We see that under the same extracellular uracil conditions, impairing the ability to bind to the EQRS dramatically increased the amount of Fur4 in the cell. Our model with the proposed functions of the EQRS is able to reproduce both experimental results.
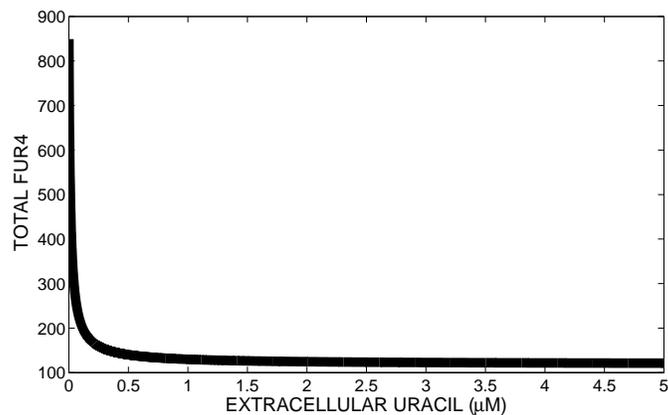
(a)
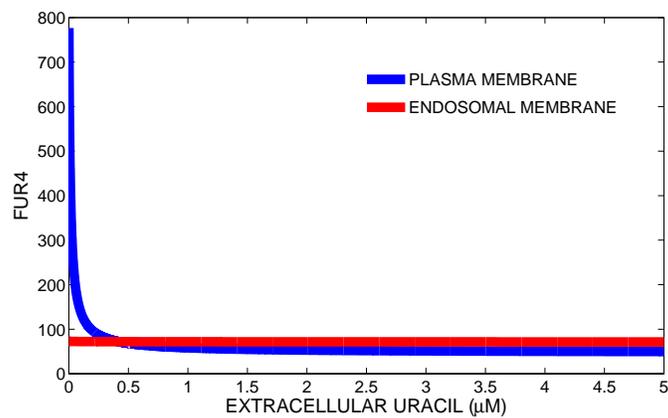
(b)
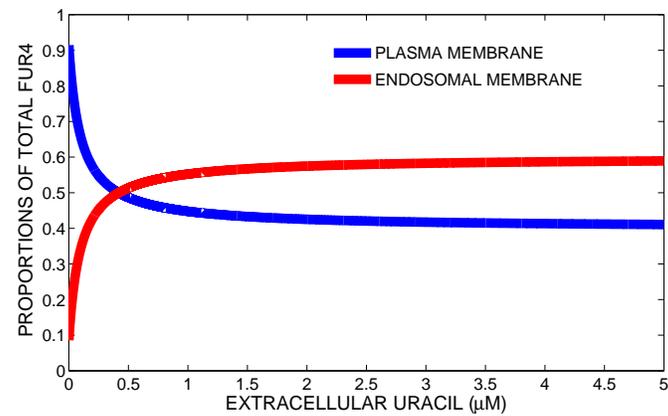
(c)

Figure 2.7: Steady states of Fur4 as S_e varies with $b = 0$ and all other parameters given in Table 2.2. (a) Total number of Fur4 in the cell. (b) Number of Fur4 localized to the plasma membrane and the endosomal membrane. (c) Proportion of total Fur4 localized to the plasma membrane and the endosomal membrane.
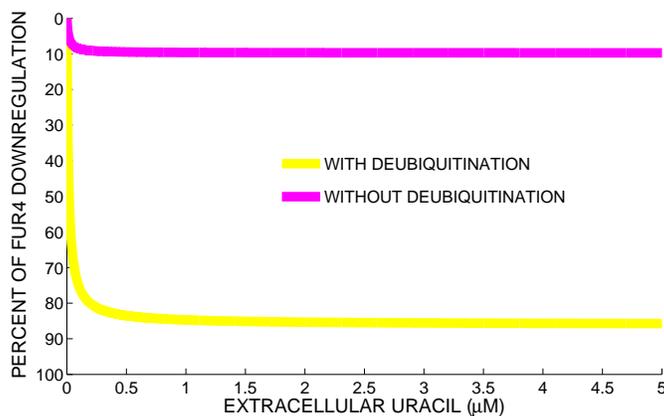
(a)
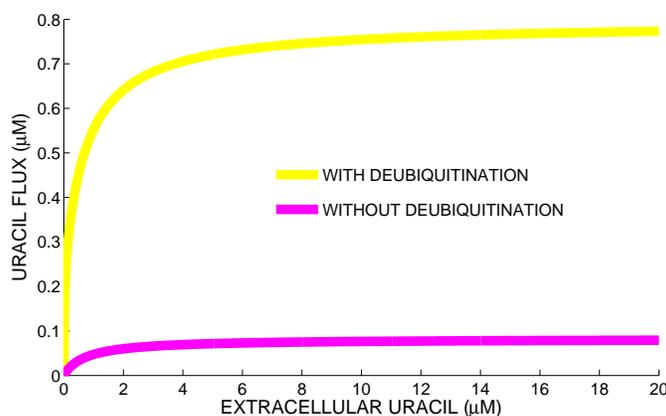
(b)

Figure 2.8: Effects of deubiquitination. (a) Percent of downregulation with and without deubiquitination, $b = 1$ and $b = 0$, respectively. (b) Uracil flux across the plasma membrane with and without deubiquitination, $b = 1$ and $b = 0$, respectively.

2.4.3 Endosomal Membrane Fur4

To gain better insight into the function of the observed pool of endosomal Fur4, we use our model to perform a time course experiment. We set the extracellular uracil to 2μM and run the ODE to steady state. We then change the extracellular uracil to 0.5μM to observe the transient behavior before reaching a new steady state. Figure 2.10 shows that upon a sudden decrease in extracellular uracil, the cell quickly moves Fur4 from the endosomal pool to the plasma membrane. Having a pool of extracellular uracil allows the cell to very quickly respond to decreasing extracellular uracil. This seems reasonable as the cell would benefit from having more Fur4 in the plasma membrane soon after the environmental change in order to uptake as much uracil as it can under starvation conditions.
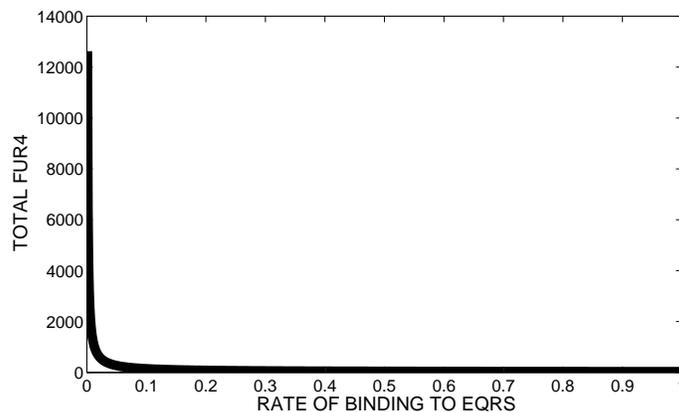
Figure 2.9: Effects of losing the retention function of the EQRS with no deubiquitination with fixed extracellular uracil, $S_e = .1$
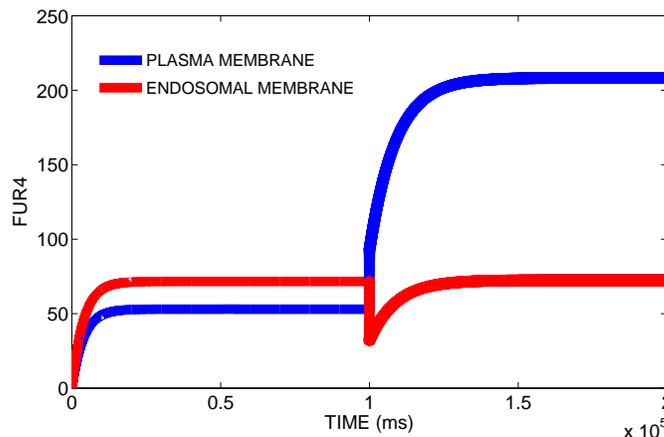
Figure 2.10: Time course experiment for rapid change in extracellular uracil.

2.5   Discussion

The goal of our model was to gain insight into how ubiquitin-dependent regulation of Fur4 is achieved. We suggest that ubiquitin is involved in three distinct events in this system. First, Fur4 is able to be ubiquitin tagged by Rsp5 through the mechanisms elucidated by Keener et al. in [22]. Second, Fur4 may lose its ubiquitin tag at the endosome by interaction with Ubp2. Third, ubiquitinated Fur4 may interact with the EQRS and thereby be retained at the endosome.

By including these three roles for ubiquitin (or lack thereof in the case of deubiquitination), we were able to create a model that reproduces the known experimental results in situ. Our results show downregulation of Fur4 in response to increases in extracellular uracil which is well-known experimentally. What has not been clear experimentally is the

regulatory mechanisms of Fur4 at the early endosome. Our results suggest that deubiq-uitination at the endosome is crucial for this regulatory system to function properly as is retention. Without deubiquitination, we do not see marked downregulation of Fur4 in response to increasing extracellular uracil in our model. Without deubiquitination on the endosome, the cell is rapidly degrading Fur4 under all extracellular conditions which is both energetically unfavorable and unlikely as it prevents efficient uracil uptake.

With our proposed roles for the EQRS, we are also able to make sense of the seemingly contradictory experimental results that have been observed. We argue that deletion of Ubp2 causes impaired retention function and impaired deubiquitination while mutating the Ubp2 active site only causes impaired deubiquitination. Removing deubiquitination results in very low levels of Fur4 while decreasing retention function and removing deubiquitination results in increased levels of Fur4 in the cell mimicking the experimental results.

Surprisingly, our results predict an almost constant pool of endosomal Fur4 over a wide range of extracellular conditions. Using a time course experiment, our model suggests that upon a switch to starvation conditions, the cell will quickly move some of this endosomal Fur4 pool to the plasma membrane giving the cell a better chance for maintaining uracil uptake. It will be interesting to see if this prediction holds in the lab.

As a whole, our work suggests that the EQRS plays a very important role in ubiquitin-dependent Fur4 regulation.

CHAPTER 3

ESCRT-MEDIATED PROTEIN SORTING

3.1 ESCRT Overview

In this chapter, we focus on the mechanism by which ubiquitinated proteins are packaged for inclusion into ILVs in the MVB pathway. This process is downstream from the events involving the early endosome we discussed in Chapter 3. Endosomal Sorting Complexes Required for Transport (ESCRTs) are responsible for sorting ubiquitinated proteins (cargoes) from untagged proteins (non-cargoes) on the membrane of late endosomes and packaging them for inclusion into ILVs. Since their discovery nearly 15 years ago, the ESCRT literature has been quickly growing. In 2010, Hurley reviewed the current state of ESCRT knowledge and asserted that this may be the last time such a review would be possible with the knowledge base expanding so quickly [19]. While much has been discovered about the structure and function of the individual ESCRT proteins, it still remains unclear as to how they assemble on the endosomal membrane and facilitate protein sorting during the process of ILV formation. In this work, we seek to better understand the process of ESCRT-mediated protein sorting.

The ESCRT proteins assemble on the endosomal membrane through interactions with each other, the lipid phosphatidylinositol 3-phosphate (PI(3)P), and cargo mediated by ubiquitin [19, 49]. There are four distinct ESCRT proteins: ESCRT-0, ESCRT-I, ESCRT-II, and ESCRT-III. At least 11 different ubiquitin binding domains have been discovered among the ESCRT-0, -I, and -II, all of which have low binding affinities [42]. There is experimental evidence that the ESCRT components bind in a linear fashion, i.e., ESCRT-0 binds to ESCRT-I, ESCRT-I binds to ESCRT-II, and ESCRT-II binds to ESCRT-III (Figure 3.1). This linear order of ESCRT binding does not, however, necessitate a linear geometry for ESCRT complex structures on the endosomal membrane as individual ESCRTs are reported to have stable cores while being flexibly connected [48]. Upon ILV formation, the ESCRTs are released back into the cytoplasm while cargo is included in ILVs.

Questions remain as to how sorting and packaging of cargo are achieved. Various
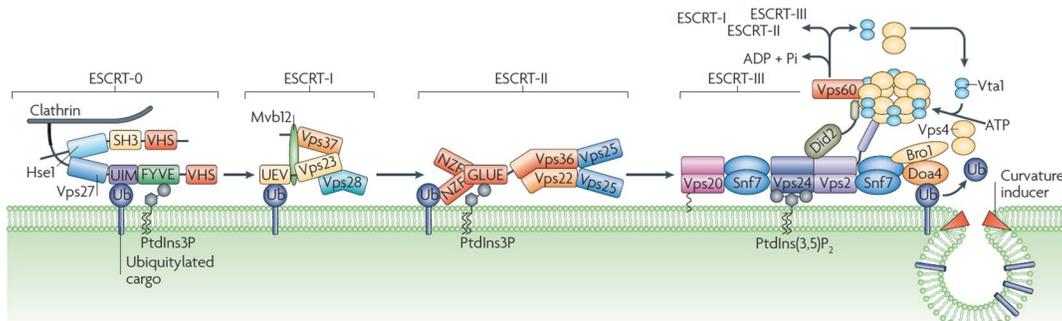
Figure 3.1: The ESCRT complexes. Reprinted with permission from [48]

hypotheses for functional models of sorting have previously been described. In 2002, Babst et al. presented the first functional sorting model suggesting that ESCRT-I binds to cargo leading to recruitment of ESCRT-II followed by ESCRT-III. ESCRT-III was then proposed to be the main player in sorting and packaging cargo into ILVs [2]. Subsequent research has shown that ESCRT-III does not contain ubiquitin binding domains and that ESCRT-0 plays a crucial role in this process. In 2007, Nickerson et al. presented a concentric circle model for ESCRT sorting with a strict arrangement of the ESCRT proteins centered around an ESCRT-0 hub [34]. While this model meets known expermental data, it is not immediately clear how cargo is concentrated into vesicles in this model. As cargo are packaged into vesicles, a circular model is appealing. Without specifying the exact arrangements of the ESCRTs, we previously hypothesized that the ESCRTs polymerize into a ring-like structure with cargo initially outside the ring. Cargo would then need to move through the ESCRT ring for inclusion into ILVs. We hypothesized that Doa4 may be localized to the interior of this ring, thereby removing the ubiquitin tag preventing the cargo from leaving once inside. A simple cellular automata model revealed that this model did not seem possible without far-reaching assumptions not yet founded in experimental evidence. In 2009, Shields et al. suggested yet another functional model for protein sorting with ESCRT-0, -I, and -II each recognizing a different subset of ubiquitinated cargo in order to facilitate the sorting of a wide variety of ubiquitinated transmembrane proteins [40]. Even more hypotheses for ESCRT sorting have been proposed in addition to those already mentioned [7, 43, 30]. It is clear from the wide range of hypotheses, from those considered experts in the ESCRT field, that sorting and packaging into ILVs is not well understood.

Recent data from our collaborators provide new insight into ESCRT-mediated protein sorting. In the remainder of this chapter, we present an overview of the experimental results from our collaborators and our mathematical work which was published in Traffic in 2013,

"Binding to Any ESCRT Can Mediate Ubiquitin-Independent Cargo Sorting" [28]. Figures and text are reproduced in this dissertation with permission.

3.2   Binding to Any ESCRT Can Mediate Ubiquitin-Independent Cargo Sorting

It has been shown that the ESCRT proteins themselves are ubiquitinated. This raises the question as to why the ESCRTs are recycled to the cytoplasm and are not packaged into ILVs if ubiquitin is the only necessary tag for a cargo protein to be sorted and packaged. To better understand the rules that distinguish cargo from ESCRTs, the trafficking of artificial ESCRT-like protein fusions (GFP-FYVE) was analyzed. Interestingly, these studies showed that binding to any ESCRT can mediate cargo sorting. Even ESCRT-III, which lacks the ability to interact with cargo under normal circumstances, was able to mediate inclusion of ESCRT-like protein fusions in ILVs. Further, these protein fusions had the ability to behave as cargo, leading to inclusion in ILVs, and to behave as ESCRTs, recycling back to the cytoplasm. These studies revealed that ESCRT-like proteins with high affinity for the endogenous ESCRT machinery behaved like ESCRTs while ESCRT-like proteins with lower affinity for ESCRT behaved like cargo. These suggest that ESCRTs and their cargoes are very similar. To explain these results, we hypothesize that a flexible network of ESCRT and cargo is sufficient to achieve proper sorting.

We propose a model in which cargo triggers the polymerization of ESCRTs into flexible and unstructured networks of both ESCRTs and cargoes on the endosomal membrane. While the structure of individual ESCRTs and their interactions with each other is well defined, we do not specify an ordered arrangement for ESCRTs within networks. A simplification of this functional model is 2D network polymerization with only two components: ESCRTs and cargoes, where we do not account for the specific details of each ESCRT complex. In order to be a valid hypothesis, this model must be able to efficiently sort cargo while excluding non-cargo. To address this, we have developed a cellular automata model for the polymerization of 2D ESCRT-cargo networks.

For this model we use a 2D hexagonal grid which mimics the endosomal membrane and we allow proteins to move and interact randomly. Our model contains four types of proteins: ESCRTs containing one cargo-binding site and up to five ESCRT-ESCRT interaction sites, cargoes, non-cargoes, and GFP-FYVE (cargo mimics). The cargo-mimic in our model represents the artificial fusions used in the lab and is able to interact with ESCRTs through an ESCRT-ESCRT interaction site. The details of our model are as follows.

3.2.1   Cellular Automata Model

MATLAB was used to run all computer simulations. A cellular automata model was employed on a two-dimensional hexagonal grid. The grid size used in all simulations is 28 by 28. The simulation involves randomly selecting grid points and updating the grid according to a given set of rules for assigning probabilities of updates according to the current state of the grid. Inputs to the simulation include: 1) initial proportions of cargo, non-cargo, and GFP-FYVE construct in the grid, 2) number of ESCRT - ESCRT binding sites, 3) parameters specifying propensities for diffusion, ESCRT - ESCRT binding, ESCRT -cargo binding, ESCRT - GFP-FYVE binding , ESCRT removal, unbinding, and critical vesicle size. Every protein in the grid is associated with a network which includes all the proteins it is connected to through bonds. An unbound protein is simply considered a network of one protein. As proteins bind together, larger networks are formed. Updates may affect the entire network of the grid point selected and not just the individual grid point selected to update. Potential updates are determined by the following rules for each type of update.

3.2.1.1   Rules for ESCRT Complex Formation

1. Recruitment of ESCRTs

ESCRTs are recruited to an empty grid point if there is at least one cargo in an adjacent grid point. The recruited ESCRT is bound to an adjacent cargo.

2. Removal of ESCRTs

Unbound ESCRTs may be removed from the grid. Upon removal, the current grid point becomes empty. Due to our choice of parameters, it is almost certain that an unbound ESCRT will be removed from the grid.

3. Binding

The hexagonal grid allows for up to six bindings per protein. However, cargo and GFP-FYVE are limited to one binding. ESCRTs have one cargo binding site and up to five ESCRT binding sites. The number of allowed ESCRT binding sites is specified as an input. If GFP-FYVE is present in the simulation, it may occupy one of the ESCRT binding sites of an ESCRT. Cargo may bind to one ESCRT. GFP-FYVE may bind to one ESCRT. ESCRTs may bind to one cargo, one GFP-FYVE (if present), and up to five other ESCRTS.

4. Unbinding

Cargo may only unbind if its network is an ESCRT cargo pair. If a cargo is part of a network of three or more proteins, the cargo is prohibited from unbinding. An

ESCRT may unbind if it has only a single bond with another ESCRT and no other interactions. GFP-FYVE is prohibited from unbinding once it is bound to an ESCRT.

5. **Diffusion**

   Networks of any size are allowed to move in any of six directions (up left, down right, up right, down left, left, right) if all of the adjacent grid spaces to the network in that direction are empty. If any adjacent space is occupied by a protein, including adjacent spaces in the interior of the network, movement of the network is prohibited.

6. **Rotation**

   Any protein bound to only one other protein is allowed to rotate clockwise if the adjacent grid space in the clockwise direction is empty. Similarly, a protein bound to only one other protein is allowed to rotate in the counter-clockwise direction.

7. **Removal of networks and trapped non-cargoes**

   As networks grow they may be removed from the system mimicking vesicle formation. An input to the simulation is critical vesicle size, which is defined as a sufficient number of proteins included in a network in order to form a vesicle. We calculate the value $V(N) = \frac{1}{1+e^{-(N-C)}}$ where N is the size of the network defined by the number of proteins in the network and C is the critical vesicle size. As a function of N, V is a sigmoid function centered at C. If the computed value of V is greater than a randomly generated number, the entire network is removed from the grid and every grid point occupied by the network becomes empty. As a network grows it is more likely to be removed from the grid as V is an increasing function. A network rarely grows much larger than the critical vesicle size due to the nature of V(N). In addition, small networks are almost never removed from the grid. It is possible that non-cargo may be trapped in a network which is selected to be removed from the grid. A non-cargo is considered trapped if 1) it is in an interior row of the current network and 2) in each interior row of the network, it is in an interior column of that row. If a non-cargo is trapped when a network is removed, it is also removed with the network and the grid point occupied by that non-cargo becomes empty. This is the only way that non-cargo can be removed from the grid.

3.2.1.2 Determination of Probabilities

The probability of each allowable update for a selected non-empty grid point is determined by the propensity for that update scaled by the sum of propensities for all allowable updates. Therefore, the probabilities of all possible updates sum to one. Further, if only

one update is allowable, that update will be selected with probability one. If an empty grid point is selected, an ESCRT will be recruited with probability one if at least one cargo occupies an adjacent grid point. Otherwise, the empty grid space remains empty.

3.2.1.3   Simulation

To begin each simulation, a random grid is generated based on the initial proportions of cargo, non-cargo, and GFP-FYVE specified. Following the initial grid setup, the simulation repeats as follows until a specified number of runs are completed. Updates to the grid are selected based on the calculated probabilities.

1. Randomly select a grid point.

2. Determine the occupancy type of all the grid points adjacent to the current grid point.

3. Is the current grid point empty?

    - Yes: Move to step 4.

    - No: Move to step 5.

4. Recruit an ESCRT?

    - Yes: Add an ESCRT to the current grid point, bind the ESCRT to an adjacent cargo, and begin again with step 1.

    - No: Begin again with step 1.

5. Remove network of current grid point?

    - Yes: Remove network (along with any trapped non-cargo) and begin again with step 1.

    - No: Move to step 6.

6. Determine what updates are possible for the current grid point or the network of the current grid point.

7. Is there at least one possible update?

    - Yes: Randomly select one of the possible updates and update the grid. Begin again with step 1.

    - No: Begin again with step 1.

3.2.1.4   Propensity parameters

While initial proportions and number of ESCRT - ESCRT bindings varied for the results shown, the propensity parameters used for all simulations are as follows:

Diffusion (for all proteins): 1, ESCRT ESCRT binding: 10, ESCRT Cargo binding: 10, ESCRT GFP-Fyve binding: 10, ESCRT Removal: 1000, Unbinding: 1, and Critical Vesicle Size: 20.

We set the propensity parameters for binding to be equal, independent of protein type. However, since the updates are determined by the calculated probabilities, an ESCRT (with 5 ESCRT ESCRT interaction sites) is usually more likely to associate with another ESCRT than a cargo since it has 5 ESCRT interaction sites and only one cargo interaction site. In this type of probability driven model, binding affinities are not explicitly specified, rather, the qualitative behavior of the protein interactions were accounted for.

For each set of initial proportions of proteins 20 simulations were completed. The amounts of each type of protein (cargo, non-cargo, ESCRT, and GFP-FYVE) were recorded at 100 intervals of 2500 steps of each simulation. The data from each simulation were normalized by dividing by the total number of grid spaces (in this case 28^2) giving a grid concentration of each protein type. The average and standard deviation of the 20 simulations were calculated. Movie frames were captured every 250 runs of a simulation and include 700 frames. Still images of the same simulation were recorded every 25,000 runs of the simulation.

3.2.2   Results

We ran our cellular automata model for different starting conditions and found that sorting efficiency and fidelity was high. Sample snapshots from a run of our simulation and our results are given in Figure 3.2. Figure 3.3 gives results of the model under specified conditions. Our model is indeed able to sort cargo while excluding non-cargo in each case. This suggests that a flexible network of ESCRT-cargo is sufficient for proper sorting contrasting the previously proposed highly ordered models of ESCRT-cargo network formation. In addition, the inclusion of a cargo-mimic (GFP-FYVE) did not disrupt the sorting efficiency and both cargoes and GFP-FYVE were efficiently sorted. We ran the model for different number of possible ESCRT-ESCRT interactions and the sorting efficiency was maintained with 3, 4, or 5 ESCRT-ESCRT interactions.
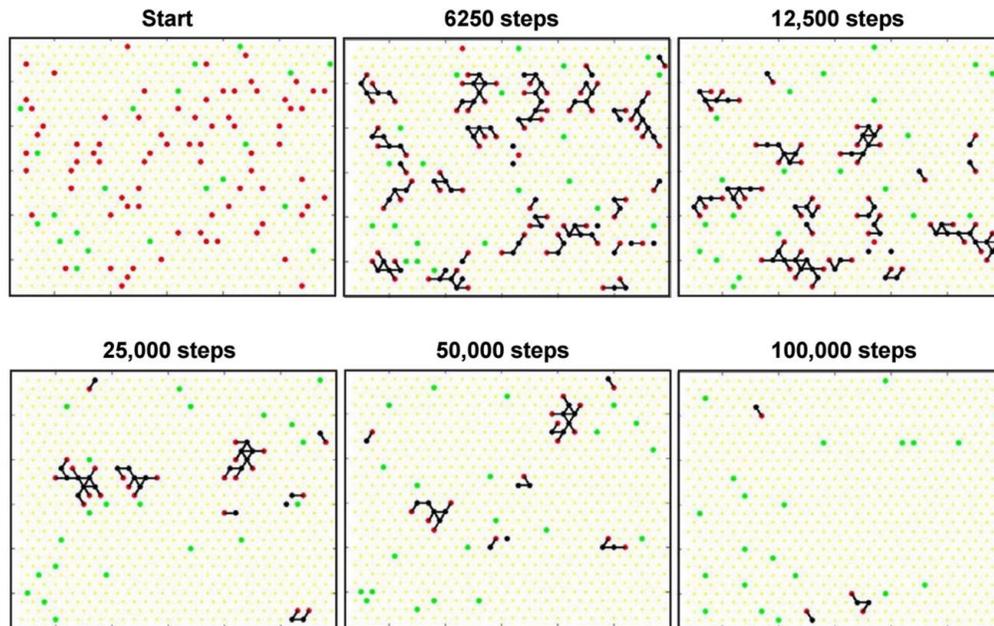
Figure 3.2: Snapshots of a simulation with five 5 ESCRT-ESCRT interaction sites. Cargoes are in red, non-cargoes in green, and ESCRTs in black. Black lines indicate interactions between components. A movie of this simulation can be found in Supporting Information of [28].

3.3   Discussion

In order to achieve proper sorting the specified rules for updates were found to be essential. Of particular importance is the rule that ESCRTs are recruited and stabilized on the membrane through interactions to cargo. Without this rule, ESCRTs are able to polymerize into large networks without any cargo present. This rule is consistent with data demonstrating the necessity of ubiquitinated cargo for ESCRT function [27]. In our current model, we also make the assumption that ESCRT interactions with cargo are stable in networks of more than three proteins although it is well known that UBDs have very low affinity. Recent evidence suggests the possibility of more stable interactions with cargo by showing that two domains of ESCRT-0 function together in binding cargo [42]. This provides some justification for this proposed rule, at least in the case of ESCRT-0.

With the proper rules for sorting as specified, we observe that a flexible network of ESCRTs is enough to achieve efficient sorting with the exclusion of non-cargo. The number of ESCRT-ESCRT interactions also does not appear to have an impact on sorting. This result is promising as no individual ESCRT has been shown to have 5 ESCRT-ESCRT interactions sites. Further, inclusion of a cargo-mimic did not interupt the sorting efficiency.

While these are important qualitative results, we recognize drawbacks to our current
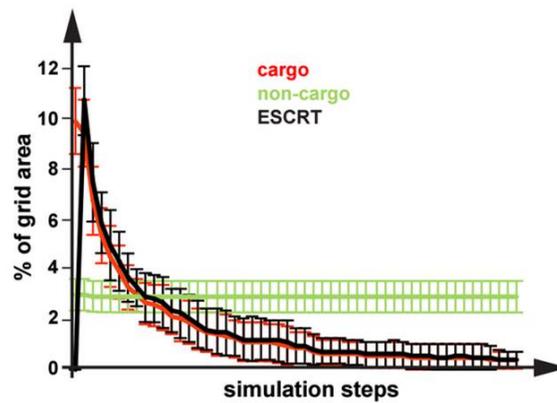
(a)
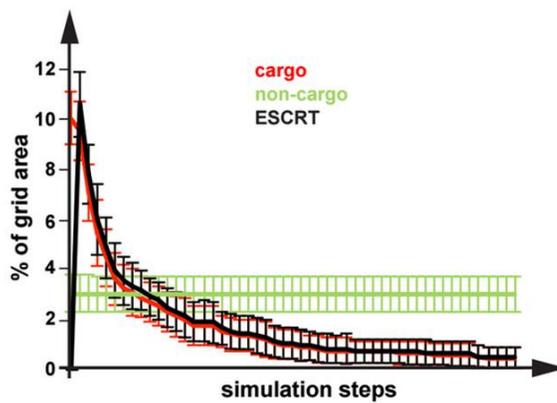
(b)
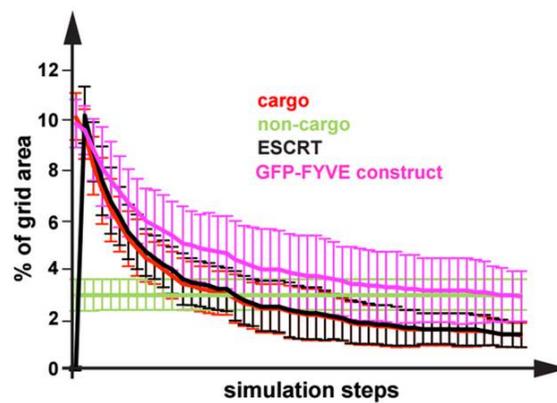
(c)

Figure 3.3: Average % of grid area occupied by different components under the specified conditions for 20 simulations. (a) 5 ESCRT-ESCRT interaction sites. (b) 3 ESCRT-ESCRT interaction sites. (c) Inclusion of GFP-FYVE with 5 ESCRT-ESCRT interaction sites

cellular automata model. Our model only allows each ESCRT to bind to one cargo at a time. While this may be likely for ESCRT-I and ESCRT-II, it has been shown that ESCRT-0 is able to simultaneously bind multiple ubiquitinated cargoes [30]. In order to assess the effects of this binding we would need to specify different ESCRT components in our model. While the concept of a cellular automata model is quite simple, appropriate coding of the specified rules is not trivial. Including more ESCRTs with different rules would likely not be worth the effort as our results would still be purely qualitative and we believe we will still be able to achieve high fidelity sorting. Further, it is not clear how to appropriately include true reaction rates in our cellular automata model as there is no true time. For example, diffusion and binding rates are not accurately modeled in this system but are rather specified by propensities. The probability of a certain update is highly dependent on the number and nature of allowable updates rather than on true kinetic rate parameters.

We began to explore ways to parameterize our model more appropriately and realized that there is little known about how to appropriately model protein interactions on membranes. Binding affinities, such as those reported for UBDs, are typically reported as dissociation constants (K_ds) which have units of concentration. These units to not apply to binding affinities on membranes. We, likewise, realized that there is no well accepted way to transform a three dimensional dissociation constant, with units of concentration, to a two dimensional dissociation constant. Not only do we not know how to include proper rates into our model, we also do not have proper experimental rates for interactions of proteins on membranes.

As a conclusion to this chapter, we propose that a flexible ESCRT-cargo network is sufficient to achieve efficient sorting. We further explore ESCRT mediated protein sorting in Chapter 5 after exploring the general consequences of dimensionality in protein interactions in the next chapter.

CHAPTER 4

DIMENSIONAL DEPENDENCE OF
BINDING KINETICS

4.1 The Dissociation Constant, K_d

Binding affinities of proteins which interact with one another are usually expressed in terms of K_d, the dissociation constant. As a simple mathematical example, consider the following binding reaction between two proteins, A and B, which are able to reversibly bind forming complex C,

$$A + B \underset{k_-}{\overset{k_+}{\rightleftarrows}} C. \tag{4.1}$$

Using the law of mass action, we model the change in concentration of complex C,

$$\frac{dC}{dt} = k_+ AB - k_- C. \tag{4.2}$$

If the proteins are in a three-dimensional space measured in units of volume concentration, k_+ has units of $(\text{concentration} \cdot \text{time})^{-1}$ and k_- has units of $(\text{time})^{-1}$. At steady state, it follows that

$$\frac{k_-}{k_+} = \frac{AB}{C}. \tag{4.3}$$

The above value is defined as the K_d for this reaction and has units of volume concentration in three-dimensional space. This is the traditional definition of K_d. However, if the proteins are membrane-bound, their concentrations are measured in units of areal concentration which leads to a K_d with units of areal concentration not volume concentration. Most experimentally measured K_ds are measured in solution and given in units of volume concentration. This is especially true in the case of soluble proteins which can also be membrane bound. This begs the question, what does K_d mean in two dimensions and is there a way to convert a traditional three-dimensional K_d into a two-dimensional K_d in order to appropriately model protein interactions on membranes? Others have also noticed this problem of understanding dissociation constants in two dimensions. The first attempt

we are aware of to address this issue were presented by Bell et al. as described in [3, 50]. They suggest a very simple conversion of between two- and three-dimensional K_ds in the context of bonds between cell-cell adhesion receptors,

$$K_d^{(2)} = \xi K_d^{(3)}. \tag{4.4}$$

This conversion relies on what is termed the confinement length, ξ. If we think of traditional x, y, and z coordinates with the x-y plane being the membrane of interest, ξ represents the z length of the molecule confined to the membrane during cell-cell adhesion. Indeed this conversion produces the correct units for K_ds in both dimensions and is appealing in its simplicity. In 2011, Wu et al. addressed the transformation of three-dimensional binding affinities to two-dimensional binding affinities in the context of cadherin clustering. [50] They recognize that while experimental studies have been performed to determine the confinement length, ξ, for different proteins [10] there is much variation among these measurements. They use very specific properties of cadherin molecules to formulate more complex transformations between 3D and 2D K_ds for this specific case. They specify two types of cadherin interactions termed trans and cis interactions. Trans interactions are interactions between cadherin molecules on different cell membranes during adhesion and cis interactions are interactions between cadhering molecules on the same membrane. Cis interactions are the type of interactions we are currently interested in. Wu et al. use very specific properties of caherin molecules to formulate more complex transformations between 3D and 2D K_ds than a simple confinement length for both trans and cis interactions. Dustin et al. also suggest that relationships between K_ds in two and three dimensions for cell-cell adhesion molecules are much more involved than the previous simple transformation.

As opposed to trying to better determine a confinement length for the specific proteins we are interested in, we have chosen to address the problem from an entirely different viewpoint beginning with the essential features of protein interactions. In the remainder of this chapter, we consider the essential features of protein interactions and how to calculate binding rates and dissociation constants of protein interactions in one, two, and three dimensions using stochastic methods. Many have worked to determine protein interaction rates in three dimensions. Less work has been done to determine binding rates in lower dimensions. In the 1980s, diffusion limited interaction rates in membranes were studied by considering the probability of first return for a random walk on a lattice [44]. Here we present a novel approach to determining interaction rates by considering the general problem of two particle interaction in one, two, and three dimensions. Our calculations do not rely on the interaction being either diffusion limited or reaction limited. Further, we explore

protein associations in domains with both periodic boundaries and no flux boundaries. We also calculate dissociation constants for reversible binding reactions and provide a general transformation between dissociation constants in two and three dimensions. Our work suggests that transformations of dissociation constants between dimensions is, in general, much more complex than previously suggested.

4.2   Essential Features of Protein Interactions

At its most basic description, protein-protein interaction is defined by interactions between a region of one protein and a region of another. These regions that interact are termed binding sites. With these simple definitions in mind, there are essentially only two necessary conditions for protein protein interaction.

1. Proteins must be in sufficient proximity to allow for binding.

2. Binding sites must be properly oriented in order to interact.

These are the two features at the core of our binding models. We argue that dimensionality affects both features. Diffusion coefficients for proteins are indeed different on membranes than in solution [38]. In addition, proteins associated with membranes are fixed in a particular orientation as directed by their membrane associations.

4.3   Orientational Constraints

To better understand the effects of orientational constraints, we introduce a highly simplified model with two cubical proteins, each of which contain one binding site on one of their faces. Assuming that proteins are in sufficient proximity to bind, we calculate the probabilities that they are oriented correctly to facilitate binding under the assumption that all orientations are equally likely. In our model, we allow for 6 different orientations in three dimensions and 4 different orientations in 2 dimensions. One can think of this highly simplified model as two dice on a table that are close together which may only interact if the adjacent surfaces are both showing the number 3. In three dimensions, each die has a 1 in 6 chance of being oriented correctly. As the dice are independent, this leads to a probability of $\frac{1}{36}$ that they are able to interact. Now suppose that to change between orientations the dice must remain on the table at all times. This leads to only 4 possible configurations for each die and the probability of interaction depends on where the face showing the number 3 is positioned relative to the surface of the table. If both die have the face with number 3 as a possible interaction face, there is a 1 in 4 chance of each protein being oriented

correctly leading to an interaction probability of $\frac{1}{16}$, over double the probability calculated in three dimensions. On the other hand, if the face showing the number 3 is not in the plane of interaction, the probability of interaction is 0 as the binding sites will never come in contact with one another. This simple example is depicted in Figure 4.1 and illustrates the point that orientational constraints greatly affect proteins' abilities to interact with one another. We, therefore, argue that having an understanding of location of binding sites in two and three dimensions is essential for understanding K_ds and how to convert them between dimensions.

4.4   Discrete Space Binding Model

We now present a discrete space model for two particle interactions in one, two, and three dimensions. With the two essential features of protein interaction in mind, we seek to calculate the expected time of interaction between two particles. As opposed to tracking the two particles independently, we instead track the distance between the two particles thereby decreasing our state space. We begin by outlining this model in one dimension. We discretize our state space using the interaction radius of the two particles which we call h. We define h as the distance of sufficient proximity between two particles that allows for interaction. For a one-dimensional domain of length L, we define N to be $\frac{L}{h}$, the number of grid points on which the proteins move. To track the distance between the particles, we need to determine the state space for our distance model. In one dimension, the distance state space is the number of grid spaces between the proteins. If the proteins occupy the same grid point, we call this state 0 as they are considered to be "zero distance" apart on
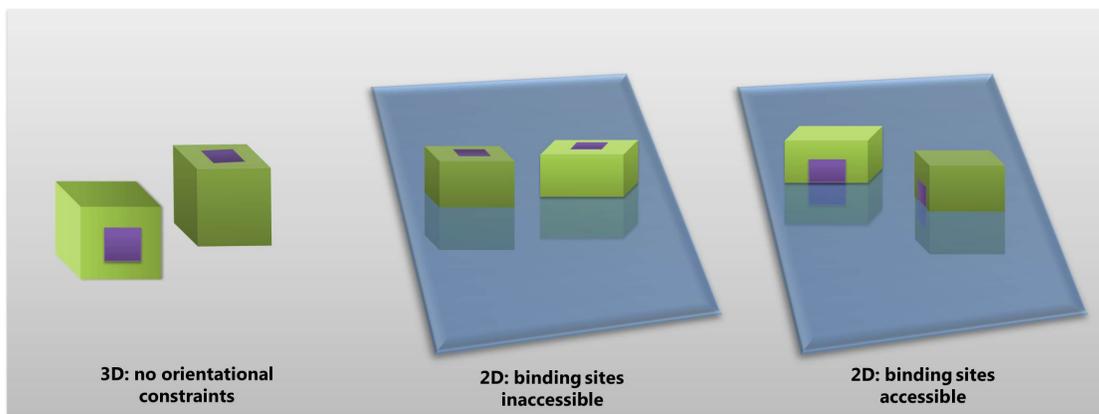
Figure 4.1: Example of orientational constraints in 3D vs. 2D.

the grid. If the proteins occupy adjacent grid spaces, they are in state 1, and so on. The maximal possible distance between the particles depends on the boundary conditions of the grid on which the proteins move. If the particles move on a periodic domain, the maximal possible distance, M_p, is $\frac{N}{2}$. If the particles move on a domain with no flux boundaries, the maximal possible distance, M_{nf}, is $N - 1$.

We assume that both particles diffuse at a rate δ and if they occupy the same grid space will bind at rate k. Taking into consideration the orientational constraints, we define $k = \hat{k}\rho$ where ρ represents the probability to bind given the two particles occupy the same grid space and have "zero distance". In general $\rho = \rho^{(d)}$ as it is dimensionally dependent as it depends on orientational constraints. The rate $\hat{k}$ represents the true rate of binding, i.e., the rate at which two binding sites interact when they are oriented correctly and in sufficient proximity. We argue that $\hat{k}$ is a feature of the binding regions themselves and is independent of dimension. This model setup leads to the following state diagram where the state I represents interaction of the particles and state M is the maximal distance, dependent on boundary conditions:

$$I \overset{k}{\leftarrow} 0 \rightleftarrows 1 \rightleftarrows 2 \rightleftarrows \cdots \rightleftarrows M. \tag{4.5}$$

Given random starting positions for both proteins, we seek to determine the time it takes for them to interact, a stochastic exit time problem. The rates at which proteins transition between states in this distance space are dependent on the boundary conditions we impose on the particles. If the particles move in a domain with periodic boundaries, we are able to compute the transitions exactly. In the case of no flux boundaries, we can compute approximate transitions assuming that all possible ways to be in a given difference state are equally likely. To compute these rates, we calculate the number of possible ways to transition from state i to state j and divide by the number of possible ways to be in state i given that each particle moves at rate δ. In one dimension, we now have two state diagrams depending on boundary conditions.

In the case of periodic boundaries, we take $M = M_p$ and use the following transition diagram:

$$I \overset{k}{\leftarrow} 0 \underset{2\delta}{\overset{4\delta}{\rightleftarrows}} 1 \underset{2\delta}{\overset{2\delta}{\rightleftarrows}} 2 \underset{2\delta}{\overset{2\delta}{\rightleftarrows}} \cdots (M-2) \underset{2\delta}{\overset{2\delta}{\rightleftarrows}} (M-1) \underset{4\delta}{\overset{2\delta}{\rightleftarrows}} M. \tag{4.6}$$

In the case of no flux boundaries, we take $M = M_{nf}$ and use the following transition diagram:

$$I \overset{k}{\leftarrow} 0 \underset{2\delta}{\overset{\frac{4(N-1)\delta}{N}}{\rightleftarrows}} 1 \underset{2\delta}{\overset{\frac{2(N-2)\delta}{(N-1)}}{\rightleftarrows}} 2 \underset{2\delta}{\overset{\frac{2(N-3)\delta}{(N-2)}}{\rightleftarrows}} \cdots (M-2) \underset{2\delta}{\overset{\frac{2(N-(M-1))\delta}{(N-(M-2))}}{\rightleftarrows}} (M-1) \underset{2\delta}{\overset{\frac{2(N-M)\delta}{(N-(M-1))}}{\rightleftarrows}} M. \tag{4.7}$$

For the remainder of our computations, we use the case where the particles move in a domain with periodic boundaries and note that the calculations are similar in the no flux case. We now write down the differential equation of this process with P_j being the probability of being in state j,

$$\frac{d}{dt}\begin{pmatrix} P_I \\ P_0 \\ P_1 \\ \vdots \\ \vdots \\ P_M \end{pmatrix} = \begin{pmatrix} k & & & & & \\ -(k+4\delta) & 2\delta & & & & \\ 2\delta & -4\delta & 2\delta & & & \\ & \ddots & \ddots & \ddots & & \\ & & 2\delta & -4\delta & 4\delta & \\ & & & 2\delta & -4\delta & \end{pmatrix}\begin{pmatrix} P_I \\ P_0 \\ P_1 \\ \vdots \\ \vdots \\ P_M \end{pmatrix}. \tag{4.8}$$

We now wish to calculate the expected time for the two particles to interact given that they have random starting positions on the grid. Alternatively, this is the time it takes to leave the system through state I. We call the expected interaction time, T, and calculate it as follows with details included in Appendix B for calculating general discrete space stochastic exit times:

$$T = -\mathbf{1}^T W^{-1} P_o \tag{4.9}$$

where

$$W = \begin{pmatrix} -(k+4\delta) & 2\delta & & & \\ 4\delta & -4\delta & 2\delta & & \\ & \ddots & \ddots & \ddots & \\ & & 2\delta & -4\delta & 4\delta \\ & & & 2\delta & -4\delta \end{pmatrix} \tag{4.10}$$

and $\mathbf{1}$ is the $(M+1)$ by 1 vector of all ones. P_o is the $(M+1)$ by 1 vector of initial probabilities in our difference states assuming random starting positions for both particles on a periodic grid,

$$P_o = \begin{pmatrix} \frac{1}{N} \\ \frac{2}{N} \\ \frac{2}{N} \\ \vdots \\ \frac{2}{N} \\ \frac{2}{N} \\ \frac{1}{N} \end{pmatrix}. \tag{4.11}$$

This initial vector, P_o is different for the case of no flux boundaries. We now must compute W^{-1}. We can rewrite the matrix W as $W = \delta L - kE$ where

$$
L = \begin{pmatrix}
-4 & 2 & & & \\
4 & -4 & 2 & & \\
& \ddots & \ddots & \ddots & \\
& & 2 & -4 & 4 \\
& & & 2 & -4
\end{pmatrix},
\tag{4.12}
$$

and

$$
E = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\tag{4.13}
$$

In this case and for any L with zero column sum, we can prove that W^{-1} is of the form

$$
W^{-1} = \frac{A}{\delta} - \frac{\Pi}{k},
\tag{4.14}
$$

where A and Π are independent of both δ and k, and $\Pi = \phi \mathbf{1}^T$ where ϕ is an element of the null space of L normalized such that $\phi(1) = 1$. This proof is included in Appendix B. We now find that the expected time for the particles to interact is

$$
\begin{aligned}
T &= -\mathbf{1}^T \left(\frac{A}{\delta} - \frac{\Pi}{k} \right) P_o \tag{4.15} \\
&= \frac{-\mathbf{1}^T A P_o}{\delta} + \frac{\mathbf{1}^T \phi \mathbf{1}^T P_o}{k}. \tag{4.16}
\end{aligned}
$$

In one dimension, $\mathbf{1}^T \phi \mathbf{1}^T P_o = N$. In general, this quantity is N^d where d specifies the dimension. We now need to determine the value of $-\mathbf{1}^T A P_o$ which we call R. We are not able to solve for R analytically so we turn to numerical methods. We use Matlab to calculate T using a linear solve to compute $W^{-1} P_o$. We then can calculate R,

$$
R = \left(T - \frac{N}{k} \right) \delta.
\tag{4.17}
$$

We do this calculation for a range of N values as we know R is a function of N only and use Matlab to fit R as a function of N.

In one dimension, we are able to fit R exactly. In the case of periodic boundaries we calculate

$$
R = \frac{N^2 - 1}{24}.
\tag{4.18}
$$

In the case of no flux boundaries, we calculate

$$
R = \frac{N^2}{16} + \frac{N}{24} - \frac{1}{15}.
\tag{4.19}
$$

.

While the coefficients differ, we see that, with either choice of boundary conditions, R is a quadratic function of N. We now can express our expected interaction time, T, in one dimension, as follows. For periodic boundaries,

$$T = \frac{\frac{N^2-1}{24}}{\delta} + \frac{N}{k}. \tag{4.20}$$

For no flux boundaries,

$$T = \frac{\frac{N^2}{16} + \frac{N}{24} - \frac{1}{15}}{\delta} + \frac{N}{k}. \tag{4.21}$$

The above calculations can be extended to both two and three dimensions. In two dimensions, we allow the particles to diffuse on an N by N grid and we track the difference between them in both dimensions. The state transitions in two dimensions are as follows:

$$
\begin{array}{c}
I \;\leftarrow\; (0,0) \;\rightleftarrows\; (0,1) \;\rightleftarrows\; \cdots \;\rightleftarrows\; (0,M) \\
\quad\;\; \updownarrow \qquad\quad\; \updownarrow \qquad\qquad\qquad\;\; \updownarrow \\
(1,0) \;\rightleftarrows\; (1,1) \;\rightleftarrows\; \cdots \;\rightleftarrows\; (1,M) \\
\quad\;\; \updownarrow \qquad\quad\; \updownarrow \qquad\quad \updownarrow \qquad\quad \updownarrow \\
\vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \\
\quad\;\; \updownarrow \qquad\quad\; \updownarrow \qquad\quad \updownarrow \qquad\quad \updownarrow \\
(M,0) \;\rightleftarrows\; (M,1) \;\rightleftarrows\; \cdots \;\rightleftarrows\; (M,M)
\end{array}
\tag{4.22}
$$

We assume that the directions of movement are independent and can therefore use the one-dimensional rates to determine the appropriate rates in higher dimensions. The matrix W in two dimensions is M^2 by M^2 and in three dimensions W is M^3 by M^3. The W matrices in both cases are block diagonal and the form of W^{-1} as given in (4.14) still holds as the matrices again may be written as $W = \delta L - kE$ with L having zero column sum. Therefore, we can calculate the value of R in higher dimensions as we have above. However, computing $W^{-1}P_o$ becomes increasingly computationally intensive as N increases in two and three dimensions. Due to this issue, we are limited in the range of N values we can use to fit R in higher dimensions. Regardless, we are still able to obtain the following forms of R.

For particles in 2D with periodic boundaries,

$$R = \left(\frac{1}{12}N^2 + 0.1758N - 2.112\right)\ln(N). \tag{4.23}$$

For particles in 2D with no flux boundaries,

$$R = \left(\frac{12}{125}N^2 + 0.4413N - 1.5443\right)\ln(N). \tag{4.24}$$

For particles in 3D with periodic boundaries,

$$R = 0.1264N^3 - 0.1128N^2 - 0.0023N + 0.0437. \tag{4.25}$$

For particles 2D with periodic boundaries,

$$R = 0.1277N^3 + 0.2001N^2 - 0.6116N + 0.8163. \tag{4.26}$$

If we assume that $L >> h$, which is reasonable for most protein protein interactions, then $N >> 1$ and we only consider the leading order behavior. As such, we can summarize the expected interaction times as follows. We use subscripts to denote boundary conditions and superscripts to denote dimension. We find

$$T_p^{(1)} = \frac{\frac{1}{24}N^2}{\delta} + \frac{N}{k}, \tag{4.27}$$

$$T_{nf}^{(1)} = \frac{\frac{1}{16}N^2}{\delta} + \frac{N}{k}, \tag{4.28}$$

$$T_p^{(2)} = \frac{\frac{1}{12}N^2 \ln N}{\delta} + \frac{N^2}{k}, \tag{4.29}$$

$$T_{nf}^{(2)} = \frac{\frac{12}{125}N^2 \ln N}{\delta} + \frac{N^2}{k}, \tag{4.30}$$

$$T_p^{(3)} = \frac{0.1264N^3}{\delta} + \frac{N^3}{k}, \tag{4.31}$$

$$T_{nf}^{(3)} = \frac{0.1277N^3}{\delta} + \frac{N^3}{k}. \tag{4.32}$$

To validate these results, we perform Gillespie simulations of this process in each case and calculate the average interaction time over many simulations. We observe good agreement between the interaction times reported above and the average interaction times of our Gillespie simulations. The agreement is better for periodic boundary conditions than it is for no flux boundary conditions which is expected as the no flux transition rates in our model are only approximate. For example, consider a 1D grid with $N = 10$. Let $\delta = 1$ and $k = .1$ both with units of inverse seconds. For periodic boundaries $T_p^{(1)} = \frac{1}{24}10^2 + \frac{10}{.1} \approx 104.1667$sec and the mean interaction time of 50,000 Gillespie simulations is 104.2193sec. For no flux boundaries $T_{nf}^{(1)} = \frac{1}{16}10^2 + \frac{10}{.1} = 106.25$sec and the mean interaction time of 50,000 Gillespie simulations is 108.611sec.

Our calculations in this section suggest that the expected interaction times for two diffusing particles are dependent on diffusion, true binding rate, probability of correct orientation, size of the domain, interaction radius of the particles, and dimension. We explore this interaction time problem further in the next three sections and then employ our results to discuss dissociation constants in one, two, and three dimensions.

4.5   Result of Periodic Boundaries

If we assume the particles move on a grid with periodic boundaries we find that tracking the distance between the two particles, each moving with diffusion rate δ, is the same

as tracking the position of one particle with diffusion rate 2δ. We can easily see this by comparing the transition matrices in both cases. Suppose we have one particle moving on a 1D grid of size N. This particle may interact at a rate k if it is at position 0. The W matrix (N by N) for this problem is

$$W = \begin{pmatrix} -(k+2\delta) & \delta & & & \delta \\ \delta & -2\delta & \delta & & \\ & & \ddots & \ddots & \ddots \\ & & & \delta & -2\delta & \delta \\ \delta & & & & \delta & -\delta \end{pmatrix}. \tag{4.33}$$

Now consider that instead of tracking the position of the particle, we track its distance from position 0. This results in the following matrix where the periodic terms are collapsed onto the off diagonals. This new W matrix has size $(M+1)$ by $(M+1)$,

$$W = \begin{pmatrix} -(k+2\delta) & \delta & & \\ 2\delta & -2\delta & \delta & \\ & & \ddots & \ddots & \ddots \\ & & & \delta & -2\delta & 2\delta \\ & & & & \delta & -\delta \end{pmatrix}. \tag{4.34}$$

This matrix is the same as the original W matrix (4.10) scaled by 2. Therefore, we observe that the interaction time problem of two particles, both moving at rate δ, on a periodic domain of size N is the same as the interaction time problems for one particle, moving at rate 2δ on a periodic domain of size N. We employ this fact in the next two sections.

4.6   Eigenvalue Decomposition

The computations in Section 5.4 rely on numerical computation of W^{-1}. As previously explained, this computation becomes very computationally intensive for higher dimensions. We are limited for the N values for which we can compute R. We address this here by computing R in a different way. Recall that $R = -\mathbf{1}^T A P_o$. Suppose we guess that $A P_o = \alpha P_o + \Phi \beta$ where α is a scalar, β is a matrix, and Φ is the matrix of eigenvectors corresponding to non-zero eigenvectors Λ of L, $L\Phi = \Phi\Lambda$.

With the given forms of $W = \delta L - kE$ and $W^{-1} = \frac{A}{\delta} - \frac{\Pi}{k}$, we have the following two conditions which we use to solve for α and β:

$$LAP_o = P_o - e, \tag{4.35}$$

$$EAP_o = 0. \tag{4.36}$$

With our guess of $AP_o = \alpha P_o + \Phi\beta$, (4.35) yields,

$$LAP_o = \alpha LP_o + L\Phi\beta = P_o - e. \tag{4.37}$$

Since P_o is in the nullspace of L, it follows that

$$L\Phi\beta = P_o - e. \tag{4.38}$$

Using the fact that $L\Phi = \Phi\Lambda$, we see that

$$\Phi\Lambda\beta = P_o - e. \tag{4.39}$$

Let Ψ represent the matrix of left eigenvectors of L, $\Psi L = \Lambda\Psi$. If we choose the eigenvectors such that they form a biorthogonal set, then $\Psi\Phi = I$. In fact, for our case, it turns out that $\Psi = \Phi^*$ as L is a self adjoint matrix. Therefore,

$$\Lambda\beta = \Phi^*(P_o - e). \tag{4.40}$$

Inverting Λ and applying to both sides yields,

$$\beta = \Lambda^{-1}\Phi^*(P_o - e). \tag{4.41}$$

We now know β and from (4.36) we see that

$$EAP_o = \alpha EP_o - E\Phi\Lambda^{-1}\Phi^*(P_o - e) = 0. \tag{4.42}$$

Therefore,

$$\alpha = \frac{[E\Phi\Lambda^{-1}\Phi^*(P_o - e)](1)}{P_o(1)}. \tag{4.43}$$

where $P_o(1)$ represents the first element of the vector P_o and $[E\Phi\Lambda^{-1}\Phi^*(P_o - e)](1)$ represents the first element of the vector $[E\Phi\Lambda^{-1}\Phi^*(P_o - e)]$. Recall that we wish to calculate R which is $-\mathbf{1}^T AP_o$. With our guess of $AP_o = \alpha P_o + \Phi\beta$ we see that,

$$-\mathbf{1}^T AP_o \;=\; -\mathbf{1}^T \alpha P_o - \mathbf{1}^T \Phi\beta \tag{4.44}$$
$$=\; -\alpha\mathbf{1}^T P_o \tag{4.45}$$
$$=\; -\alpha. \tag{4.46}$$

Therefore, $R = -\alpha$ which may be calculated from (4.43). We now can determine R using the eigenvalues and eigenvectors of L. As mentioned in the previous section, to solve our problem on a periodic domain, it suffices to solve the problem of one particle moving on a

periodic domain of size N. Therefore, we compute the eigenvalues and eigenvectors for the L matrix below of size N by N,

$$L = \begin{pmatrix} -4 & 2 & & & 2 \\ 2 & -4 & 2 & & \\ & \ddots & \ddots & \ddots & \\ & & 2 & -4 & 2 \\ 2 & & & 2 & -4 \end{pmatrix}. \tag{4.47}$$

The non-zero eigenvalues and corresponding eigenvectors for this matrix can be written as follows with j corresponding to the component of the kth eigenvector. In this case, we have $N-1$ eigenvectors each with N components since we are only considering the non-zero eigenvalues and their corresponding eigenvectors. The eigenvalues and eigenvectors are

$$\lambda_k = -8\sin^2\left(\frac{\pi k}{N}\right), \tag{4.48}$$

$$\phi_{(k,j)} = \frac{1}{\sqrt{N}}e^{\left(\frac{2\pi ijk}{N}\right)}, \tag{4.49}$$

$$\phi^{*j}_{(k,j)} = \frac{1}{\sqrt{N}}e^{\left(\frac{-2\pi ijk}{N}\right)}. \tag{4.50}$$

We can now compute R with the following sum:

$$R = \frac{1}{P_o(1)}\sum_{m=1}^{N}\sum_{n=1}^{N-1}\frac{\phi_{(n,1)}\phi^*_{(n,m)}}{\lambda_n}(P_o - e)(m). \tag{4.51}$$

For the 1D periodic case, we again see that $R = \frac{N^2-1}{24}$. This eigenvalue decomposition provides an alternative way to solve for R that is computationally less intensive and the exit time results hold.

4.7 Continuous Space Binding Model

We wish to solve this problem in continuous space as well as discrete space. To do this, we must solve the continuous exit time problem associated with the following stochastic differential equations. We again employ the fact that solving for the interaction time of two particles on a periodic domain is identical to solving for the interaction time of 1 particle on a periodic domain with twice the rate of diffusion. D in this case is a true diffusion coefficient not a diffusion rate. We take $D = \delta h^2$. The spatial variable, $\mathbf{x}$, is written as a vector and the size of $\mathbf{x}$ depends on the dimension we are considering.

$$\frac{\partial}{\partial t}P(0,\mathbf{x},t|0,\mathbf{y},s) = 2D\frac{\partial^2}{\partial\mathbf{x}^2}P(0,\mathbf{x},t|0,\mathbf{y},s) - k(\mathbf{x})P(0,\mathbf{x},t|0,\mathbf{y},s) \tag{4.52}$$

$$\frac{\partial}{\partial t}P(1,t|0,\mathbf{y},s) = k(\mathbf{x})P(0,\mathbf{x},t|0,\mathbf{y},s) \tag{4.53}$$

with

$$k(\mathbf{x}) = \begin{cases} k, & \text{if } \mathbf{x} < \mathbf{h} \\ 0, & \text{otherwise} \end{cases} \tag{4.54}$$

P, in this case, represents the probability of being in a particular state in our state space which is continuous in time and space but discrete in the binding state, bound or unbound. Hence, the first independent variable represents bound (0) or unbound (1), the second independent variable represents the location of the particle, and the third independent variable represents time. Therefore, $P(0, \mathbf{x}, t|0, \mathbf{y}, s)$ is the probability of being unbound at position $\mathbf{x}$ at time t given that the particle started at position $\mathbf{y}$ at time s and $P(1, t|0, \mathbf{y}, s)$ represents the probability that the particle is bound at time t given that the particle was unbound at time s at position $\mathbf{y}$. The above equations are the forward Chapman Kolmogorov equations for our problem as they describe how the probabilities change as ending time changes, not starting time.

Given that the particle may bind in a specified region of length h at rate k, we seek to determine the expected time of binding of an initially unbound particle at a random position in the domain. The exit time for an initially unbound particle starting at position $\mathbf{y}$ at time 0 is calculated by solving the following ODE where $E(\mathbf{y})$ is the expected interaction time,

$$-1 = 2DE_{\mathbf{yy}} - k(\mathbf{y})E. \tag{4.55}$$

This exit time equation can be formally derived using the methods presented by Gardiner in [12]. We calculate the expected exit time of a protein at a random starting position in the domain as follows. We call this quantity $T^{(d)}$ to match the notation of exit times calculated in discrete space. We note, again, that in continuous space we can only calculate $T^{(d)}$ for periodic boundary conditions,

$$T^{(d)} = \int_0^L \frac{1}{L^d} E(\mathbf{y}) d\mathbf{y}. \tag{4.56}$$

To obtain the desired exit time, $T^{(d)}$, it remains to solve (4.55). In one dimension, we are able to solve for (4.55) analytically by splitting it into two differential equations,

$$-1 = 2DE1_{yy} \text{ if } h \le y \le L, \tag{4.57}$$

$$-1 = 2DE2_{yy} - kE2 \text{ if } 0 \le y \le h. \tag{4.58}$$

The solution of the first equation is

$$E1(y) = \frac{-y^2}{4D} + c_1 y + c_2, \tag{4.59}$$

where c_1 and c_2 are constants.

To solve the second equation, we observe that a particular solution is given by $E2_p = \frac{1}{k}$. We then approximate the solution of the homogeneous differential equation, $0 = 2DE2_{yy} - kE2$, using a power series approximation which gives $E2_h = c_3 + c_4 y + \frac{kc_3 y^2}{4D} + \frac{kc_4 y^3}{12D}$ with c_3 and c_4 being constants. We now have an approximate full solution to (4.58),

$$
\begin{aligned}
E2(y) &= E2_p + E2_h & (4.60) \\
&= \frac{1}{k} + c_3 + c_4 y + \frac{kc_3 y^2}{4D} + \frac{kc_4 y^3}{12D}. & (4.61)
\end{aligned}
$$

We determine the constants by requiring that the solution and its derivative be continuous at h and at the boundary as the domain is periodic. We now have a full solution for our exit time problem,

$$
E(y) = \begin{cases} E1(y), & \text{if } h \leq y \leq L \\ E2(y), & \text{if } 0 \leq y \leq h \end{cases} \tag{4.62}
$$

Using this solution for E and (4.56), we can calculate $T^{(1)}$. Using the facts that $D = \delta h^2$ and $L = hN$, we obtain the following form which may be compared to (4.27).

$$
T^{(1)} = \frac{k^2 - 48\delta Nk - 3k^2 N + 2k^2 N^2 + 576\delta^2 N^2 + 24kN^3\delta}{24\delta k(-k + 24\delta N)} \tag{4.63}
$$

To compare this with our discrete periodic result given in (4.27), we plot them both for $\delta = 1$ and $k = .1$ for a range of N as shown in Figure 4.2. This demonstrates that, while (4.27) and (4.63) are not identical, the behavior of the exit times as a function of N match using both the discrete and continuous formulation of the problem. In addition, we see that in both cases the leading order behavior in N is identical,

$$
\lim_{N \to \infty} T^{(1)} = \frac{N^2}{24\delta}. \tag{4.64}
$$

In order to solve for the expected interaction time using (4.55) in two and three dimensions, we solve the differential equations numerically using finite differences. This yields the following block diagonal matrix equation in two dimensions,

$$
\begin{pmatrix} -1 \\ -1 \\ \vdots \\ \vdots \\ -1 \\ -1 \end{pmatrix} = \frac{1}{h^2} \begin{pmatrix} A1 & B & & & & B \\ B & A & B & & & \\ & B & A & B & & \\ & & \ddots & \ddots & \ddots & \\ & & & B & A & B \\ B & & & & B & A \end{pmatrix} \begin{pmatrix} E_{(1,1)} \\ E_{(1,2)} \\ \vdots \\ \vdots \\ E_{(N,N-1)} \\ E_{(N,N)} \end{pmatrix}, \tag{4.65}
$$
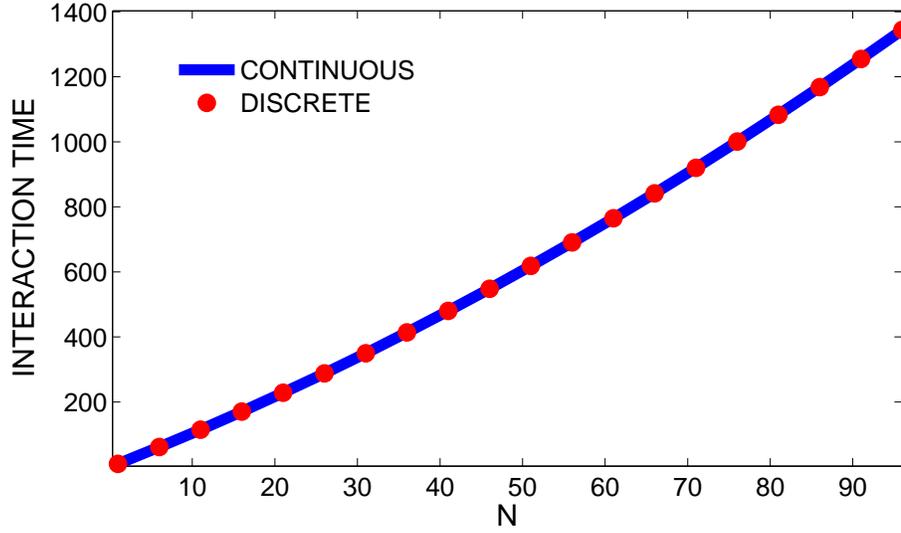
Figure 4.2: Comparison of exit times computed in discrete and continuous space with $\delta = 1$ and $k = .1$.

with

$$
A1 = \begin{pmatrix}
-8D - k & 4D & & & & & 4D \\
4D & -8D & 4D & & & & \\
 & 4D & -8D & 4D & & & \\
 & & & \ddots & \ddots & \ddots & \\
 & & & & 4D & -8D & 4D \\
4D & & & & & 4D & -8D
\end{pmatrix}, \tag{4.66}
$$

$$
A = \begin{pmatrix}
-8D & 4D & & & & & 4D \\
4D & -8D & 4D & & & & \\
 & 4D & -8D & 4D & & & \\
 & & & \ddots & \ddots & \ddots & \\
 & & & & 4D & -8D & 4D \\
4D & & & & & 4D & -8D
\end{pmatrix}, \tag{4.67}
$$

and

$$
B = \begin{pmatrix}
4D & & & & & \\
 & 4D & & & & \\
 & & 4D & & & \\
 & & & \ddots & & \\
 & & & & 4D & \\
 & & & & & 4D
\end{pmatrix} \tag{4.68}
$$

.

Recalling that $D = \delta h^2$ and using the discrete sum, $\mathbf{1}^T \left(\frac{1}{N^2} E \right)$, in place of the integral in (4.56), we observe that the continuous space problem solved with finite differences is

identical to the discrete problem solved in Section 4.4. The same behavior holds in three dimensions as well. However, the continuous formulation of our problem solved with finite differences does not require, h, to be the space step. The radius of interaction, h, is easily discretized over any choice of space step. As a conclusion of our continuous space results, we observe the same behavior for expected interaction times as calculated using discrete space.

4.8 Summary of Exit Time Results and Reaction Rates

In the previous four sections, we have outlined our work on calculating the expected interaction times for two particles, both with diffusion rate δ, which interact at rate k if they are in sufficient proximity to one another. We have seen that approaching this problem in discrete space as well as continuous space led us to the same forms for our results. Importantly, our discrete model demonstrates that these forms hold in the cases of both periodic boundaries and no flux boundaries. The exit time results are as follows with $\alpha^{(1)}$, $\alpha^{(2)}$, and $\alpha^{(3)}$ representing the coefficients reported in Section 5.4 which are also dependent on boundary conditions:

$$T^{(1)} = \frac{\alpha^{(1)} N^2}{\delta} + \frac{N}{k}, \tag{4.69}$$

$$T^{(2)} = \frac{\alpha^{(2)} N^2 \ln(N)}{\delta} + \frac{N^2}{k}, \tag{4.70}$$

$$T^{(3)} = \frac{\alpha^{(3)} N^3}{\delta} + \frac{N^3}{k}. \tag{4.71}$$

Interestingly, the form of the diffusive piece of these exit times is similar to those calculated for the expected diffusion time for a particle to reach a small region in the center of one-, two-, and three-dimensional domains [5]. However, our interaction times are for two particle interactions and also include the addition of a reaction term which also depends on dimension.

To turn these into reaction rates, we simply need invert them. Recalling that $N = L/h$ and letting $\delta = \frac{D}{h^2}$, where D is the diffusion coefficient of both particles, we obtain reaction rates dependent on dimension,

$$r^{(1)} = \frac{k\left(\frac{h}{L}\right)}{1 + \alpha^{(1)}\left(\frac{kh^2}{D}\right)\left(\frac{L}{h}\right)}, \tag{4.72}$$

$$r^{(2)} = \frac{k\left(\frac{h}{L}\right)^2}{1 + \alpha^{(2)}\left(\frac{kh^2}{D}\right)\ln\left(\frac{L}{h}\right)}, \tag{4.73}$$

$$r^{(3)} \quad = \quad \frac{k\left(\frac{h}{L}\right)^3}{1 + \alpha^{(3)}\left(\frac{kh^2}{D}\right)}. \tag{4.74}$$

While not explicitly specified in the above reaction rates, we note that for the problem of protein dimerization, D and k are also dimensionally dependent. Recall that the rate k is $\hat{k}\rho^{(d)}$ where $\hat{k}$ is the true binding rate independent of dimension and $\rho^{(d)}$ is the probability of correct orientation which is dependent on dimension. Presumably, diffusion coefficients are also dimensionally dependent and depend on the viscosity of the medium the proteins are in [38]. These reaction rates are simply rates with units of inverse time unlike the traditional k_+ of a binding event.

4.9   Mass Action Kinetics and Kds

In order to compare the computed reaction rates to what is typically considered the k_+ of a binding event and to calculate K_ds, we consider these rates in mass action kinetics. Consider the dimerization reaction of a n identical particles. With n particles in a given domain, there are $\binom{n}{2}$ possible reactions. We assume the particles bind at rate $r^{(d)}$, depending on the dimension of the domain (d), and track the number of monomers with the following differential equation:

$$\frac{dn}{dt} \quad = \quad -2\binom{n}{2}r^{(d)}, \tag{4.75}$$

$$= \quad -2\frac{n(n-1)}{2}r^{(d)}. \tag{4.76}$$

If we assume that n is large, we can approximate this with,

$$\frac{dn}{dt} \quad = \quad -n^2 r^{(d)}. \tag{4.77}$$

$$\tag{4.78}$$

Now let us assume that the size of our domain is L^d. The concentration of monomers is then $c = \frac{n}{L^d}$ which gives

$$\frac{dc}{dt} \quad = \quad -c^2 r^{(d)} L^d. \tag{4.79}$$

$$\tag{4.80}$$

This gives reaction rate equations in one, two, and three dimensions.

In 1D,

$$\frac{dc}{dt} = -c^2 \left(\frac{k \left(\frac{h}{L} \right)}{1 + \alpha^{(1)} \left(\frac{kh^2}{D} \right) \left(\frac{L}{h} \right)} \right) L \tag{4.81}$$

$$= -c^2 \left(\frac{kh}{1 + \alpha^{(1)} \left(\frac{kh^2}{D} \right) \left(\frac{L}{h} \right)} \right). \tag{4.82}$$

In 2D,

$$\frac{dc}{dt} = -c^2 \left(\frac{k \left(\frac{h}{L} \right)^2}{1 + \alpha^{(2)} \left(\frac{kh^2}{D} \right) \ln \left(\frac{L}{h} \right)} \right) L^2 \tag{4.83}$$

$$= -c^2 \left(\frac{kh^2}{1 + \alpha^{(2)} \left(\frac{kh^2}{D} \right) \ln \left(\frac{L}{h} \right)} \right). \tag{4.84}$$

In 3D,

$$\frac{dc}{dt} = -c^2 \left(\frac{k \left(\frac{h}{L} \right)^3}{1 + \alpha^{(3)} \left(\frac{kh^2}{D} \right)} \right) L^3 \tag{4.85}$$

$$= -c^2 \left(\frac{kh^3}{1 + \alpha^{(3)} \left(\frac{kh^2}{D} \right)} \right). \tag{4.86}$$

The above reaction rates have units of $(\text{concentration} \cdot \text{time})^{-1}$ consistent with traditional k_+ reaction rates as described in Section 5.1. We call these reaction rates $k_+^{(d)}$. Now, to determine dissociation constants, we must allow this reaction to be reversible. Let b be the concentration of dimers and k_- be the unbinding rate of dimers. We assume that the unbinding rate is simply a property of the proteins and is, therefore, independent of dimension. This is consistent with experimental results [9]. The reversible dimerization reaction is

$$c + c \underset{k_-}{\overset{k_+^{(d)}}{\rightleftarrows}} b. \tag{4.87}$$

In the case of reversible dimerization, we can now calculate the K_ds in each dimension as defined by $K_d^{(d)} = \frac{k_-}{k_+^{(d)}}$. We obtain the following K_ds for dimerization reactions in one, two, and three dimensions.

In 1D,

$$K_d^{(1)} = \frac{k_- \left(1 + \alpha^{(1)} \left(\frac{kh^2}{D}\right) \left(\frac{L}{h}\right)\right)}{kh}.$$ (4.88)

In 2D,

$$K_d^{(2)} = \frac{k_- \left(1 + \alpha^{(2)} \left(\frac{kh^2}{D}\right) \ln\left(\frac{L}{h}\right)\right)}{kh^2}.$$ (4.89)

In 3D,

$$K_d^{(3)} = \frac{k_- \left(1 + \alpha^{(3)} \left(\frac{kh^2}{D}\right)\right)}{kh^3}.$$ (4.90)

The above results can easily be applied to nonidentical particle interactions. If we consider two distinct spherical proteins with diffusion coefficients D_1 and D_2 and radii h_1 and h_2, respectively, we can use the following values for D and h in the equations above:

$$D = \frac{D_1 + D_2}{2},$$ (4.91)

$$h = h_1 + h_2.$$ (4.92)

Note that the probability of correct orientation, $\rho^{(d)}$, which is built into k, is also dependent on the location of binding sites on each particle. With these assumptions, we can apply (4.88), (4.89), and (4.90) to interactions between any two proteins.

We now argue that the dissociation constant, K_d, is highly dependent on dimension. In order to accurately model K_d, there are many necessary pieces of information including the size of the proteins, location of binding sites, orientational constraints in 1D and 2D, diffusion coefficients in all dimensions, and, interestingly, length of the domain in 1D and 2D. Our model suggests that one- and two-dimensional K_ds depend on the length of the domain in which the proteins interact. We are interested to see if this observation holds experimentally. For well-mixed systems, we observe that diffusion and dimensionality play a role in determining dissociation constants.

4.10   Limiting Cases of Reaction Rates

We now examine our reported $k_+^{(d)}$ values in the diffusion limited case and the reaction limited case. In the diffusion limited case, we assume that $k >> \frac{D}{h^2}$; therefore, we evaluate our reaction rates in the limit as k goes to infinity.

$$\lim_{k \to \infty} k_+^{(1)} = \frac{D}{\alpha^{(1)} L}$$ (4.93)

$$\lim_{k \to \infty} k_+^{(2)} = \frac{D}{\alpha^{(2)} \ln\left(\frac{L}{h}\right)} \tag{4.94}$$

$$\lim_{k \to \infty} k_+^{(3)} = \frac{Dh}{\alpha^{(3)}} \tag{4.95}$$

We note that $k_+^{(3)}$ is of the same form as the well-known Smoluchowski result for diffusion limited protein interactions as reported in [5].

In the reaction limited case, we assume $\frac{D}{h^2} >> k$ and evaluate our reaction rates in the limit as $\frac{D}{h^2}$ goes to infinity.

$$\lim_{\frac{D}{h^2} \to \infty} r^{(1)} = kh \tag{4.96}$$

$$\lim_{\frac{D}{h^2} \to \infty} r^{(2)} = kh^2 \tag{4.97}$$

$$\lim_{\frac{D}{h^2} \to \infty} r^{(3)} = kh^3 \tag{4.98}$$

In the reaction limited case, we only need consider the true binding rate, orientational constraints, and the size of the interaction radius.

4.11   Dimerization in 2D vs. 3D

To demonstrate the differences between K_ds in two and three dimensions, we again consider a simple reversible dimerization reaction.

$$\frac{dn}{dt} = -2r^{(d)}n^2 + 2k_- n_2, \tag{4.99}$$

$$\frac{dn_2}{dt} = r^{(d)}n^2 - k_- n_2. \tag{4.100}$$

We include the conservation equation below for the total number of monomers, n_t,

$$n_t = n + 2n_2. \tag{4.101}$$

If we assume that the parameters in three dimensions are identical to the parameters in two dimensions, we observe that proteins are more likely to dimerize in 2D than they are in 3D. We solve the above system of equations at steady state. Figure 4.3 presents a plot of the number of dimers at steady state in both 3D and 2D with domain sizes L^3 and L^2, respectively. This simple example demonstrates the pure dimensional dependence of binding without taking into account orientational constraints and dimensional dependence of diffusion.
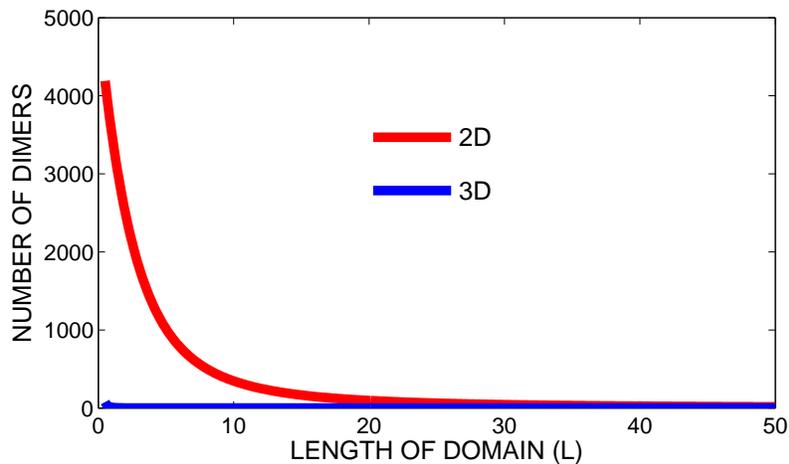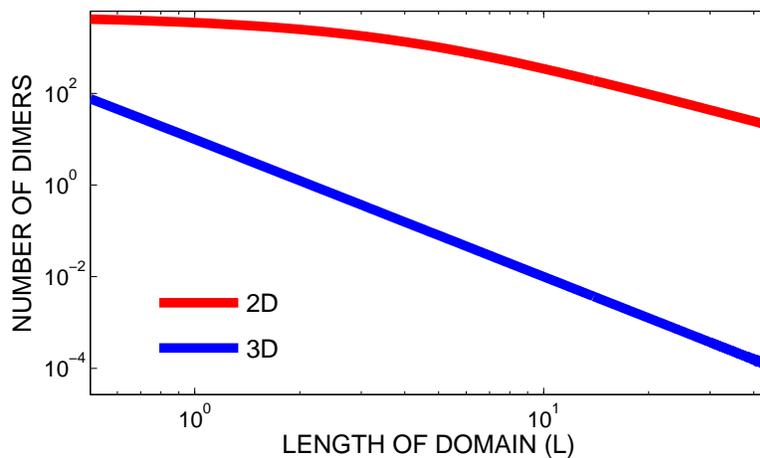
Figure 4.3: Dimerization in 2D vs. 3D. (a) Number of dimers in 2D vs. 3D as the size of the domain varies. Parameters: $n_t = 10000$, $h = .001$, $s = 1$, $D = 100$, $k = 100$. We use the following coefficients for this model from section 4.4, $\alpha^{(2)} = \frac{1}{12}$, $\alpha^{(3)} = .1264$. (b) Same plot as in (a) on a log scale.

4.12 Transforming Kds from 3D to 2D

We now propose a new method of transforming K_ds between two and three dimensions. Assuming the three-dimensional K_d, $K_d^{(3)}$, is known, we present the following conversion to the two-dimensional K_d, $K_d^{(2)}$, for a binding reaction.

$$K_d^{(2)} = K_d^{(3)} \frac{L\rho^{(3)} \left(1 + \alpha^{(2)} \left(\frac{k^{(2)} h^2}{D^{(2)}}\right)\right) \ln\left(\frac{L}{h}\right)}{\rho^{(2)} \left(\frac{L}{h}\right) \left(1 + \alpha^{(3)} \left(\frac{k^{(3)} h^2}{D^{(3)}}\right)\right)} \tag{4.102}$$

It is clear that much is involved in K_d conversions. There is a length scaling by L but this is merely part of the conversion. Recalling $k = \hat{k} P \rho^{(d)}$ where $\rho^{(d)}$, the probability of correct

orientation is dimensionally dependent, $k^{(2)} = ak^{(3)}$ where a is a constant determined by the orientational constraints in two dimensions. The true binding rate, $\hat{k}$, is unknown and it is unlikely that it can be measured experimentally. However, if the $K_d^{(3)}$ is known as are the other parameters, we can solve for the true binding rate, $\hat{k}$. To convert between dimensions, we also must understand how diffusion is affected by dimension for the particular proteins of interest. Further, we need to know something about the structure of the interacting molecules. The size of the proteins determines the value for h. In addition to overall size of the protein, we need to know the position and size of the binding sites to calculate the probability of correct orientation. Lastly, we need to know the size of the domain on which the proteins interact. We conclude that conversions are anything but simple and many pieces of information about the proteins involved must be known to accurately convert between dimensions. We employ this conversion in the next chapter to better model ESCRT binding events.

4.13   Discussion

In this chapter, we have taken an interlude from exploring the roles of ubiquitin in order to better understand the dimensional dependence of protein interactions. We highlighted the essential features of any protein-protein interaction which are the ideas behind all the work in this chapter.

1. Proteins must be in sufficient proximity to allow for binding.

2. Binding sites must be properly oriented in order to interact.

Using a very simple model, we demonstrated that orientational constraints can have a great effect on binding affinities in one, two, and three dimensions. With these ideas in mind, we calculated the expected interaction times for two proteins using both discrete and continuous space models which showed consistent results. In our discrete model, we are able to calculate expected interaction times for proteins diffusing in both periodic and no flux domains. We argue that the overall behavior of interaction times are dimensionally consistent regardless of boundary conditions. Importantly, we observe that in the case of periodic boundaries we can reduce a two particle problem to a one particle problem with twice the rate of diffusion. This observation allows for further analysis for our continuous space model. We invert our calculated expected interaction times to obtain reaction rates and use mass action kinetics to determine values for k_+ which are dimensionally dependent. Importantly, in the diffusion limited case, the form of our result in 3D is consistent with that

of Smoluchowski as presented in [5]. To our knowledge, most binding models assume either the diffusion limited case or the reaction limited case. As opposed to being strictly diffusion limited or strictly reaction limited, our calculated rates include contributions from both diffusion and reaction. We argue that k_- is simply a property of the interacting proteins which has also been observed experimentally [9]. Using the dimensionally dependent values for k_+ and dimensionally independent value of k_- we are able to define dimensionally dependent values of K_d. We observe that K_ds are highly dimensionally dependent. Even if all rates are identical in two and three dimensions, the difference in dimension alone leads to different K_ds. Importantly, our results suggest that one- and two-dimensional K_ds are dependent on the length of the domain in which the proteins interact. This means that the size of the domain on which experimentalists measure a two dimensional K_d will affect the reported K_d.

A main goal of this chapter was to better understand how to convert between K_ds in two and three dimensions. For proteins that can be both soluble and membrane bound, K_ds are often reported as traditional 3D K_ds. It would, therefore, be extremely useful to be able to transform these reported results into two-dimensional K_ds in order to model protein binding on membranes. Here we suggest that this transformation can be achieved but is very involved. In particular, much needs to be known about the proteins involved to appropriately transform K_ds between two and three dimensions, including the size and structure of the proteins, the size and position of the binding sites, the orientational constraints in two dimensions, and the difference in diffusion coefficients. We give examples of such transformations in the next chapter for ESCRT bonds.

Taken together, the results in this chapter demonstrate that dimension must be considered when modeling binding events. The k_+ rates of binding events are heavily influenced by dimension. As such, transforming dissociation constants between dimensions is not simple and requires much information about the proteins involved.

CHAPTER 5

ESCRT SORTING REVISITED

In this chapter, we employ our results on the dimensional dependence of protein interactions to again consider the mechanisms behind ESCRT-mediated sorting of ubiquitinated proteins. We, first, use reported rates to determine an approximate two-dimensional K_d for ESCRT-cargo interactions mediated by ubiquitin. The reported three-dimensional K_ds for UBDs to ubiquitin imply very weak binding of ESCRTs to cargoes. However, cargoes and ESCRTs interact on the endosomal membrane, not in solution. As such, we ask, what is the two-dimensional K_d of UBDs and does being membrane-bound enhance their affinity thereby making these interactions more stable? Next, we present a new differential equation model of ESCRT-cargo network polymerization using our computed two-dimensional interaction rate, $r^{(2)}$. We simulate this model using a Gillespie algorithm in order to gain insights into time it takes to sort a given amount of ubiquitinated proteins on the endosome with flexible ESCRT-cargo protein networks.

5.1 Transforming ESCRT K_ds

The three-dimensional dissociation constants have been measured for ESCRT UBDs and range from $\sim 100-600\mu M$ as reported in [42]. Most of these measurements are for the UBDs and ubiquitin only, not ubiquitinated cargo. Recall that we define cargo as ubiquitinated transmembrane proteins which are not soluble. As such there is not a way to measure three dimensional K_ds of cargo ESCRT interactions. However, the weak binding affinities of UBDs is often presented as a quandary in understanding the mechanisms of protein sorting. We hypothesize that the affinity of ESCRT UBDs is much different on membranes than in solution. To demonstrate the effects of dimensionality on the K_ds of ESCRT, UBDs we assume that our ESCRT-cargo interaction has a three-dimensional K_d of $100\mu M$, similar to the K_d of the UBD itself, and compute the corresponding two dimensional K_d. We wish to employ the following transformation, given in Section 4.12, to better understand ESCRT binding events on membranes.

$$K_d^{(2)} = K_d^{(3)} \frac{\rho^{(3)} L \left(1 + \alpha^{(2)} \left(\frac{k^{(2)} h^2}{D^{(2)}}\right)\right) \ln\left(\frac{L}{h}\right)}{\rho^{(2)} \left(\frac{L}{h}\right) \left(1 + \alpha^{(3)} \left(\frac{k^{(3)} h^2}{D^{(3)}}\right)\right)} \tag{5.1}$$

In order to perform this transformation for a specific three-dimensional K_d, we need many pieces of information about the specific binding event. For example, to determine the dissociation constant for ESCRT-cargo binding on membranes mediated by UBDs, we need to know the following information,

- size of the ESCRT,

- size of the cargo,

- size of UBDs,

- location of UBDs,

- number of UBDs on the ESCRT,

- three dimensional K_d for UBD,

- length of domain in two dimensions,

- true binding rate, $\hat{k}$, and

- diffusional differences in 3D and 2D.

It is immediately obvious that this transformation is anything but simple. However, we are able to obtain or estimate most of this information from the literature. As we do not wish to distinguish between ESCRTs at this point, we model a generic ESCRT with one UBD. This may be a sufficient model for ESCRT-I or ESCRT-II. However, ESCRT-0 has at least 8 UBDs that are able to bind to multiple cargoes simultaneously [30]. As a first model, however, we wish to understand the effects of membrane association on single ESCRT-cargo interactions. We therefore consider the case where ESCRTs have only one UBD.

Each individual ESCRT component has a different size. ESCRT-0 is reported to have a hydrodynamic radius of 7.9nm [37]. ESCRT-I has been shown to span approximately 25nm on endosomal membranes [24]. With these measurements in mind, for our generic ESCRT protein model we assume a radius of 10nm. As the ESCRTs interact with a wide variety of cargo, we simply pick our model ubiquitinated cargo to also have a radius of 10nm. These measurements lead to a computed value for the interaction radius of $h = 10 + 10 = 20$nm.

As approximated in Chapter 2, we estimate the endosome having an approximate radius of 0.5μm. This measurement is supported by [29]. If we assume a spherical endosome, this leads to a surface area of $\pi(\mu\text{m})^2$. As our calculations involve a rectangular domain of size L^2, we then compute that $L = \sqrt{\pi}\mu$m in our model.

5.1.1 Spherical Approximation

We use a spherical approximation for both ESCRTs and cargo. We assume, since the ESCRTs are able to interact when membrane-associated, the binding sites are along the equator parallel to the membrane as shown in Figure 5.1. Under this assumption, all binding sites are "reachable" in both two and three dimensions.

As ESCRT-0 is known to have at least 8 distinct UBDs as well as multiple ESCRT interaction sites, we assume that each UBD and ESCRT interaction site is small relative the circumference of the equator. We assume that each binding site of the ESCRTs is $\frac{1}{15}$ of the total circumference. This allows for 8 UBDs and up to 7 ESCRT-ESCRT interactions which is sufficient for known data. We estimate that the radii of our generic ESCRT protein and our cargo protein are both 10nm. This leads to an equator circumference of $C_E = 20\pi$nm and an approximate binding site radius of $rb = \frac{2\pi}{3}$nm. For protein interactions in three dimensions, we assume these binding sites are circular leading to a binding area of $Ab = \pi\left(\frac{2\pi}{3}\right)^2$nm^2. The surface area of the proteins in our model is $SA = \frac{4000\pi}{3}$nm^3. To calculate the probability of being correctly oriented to bind, we must know the ratio of the binding region to total interaction region of the proteins in both dimensions. We calculate that in three dimensions

$$\sigma^{(3)} = \frac{NbAb}{SA},\tag{5.2}$$

where Nb is the number of binding sites on the protein. We calculate that in two dimensions,

Figure 5.1: Spherical approximation of generic ESCRT proteins with binding sites around the equator.

$$\sigma^{(2)} = \frac{2Nbrb}{C_E}. \tag{5.3}$$

As both interacting proteins need not have the same number of active sites, we specify $\sigma_1^{(d)}$ and $\sigma_2^{(d)}$ in the calculation below. This becomes necessary if one wished to consider the case of multiple UBDs on ESCRTs while monoubiquitinated cargo only has one ubiquitin bound. We can now calculate the probability of correct orientation in both two and three dimensions, $\rho^{(2)}$ and $\rho^{(3)}$ to be

$$\rho^{(d)} = \sigma_1^{(d)} \sigma_2^{(d)}. \tag{5.4}$$

In the case of ESCRT-cargo binding with only one UBD on the ESCRT we calculate,

$$\rho^{(2)} = \left(\frac{2\left(\frac{2\pi}{3}\right)}{20\pi} \right)^2 \tag{5.5}$$

$$\approx .0044 \tag{5.6}$$

$$\rho^{(3)} = \left(\frac{\pi\left(\frac{2\pi}{3}\right)^2}{\frac{4000\pi}{3}} \right)^2 \tag{5.7}$$

$$\approx .000010823. \tag{5.8}$$

It is clear from these calculations that the probability of correct orientation is much greater in two dimensions than in three.

It is well accepted that orientational constraints are important in protein-protein interactions. We are not the first to propose a simplistic model to deal with dimensionally dependent orientational constraints [20, 4, 16]. Another approach to this problem is to consider rotational diffusion as in [41]. We could indeed use another exit time calculation to compute the time it would take for a particles starting in random orientations to rotationally diffuse until they are in a configuration capable of binding. At the present, we argue that our simple probabilistic model suffices to understand general affects of protein interactions on membranes. More detailed orientational models could be used to determine $\rho^{(d)}$ for a particular problem of interest.

5.1.2 Diffusion in Membranes

We now apply the results of Saffman and Delbrück in [38] to explore 2D and 3D diffusion. Recall that in our model, the diffusion coefficient is given by $\frac{D_1 + D_2}{2}$ where D_1 and D_2 are

the diffusion coefficients of each protein. One can use the Stokes Einstein relation to model diffusion in three dimensions giving,

$$D^{(3)} = \frac{k_B T}{6\pi\mu r}, \tag{5.9}$$

where k_B is Boltzmann's constant, T is absolute temperature, μ is viscosity of the cytoplasm, and r is the radius of the protein. Saffman and Delbrück present the following result for cylindrical proteins imbedded in a membrane,

$$D^{(2)} = \frac{k_B T}{\pi\mu_2\eta} \left(\log \frac{\mu_2\eta}{\mu a} - \gamma \right). \tag{5.10}$$

This formula includes the additional parameters μ_2, η, and γ which represent viscosity of the membrane, thickness of the membrane, and Euler's constant, respectively.

Using the parameters given in [38] which are typical of biological membranes, we calculate the following diffusion coefficients in 3D and 2D which now only depend on the radius of the protein and the thickness of the membrane,

$$D^{(3)} = \frac{4(10^{-14})}{6(10^{-2})\pi r} \frac{\text{cm}^2}{\text{sec}}, \tag{5.11}$$

$$D^{(2)} = \frac{\cdot 4(10^{-14})}{\pi\eta} \left(\log \frac{\eta}{(10^{-2})r} - .5772 \right) \frac{\text{cm}^2}{\text{sec}}. \tag{5.12}$$

It has been reported that biological membranes of lipid bilayers are between 3 and 4nm [31, 1]. For cargo, which span the membrane, we use $\eta = 3.5(10^{-7})$cm. ESCRTs bind to lipids in the bilayer and do not span the membrane. Therefore, for ESCRTs which are not bound to cargo, we use $\eta = 1.25(10^{-7})$cm as ESCRTs interact with only one layer of the lipid bilayer. We calculate that for the generic ESCRT in our model, $D^{(2)} = 5.145(10^{-8}) \left(\frac{\text{cm}^2}{\text{sec}} \right)$ and $D^{(3)} = 2.122(10^{-7}) \left(\frac{\text{cm}^2}{\text{sec}} \right)$. For the cargo in our model, we calculate $D^{(2)} = 3.2183(10^{-8}) \left(\frac{\text{cm}^2}{\text{sec}} \right)$ and $D^{(3)} = 2.122(10^{-7}) \left(\frac{\text{cm}^2}{\text{sec}} \right)$. Figure 5.2 illustrates how the diffusion coefficients depend on the size of the protein. It is clear that proteins move faster in three dimensions than two for protein radii relevant to our model.

5.1.3 K_ds of UBDs in Membranes

We are now ready to convert the K_d for ESCRT binding to ubiquitinated cargo from three dimensions to two dimensions. We do not know the value of $\hat{k}$ nor is it likely to be experimentally measured. As such, we present our K_ds as functions of the true binding rate, $\hat{k}$.
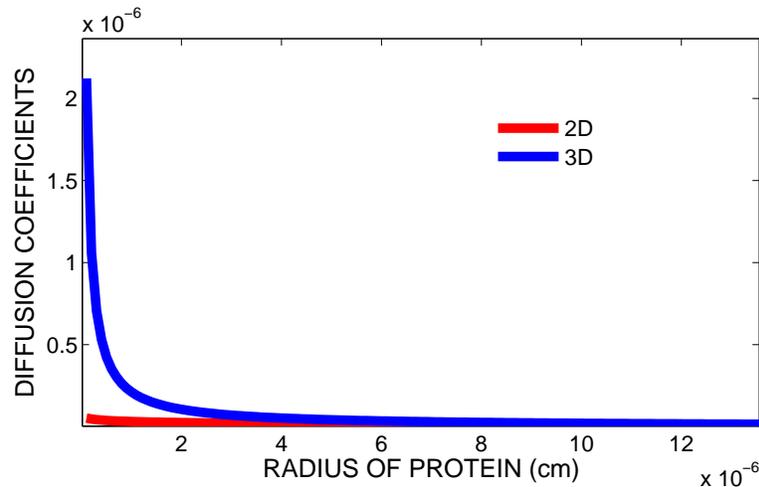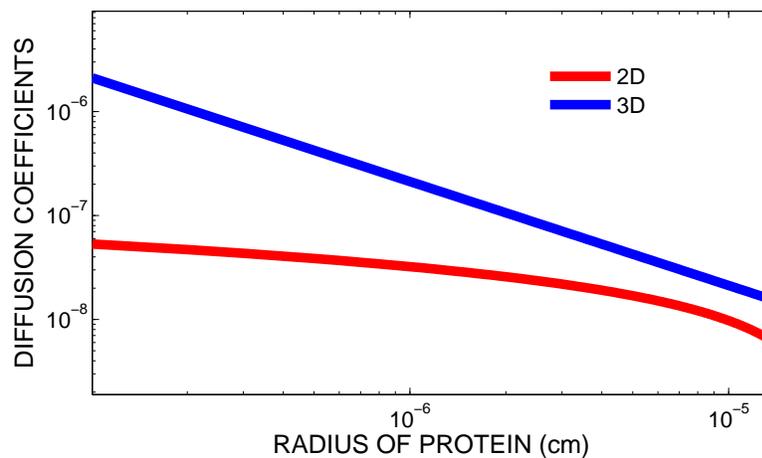
Figure 5.2: (a) Diffusion coefficients in 2D vs. 3D. (b) Same plot as in (a) on a log scale.

We would like to understand what a low affinity three-dimensional K_d such as those reported for ESCRT-cargo binding means in two dimensions. We will therefore transform $K_d^{(3)} = 100\mu M$ to determine its corresponding $K_d^{(2)}$.

Recall again the transformation derived in the previous chapter,

$$K_d^{(2)} = K_d^{(3)} \frac{\rho^{(3)} L \left(1 + \alpha^{(2)} \left(\frac{k^{(2)} h^2}{D^{(2)}}\right)\right) \ln\left(\frac{L}{h}\right)}{\rho^{(2)} \left(\frac{L}{h}\right) \left(1 + \alpha^{(3)} \left(\frac{k^{(3)} h^2}{D^{(3)}}\right)\right)}. \tag{5.13}$$

We summarize the values used for this conversion in Table 5.1. We use the coefficients for α corresponding to periodic boundaries as reported in Section 4.4. Figure 5.3 illustrates the computed two-dimensional K_ds of ESCRT-cargo binding as a function of the true binding rate $\hat{k}$.

Table 5.1: ESCRT UBD K_d Conversion Parameters

$K_d^{(3)}$	$100\mu M$
L	$\sqrt{\pi}\mu m$
h	20nm
$\alpha^{(2)}$	$\frac{1}{12}$
$\alpha^{(3)}$	.1264
k	$\hat{k}\rho^{(d)}(sec^{-1})$
$\rho^{(2)}$	0.0044
$\rho^{(3)}$	0.000010823
$D^{(2)}$	$4.1815(10^{-8})\left(\frac{cm^2}{sec}\right)$
$D^{(3)}$	$2.122(10^{-7})\left(\frac{cm^2}{sec}\right)$
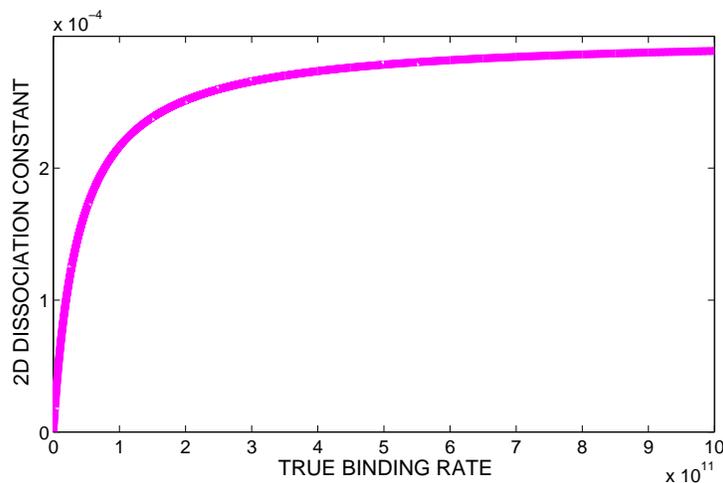
Figure 5.3: $K_d^{(2)}$ corresponding to $K_d^{(3)} = 100\mu$M as a function of true binding rate, $\hat{k}$.

To better understand what these two- and three-dimensional K_ds mean, we employ the following simple dimerization model with n being the number of monomers and n_2 the number of dimers,

$$\frac{dn}{dt} = -2r^{(d)}n^2 + 2k_-n_2, \tag{5.14}$$

$$\frac{dn_2}{dt} = r^{(d)}n^2 - k_-n_2. \tag{5.15}$$

We include the following conservation equation for the total number of monomers, n_t,

$$n_t = n + 2n_2. \tag{5.16}$$

In this equation, we are interested in total numbers as opposed to concentrations. As such, the "dissociation constant" for this dimerization model is the dimensionless parameter, $\frac{s}{r^{(d)}}$. We call this value $KD^{(d)}$ and define it as,

$$KD^{(d)} = K_d^{(d)} L^d, \tag{5.17}$$

where L represents the characteristic length of the endosome as defined in Table 5.1. Figure 5.3 demonstrates that the two-dimensional K_d is a saturating function of the true binding rate, $\hat{k}$. As such, for this dimerization model, we choose $\hat{k} = 10^{12}$, as this leads to a two-dimensional dissociation constant that is in the saturating region. Note that the following dimerization results will be amplified for smaller values of $\hat{k}$ as the converted two-dimensional K_d decreases as a function of $\hat{k}$.

We compute the number of dimers at steady state in both dimensions. The two-dimensional domain of this model is a square approximation of the endosome while the three-dimensional domain is a cube of size L^3. Figure 5.4 shows a plot of the steady state of n_2 as a function of L in both two and three dimensions. It demonstrates that dimers form much more readily in two dimensions than in three dimensions. A log scale is used to plot the steady state of n_2 as a function of L in Figure 5.4(b). This illustrates that there are approximately 10-fold more dimers on membranes than in solution. This suggests that ESCRT-cargo bonds are more stable than previously thought. It is clear from our model that using a three-dimensional K_d to describe binding events in two dimensions is not appropriate. Simply because ESCRT UBDs have weak affinity in solution does not automatically mean their bonds are weak on membranes. This result supports the use of stable bonds between ESCRT and cargo in our cellular automata model in Chapter 3.

A similar calculation could be performed to approximate the binding affinities of ESCRT-ESCRT interactions. It has been reported that ESCRT-ESCRT interactions are much stronger than that of ESCRT-cargo interactions mediated by ubiquitin. For example, experiments show that ESCRT-I and ESCRT-II interact with a K_d of 27nM [14]. We argue that a similar enhancement of affinity will be observed for ESCRT-ESCRT bonds on membranes and that the two-dimensional K_d will again be less than that of ESCRT-cargo bonds. We make this argument as $K_d^{(3)}$ appears linearly in the numerator of our conversion equation (5.13). Therefore, our model suggests that on membranes, ESCRT-ESCRT affinity is greater than that of ESCRT-cargo interactions which appears necessary for proper sorting of cargo and recycling of ESCRTs. [28]. While reported three-dimensional K_ds are not sufficient, on their own, to make claims about binding affinity on membranes, they likely suggest the hierarchy of affinities for ESCRT binding events on membranes. In summary,
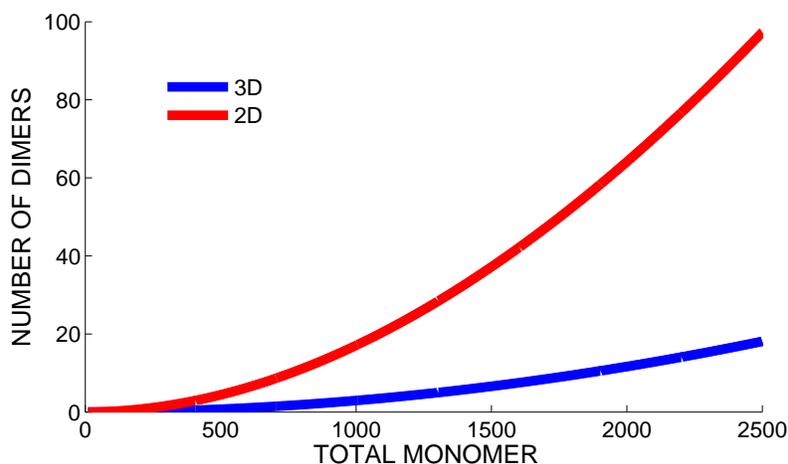
(a)
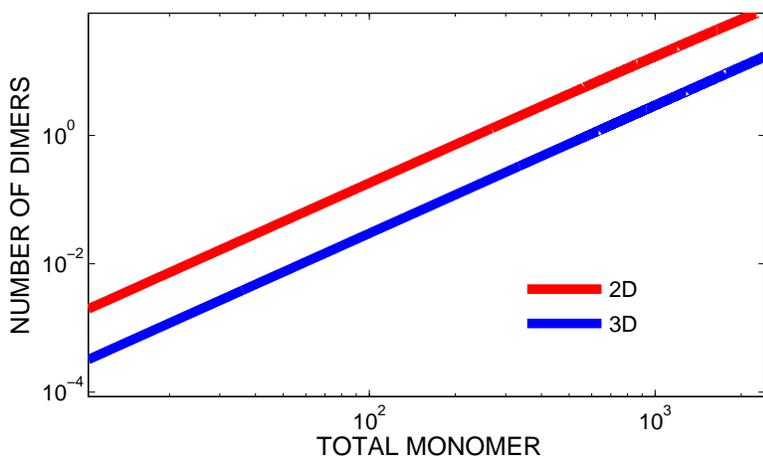
(b)

Figure 5.4: (a) Comparison of dimerization in 2D and 3D with reported $K_d^{(3)}$ and computed $K_d^{(2)}$. (b) Same plot as in (a) on a log scale.

we predict that constraining ESCRT proteins to two dimensions enhances their binding affinities from what is observed in solution.

5.2   ESCRT-cargo Network Formation Model

We now present a new model for ESCRT-mediated ubiquitin-dependent protein sorting using the rate of two-dimensional protein interaction determined in Chapter 4,

$$r^{(2)} \quad = \quad \frac{k \left(\frac{h}{L}\right)^2}{1 + \alpha^{(2)} \left(\frac{kh^2}{D}\right) \ln \left(\frac{L}{h}\right)}. \tag{5.18}$$

The results from our cellular automata model of cargo sorting in Chapter 3 suggest that a

flexible ESCRT-cargo network with given rules is sufficient to efficiently sort cargo. As such, we now aim to better understand the evolution of ESCRT-cargo networks in a continuous time model using rates of diffusion and binding instead of the propensity parameters used in our cellular automata model. We make the assumption that a single ESCRT is recruited to the endosomal membrane by a single cargo. As such, we define a monomer as one ESCRT-cargo pair. In our cellular automata model, we do not allow for ESCRT-cargo interactions to break once in networks of more than three proteins. The calculation of the two-dimensional K_d for ESCRT-cargo binding presented in the previous section suggests that the stability of the ESCRT-cargo interactions on membranes is much greater than previously thought. Under the assumption that networks of more than three proteins are, indeed, stable, we do not include unbinding in the current model. Our model considers the polymerization reaction of ESCRT-cargo monomers into ESCRT-cargo networks. We assume that each monomer has 5 ESCRT-ESCRT interaction sites to mimic our cellular automata model. Let n_i be the number of networks of exactly i monomers. Our ESCRT-cargo network polymerization model is

$$\frac{dn_i}{dt} = \sum_{j+m=i, j\neq m} r_{jm} n_j n_m + \sum_{2j=i} r_{jj}\binom{n_j}{2} - \sum_{j\neq i} r_{ji} n_j n_i - r_{ii}\binom{n_i}{2}. \tag{5.19}$$

Using (5.18), the binding rates, r_{jm}, can be written as

$$r_{jm} = \frac{\hat{k}\rho_{jm}\left(\frac{h_{jm}}{L}\right)^2}{1 + \alpha^{(2)}\left(\frac{\hat{k}\rho_{jm}h_{jm}^2}{D_{jm}}\right)\ln\left(\frac{L}{h_{jm}}\right)}, \tag{5.20}$$

where, $h_{jm} = h_j + h_m$, $D_{jm} = \frac{D_j+D_m}{2}$, and $\rho_{jm} = \sigma_j\sigma_m$. $L = \sqrt{\pi}$ as in the previous section. It remains to determine the radius of networks, diffusion coefficient, and binding ratio as functions of the size of the network. We do this as follows with ESCRTs and cargoes modeled as in the previous section.

- **Radius**

  In our model, both ESCRTs and cargoes have radii of 10nm. We make the assumption that all networks are circular. This assumption seems reasonable as our cellular automata model showed condensed structures rather than branched linear structures and since networks are invaginated to form ILVs. With this circular assumption, we approximate an ESCRT-cargo monomer occupies an area of 200π nm^2 on a membrane. Similarly we approximate that a dimer occupies $400\pi nm^2$. In general, a network of size i occupies $200i\pi nm^2$. We therefore calculate

  $$h_i = \sqrt{200i\pi} nm. \tag{5.21}$$

- **Diffusion Coefficients**

  We employ (5.12) with $\eta = 3.5nm$ as cargo spans the entire membrane. We, therefore, use the following two-dimensional diffusion coefficient which only depends on the size of the network.

  $$D_i = \frac{\cdot 4(10^{-14})}{3.5(10^{-7})\pi} \left(\log \frac{3.5(10^{-7})}{(10^{-2})h_i} - .5772 \right) \frac{\text{cm}^2}{\text{sec}}. \qquad (5.22)$$

- **Binding Ratio**

  We again assume that each binding site of the ESCRT represents $\frac{1}{15}$ of the circumference of the equator of the ESCRT. This gives the radius of each binding site $rb = \frac{2\pi}{3}$. The circumference of a network is given by, $C_E = 2\pi h_i$. We must know the number of free binding sites for a network of a given size. We employ our cellular automata model to estimate this number. We alter the model presented in Chapter 3 to include only ESCRT-cargo pairs in aggregates. There was a slight chance in the previous model of a single ESCRT being incorporated into an aggregate. Figure 5.5 shows the average number of free binding sites for networks of a given size for 100 simulations. We observe that the number of free binding sites varies linearly with network size according to the following equation,

  $$Nb(i) = 1.2621i + 4.9450. \qquad (5.23)$$

  We now compute the binding ratio as,

  $$\sigma_i = \frac{Nb_i rb}{\pi h_i}. \qquad (5.24)$$

We have now represented all the parameters in the model as functions of network size. To analyze this system of differential equations, we use a Gillespie algorithm to compute the expected sorting time. We define the sorting time as the time it takes for the average network size to reach what is considered a critical vesicle size, C. We defined $C = 20$ in our cellular automata model which corresponds to $C = 10$ in this model. Therefore, we define sorting time to be the expected time it takes for the average network size to be C.

5.2.1 Results

We specify the total number of monomers in our system to be, $C_t = 100$, and simulate (5.19) by calculating the possible rates at each Gillespie time step. Figure 5.6 depicts a sample of our Gillespie simulation. As we do not know the value of the true binding rate, we compute the expected sorting time for different values of $\hat{k}$. We fit a lognormal distribution
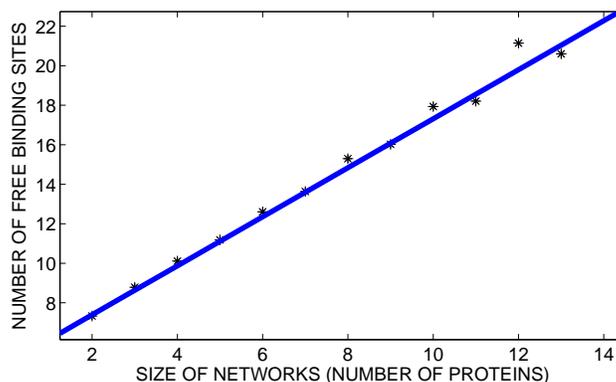
Figure 5.5: Average number of free binding sites for networks of different sizes computed from 100 simulations of the cellular automata model.

to the histogram of 5000 simulated sorting times and compute the average sorting time for each value of $\hat{k}$. Figure 5.7 provides the histogram for $\hat{k} = 10^{12} \text{sec}^{-1}$. Figure 5.8 illustrates how sorting time is affected by the true binding rate. We observe a change in behavior around $\hat{k} = 10^{-8} \text{sec}^{-1}$. For $\hat{k}$ values less than 10^{-8}sec^{-1}, the actual binding of the ESCRTs is the rate limiting step. For $\hat{k}$ greater than 10^{-8}sec^{-1}, diffusion is the rate limiting step and the true binding rate does not affect the sort times in the region. We anticipate that these reactions are diffusion limited as proteins with an affinity for one another are very likely to bind if their binding sites are in sufficient proximity and oriented together. There are no data to suggest that this is not the case in this system. If, indeed, $\hat{k}$ is such that this polymerization reaction is diffusion limited, sorting is very fast. With a total number of 100 cargo on the endosome, our calculations suggest that the average network size reaches the critical size of 20 proteins in approximately 0.27 seconds. We anticipate this sorting time is less than what should be observed experimentally. It seems plausible that including all of the ESCRTs and their specific interactions with one another will slow the sorting process down. Conversely, we predict that considering the process of vesicle formation will speed up sorting of the remaining cargo as membrane is being removed from the endosome. Our two-dimensional reaction rates are dependent on the length of the domain. As length decreases, the reaction rates increase which would lead to increased sorting speeds as ILVs form. As more experimental data come available, it will be possible to include more detail in our network polymerization model and better predict sorting time.
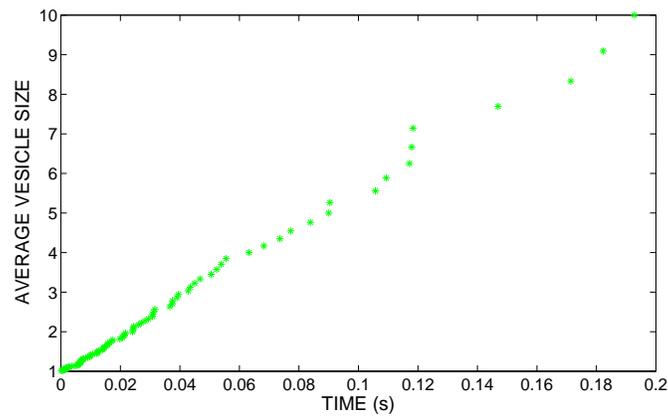
Figure 5.6: Sample Gillespie simulation of ESCRT-cargo network polymerization with time plotted against average vesicle size.
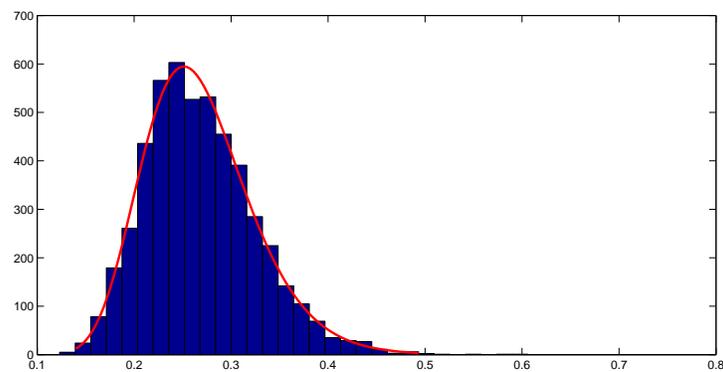
Figure 5.7: Histogram of simulated sorting times with $\hat{k} = 10^{12}$.
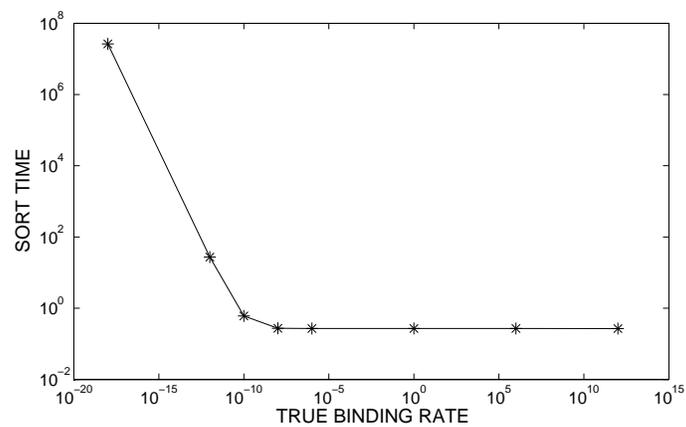
Figure 5.8: Average sorting times as a function of true binding rate, $\hat{k}$.

5.3 Discussion

In this chapter, we applied our results of dimensional dependence of protein interactions to further understand ESCRT-mediated protein sorting. We first applied our 3D to 2D transformation of dissociation constants to a model system of ESCRT-cargo interaction mediated by a single UBD. The results from this section are two fold. First, correct transformations of K_ds are very involved. Second, correct transformations are very important. Our transformation suggests that what is considered a weak bond in three dimensions can be greatly enhanced when the proteins interact on membranes. Our model transformation suggests that ESCRT UBDs may have higher affinities than previously thought, suggesting the possibility of stable ESCRT-cargo interactions on membranes. Our work suggests again that dimension is very important in binding reactions. For any soluble proteins that are able to associate with membranes and interact with other proteins, more than the three-dimensional dissociation constants must be known in order to understand protein interactions on membranes. We predict that coat proteins, such as COP and clathrin, which readily polymerize on membranes but not in cytosol will similarly show enhanced binding affinities on membranes.

In this chapter, we also applied our computed two-dimensional reaction rates to a differential equation model of ESCRT-cargo network polymerization. Using a Gillespie algorithm, we compute the expected average sorting time. We suggest that ESCRT-cargo network formation is diffusion limited. As such, it appears that sorting is a very quick process on endosomes, on the order of seconds. Understanding ESCRT polymerization is important not only for protein sorting but also for understanding the process of HIV budding. It is possible that similar models could be used to address ESCRT interactions in other settings.

CHAPTER 6

RAD18 REGULATION IN DNA DAMAGE TOLERANCE

6.1 DNA Damage Tolerance by Translesion Synthesis Polymerases

As cells frequently experience damage to their DNA due to environmental influences, it is necessary that they be employed with pathways that are able to prevent problems in DNA replication and mutagenesis. DNA lesions are one type of damage cells must address. The DNA damage tolerance (DDT) pathways function as a way for cells to bypass DNA lesions during replication. The specific type of DDT pathway we are presently interested in is bypass accomplished by translesion synthesis (TLS) polymerases. We now provide an overview of this pathway which is described in more detail by Zeman et al. in [51]. DDT is controlled by the ubiquitination of the replicative clamp proliferating cell nuclear antigen (PCNA). The primary E3 ligase responsible for this ubiquitination is Rad18. Rad18 is also involved in recruiting the correct TLS polymerase for a specific type of damage. It is clear that Rad18 requires tight control. Over-expression leads to inappropriate recruitment of TLS polymerases leading to potential errors in DNA replication. As such, it is not surprising that Rad18 has been found to be upregulated in some cancers. In addition, under-expression of Rad18 increases sensitivity to DNA damage.

DDT pathways involving TLS polymerases are tailored to respond appropriately to different types of lesions but how specificity is achieved has been poorly understood. It is known that part of this specificity involves recruitment of the correct TLS polymerase. However, not only is the recruitment mechanism unclear, it also appears there are other events involved in the responses to specific types of damage. One such event is replication fork reversal. Data suggest that two additional E3 ligases are involved in these damage specific events: helicase-like trancreption factor (HLTF) and SNF2 histone linker plant homeodomain RING helicase (SHPRH). Both HLTF and SHPRH are thought to be able to induce replication fork reversal. In addition, HLTF and SHPRH are competitive binding

partners for Rad18. It is thought that HLTF/Rad18 and SHPRH/Rad18 complexes are important for proper PCNA ubiquitination and accurate TLS polymerase recruitment under different circumstances.

Multiple studies have revealed that HLTF constitutively binds to Rad18 whereas SHPRH/Rad18 complexes are strongly induced when cells are treated with methyl methanesulfonate (MMS) [26, 33]. It is also known that SHPRH exists in much lower concentrations in the cell than does HLTF. It is the switch from HLTF/Rad18 to SHPRH/Rad18 that we address in this work.

6.2 Specific DDT Response to MMS

Experimental evidence shows that HLTF is degraded upon MMS treatment. As HLTF and SHPRH compete to bind to Rad18, it would seem that this degradation should induce SHPRH/Rad18 complexes. However, this is not the case. While necessary, HLTF degradation is not sufficient to induce SHPRH/Rad18 complex formation. Zeman et al. recently introduced an additional MMS damage inducible step, deubiquitination of Rad18 [51]. Rad18 is shown to exist in two forms in the cell: Rad18 and ubiquitinated Rad18, Rad18u. Rad18u is unable to form complexes with HLTF and SHPRH and perform other roles in the DDT pathway. Therefore, Rad18 is active while Rad18u is inactive. Additionally, it was observed that Rad18 is able to form a stable complex with Rad18u therefore introducing another competitive binding partner for Rad18. It appears that the ubiquitin binding ZnF (UBZ) domain of Rad18 is required for interaction of Rad18 with HLTF, SHPRH, and Rad18u. The proposed functional model for MMS-specific DDT is depicted in Figure 6.1.
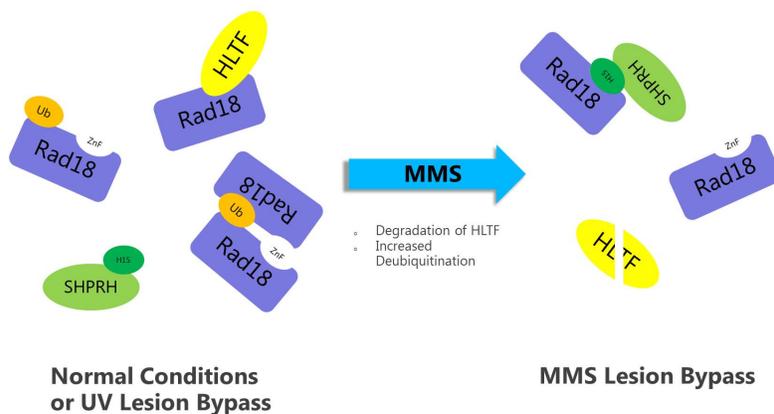
Figure 6.1: MMS-specific DNA damage response. Adapted from [51].

As a further level of ubiquitin-dependent control of Rad18, Rad18 is capable of autoubiq-uitination in complex with the E2 conjugating enzyme, Rad6 [32]. Miyase et al. suggest that two Rad18 molecules, in complex with Rad6, interact via their ZnF domain and monoubiquitinate each other as shown in Figure 6.2. It is of importance to note that autoubiquitination introduces a fourth competitive binding parter for the ZnF domain of Rad18.

Our primary question of interest is how MMS induces SHPRH/Rad18 complexes. We know that both HLTF degradation and increased deubiquitination are necessary but niether is sufficient. Zeman et al. suggest a few hypotheses for how deubiquitination is acheived in the system. They show that the apparent reduction of Rad18u is not caused by degradation of Rad18u suggesting that less of the total Rad18 in the cell is ubiquitinated after MMS treatment. This leads to two hypotheses for how MMS may induce the perceived increase in deubiquitination. First, Rad18's autoubiqutinating function may be reduced after MMS treatment leading to less Rad18u in the cell. Second, MMS may cause Rad18 to be actively targeted by DUBs thereby decreasing Rad18u. We seek to understand the general behavior of this system as well as the roles of autoubiquitination, the Rad18/Rad18u complex, and to determine any other necessary conditions to induce SHPRH/Rad18 complexes.

6.3   ODE Models

In our mathematical models, we assume that the total concentration of SHPRH and Rad18, S_t and R_t, respectively, are held fixed as neither upregulation nor degradation are presumed to play a role in the response to MMS. As we do not wish to specifically model the mechanism by which HLTF is degraded upon MMS treatment we also assume a constant
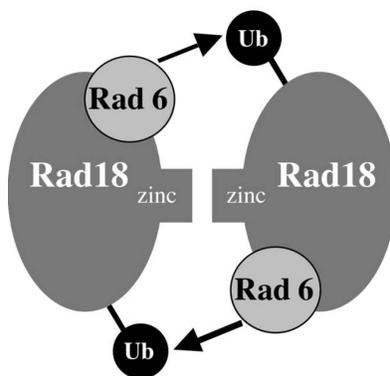
Figure 6.2: Proposed model for Rad18 autoubiquitination. Reused with permission from [32].

total concentration of HLTF, H_t, and vary this concentration as a parameter in our system to mimic MMS induced degradation.

As we wish to consider the effects of the Rad18/Rad18u complex and the effects of autoubiquitination we present three different ODE systems of increasing complexity. Model 1 includes only competitive binding of SHPRH and HLTF for Rad18. Model 2 adds the additional competitive binding partner of the Rad18/Rad18u complexes. Model 3 also includes the feature of Rad18 autoubiquitination.

The variables in our models include the following concentrations: Rad18 (R), ubiquitinated Rad18 (Ru), Rad18/Rad18u complex (RuR), Rad18/Rad18 complex (R_2), HLTF (H), SHPRH (S), HLTF/Rad18 complex (H_R), and SHPRH/Rad18 complex (S_R). Descriptions of the rates included in our models are given in Table 6.1. Unless specified, the rates are pertinent to all three models.

In all of our models, the following conservation equations hold:

$$H_t = H + H_R, \tag{6.1}$$
$$S_t = S + S_R. \tag{6.2}$$

We also use a conservation equation for the total amount of Rad18; however, this is specified for each model as it is different in every case. We now outline our ODE models.

6.3.1 Model 1

Model 1 is considered the base model for this system. It includes modeling ubiquitination of Rad18 as well as competitive binding of HLTF and SHPRH to Rad18. The system is

Table 6.1: Description of parameters

a	rate of Rad18 ubiquitination (Models 1&2)
a_1	binding rate of Rad18/Rad18 (Model 3)
a_2	unbinding rate of Rad18/Rad18
a_3	rate of autoubiquitination of Rad18
b	rate of Rad18u deubiquitination
c	binding rate of Rad18/Rad18u
d	unbinding rate of Rad18/Rad18u
e	binding rate of HLTF/Rad18
f	unbinding rate of HLTF/Rad18
g	binding rate of SHPRH/Rad18
h	unbinding rate of SHPRH/Rad18
R_t	total Rad18
H_t	total HLTF
S_t	total SHPRH

$$\frac{dR}{dt} = -aR + bRu - eRH + fH_R - gRS + hS_R, \tag{6.3}$$

$$\frac{dRu}{dt} = aR - bRu, \tag{6.4}$$

$$\frac{dH}{dt} = -eRH + fH_R, \tag{6.5}$$

$$\frac{dH_R}{dt} = eRH - fH_R, \tag{6.6}$$

$$\frac{dS}{dt} = -gRS + hS_R, \tag{6.7}$$

$$\frac{dS_R}{dt} = gRS - hS_R. \tag{6.8}$$

In this model, the conservation equations for R is

$$R_t = R + Ru + H_R + S_R. \tag{6.9}$$

We determine S and H using the conservation equations and can readily solve for the steady state values of Ru, S_R, and H_R:

$$Ru = \frac{a}{b}R, \tag{6.10}$$

$$H_R = \frac{RH_t}{R + \frac{f}{e}}, \tag{6.11}$$

$$S_R = \frac{RS_t}{R + \frac{h}{g}}. \tag{6.12}$$

It is conventient to define the following new parameters:

$$B = \frac{b}{a}, \tag{6.13}$$

$$K_2 = \frac{f}{e}, \tag{6.14}$$

$$K_3 = \frac{h}{g}. \tag{6.15}$$

B is a ratio of deubiquitination to ubiquitination while K_2 and K_3 represent the dissociation constants for SHPRH/Rad18 binding and HLTF/Rad18 binding, respectively.

It remains to solve for the steady state of R using the conservation equation which can now be expressed as a function of R alone,

$$R_t = R + \frac{R}{B} + \frac{RH_t}{R + K2} + \frac{RS_t}{R + K3}. \tag{6.16}$$

We explain our method for solving this equation following the descriptions of Models 2 and 3.

6.3.2 Model 2

Model 2 contains all the features of Model 1 with the addition of Rad18/Rad18u complexes. This introduces a third competitive binding event into the model. The equations are

$$\frac{dR}{dt} = -aR + bRu - cR \cdot Ru + dRuR - eRH + fH_R - gRS + hS_R, \quad (6.17)$$

$$\frac{dRu}{dt} = aR - bRu - cR \cdot Ru + dRuR, \quad (6.18)$$

$$\frac{dRuR}{dt} = cR \cdot Ru - dRuR, \quad (6.19)$$

$$\frac{dH}{dt} = -eRH + fH_R, \quad (6.20)$$

$$\frac{dH_R}{dt} = eRH - fH_R, \quad (6.21)$$

$$\frac{dS}{dt} = -gRS + hS_R, \quad (6.22)$$

$$\frac{dS_R}{dt} = gRS - hS_R. \quad (6.23)$$

In this model, the conservation equation for R is

$$R_t = R + Ru + 2RuR + H_R + S_R. \quad (6.24)$$

At steady state we find

$$Ru = \frac{a}{b}R, \quad (6.25)$$

$$RuR = \frac{a}{b}\frac{c}{d}R^2, \quad (6.26)$$

$$H_R = \frac{H_t R}{R + \frac{f}{e}}, \quad (6.27)$$

$$S_R = \frac{S_t R}{R + \frac{h}{g}}. \quad (6.28)$$

Model 2 introduces a fourth parameter K_1 representing the dissociation constant of Rad18/Rad18u binding which we define

$$K_1 = \frac{d}{c}. \quad (6.29)$$

With the above steady state values, it remains to solve the following equation for R:

$$R_t = R + \frac{R}{B} + 2\frac{R^2}{BK_1} + \frac{RH_t}{R + K2} + \frac{RS_t}{R + K3}. \quad (6.30)$$

6.3.3 Model 3

Model 3 includes all the features of Models 1 and 2 with the addition of autoubiqui-tination which is modeled after the hypothesis of Miyase et al. in [32]. The equations are

$$\frac{dR}{dt} = -2a_1 R^2 + 2a_2 R_2 + bRu - cR \cdot Ru + dRuR \tag{6.31}$$
$$-eRH + fH_R - gRS + hS_R,$$

$$\frac{dR_2}{dt} = a_1 R^2 - a_2 R_2 - a_3 R_2, \tag{6.32}$$

$$\frac{dRu}{dt} = 2a_3 R_2 - bRu - cR \cdot Ru + dRuR, \tag{6.33}$$

$$\frac{dRuR}{dt} = cR \cdot Ru - dRuR, \tag{6.34}$$

$$\frac{dH}{dt} = -eRH + fH_R, \tag{6.35}$$

$$\frac{dH_R}{dt} = eRH - fH_R, \tag{6.36}$$

$$\frac{dS}{dt} = -gRS + hS_R, \tag{6.37}$$

$$\frac{dS_R}{dt} = gRS - hS_R. \tag{6.38}$$

The conservation equation for R in Model 3 is

$$R_t = R + 2R_2 + Ru + 2RuR + H_R + S_R. \tag{6.39}$$

At steady state, we find

$$R_2 = \frac{a_1}{a_2 + a_3} R^2, \tag{6.40}$$

$$Ru = 2\frac{a_3}{b}\frac{a_1}{a_2 + a_3} R^2, \tag{6.41}$$

$$RuR = 2\frac{c}{d}\frac{a_3}{b}\frac{a_1}{a_2 + a_3} R^3, \tag{6.42}$$

$$H_R = \frac{H_t R}{R + \frac{f}{e}}, \tag{6.43}$$

$$S_R = \frac{S_t R}{R + \frac{h}{g}}. \tag{6.44}$$

Model 3 involves a new ratio of deubiquitination to ubiqutination, B_2, as well as a new parameter, with units of a dissociation constant, which we will call K_4. We define

$$B_2 = \frac{b}{a_3}, \tag{6.45}$$

$$K_4 = \frac{a_2 + a_3}{a_1}. \tag{6.46}$$

With the above steady state values, it remains to solve the following equation for R:

$$R_t = R + 2\frac{R^2}{K_4} + 2\frac{R^2}{B_2 K_4} + 4\frac{R^3}{B_2 K_1 K_4} + \frac{R H_t}{R + K2} + \frac{R S_t}{R + K3}. \tag{6.47}$$

6.3.4 Solving for R

In each of our models, we need to solve a nonlinear conservation equation for R and a function of H_t. However, since H_t appears linearly in every case, we are able to solve for H_t as a function of R.

For Model 1, we find

$$H_t = \left(R_t - R - \frac{R}{B} - \frac{R S_t}{R + K3} \right) \left(\frac{R + K2}{R} \right). \tag{6.48}$$

For Model 2, we find

$$H_t = \left(R_t - R - \frac{R}{B} - \frac{2R^2}{B K_1} - \frac{R S_t}{R + K3} \right) \left(\frac{R + K2}{R} \right). \tag{6.49}$$

For Model 3, we find

$$H_t = \left(R_t - R - \frac{2R^2}{K4} - \frac{2R^2}{B_2 K_4} - \frac{4R^3}{B_2 K_1 K_4} - \frac{R S_t}{R + K3} \right) \left(\frac{R + K2}{R} \right). \tag{6.50}$$

For a given parameter set, we calculate H_t for a range of R values from 0 to R_t. Since H_t is a monotone decreasing function of R, it follows that there is unique solution R for each value of H_t.

6.3.5 Method of Analysis

We do not know any of the parameters or concentrations in this system. However, for a model to be acceptable, it must be that increasing deubiquitination and decreasing HLTF leads to formation of the SHPRH/Rad18 complex while neither of these actions on their own do. Therefore, we compute the steady states of our models for a range of H_t and a range of B (or B_2 in Model 3) and plot the proportion of S_t that is in complex with Rad18. In our plots, we wish to see the behavior shown in Figure 6.3. That is, decreasing H_t while holding B fixed should leave S_R unchanged and increasing B while holding H_t fixed should also leave S_R unchanged. However, decreasing H_t and increasing B should result in an increase in S_R.

As previously mentioned, the parameters for this system are unknown. It is known that SHPRH exists in much lower quantities in the cell than does HLTF [26]. Under this assumption, we pick a value for total SHPRH concentration, S_t, and vary H_t to values much greater than S_t. The main parameters for this system are the dissociation constants for the
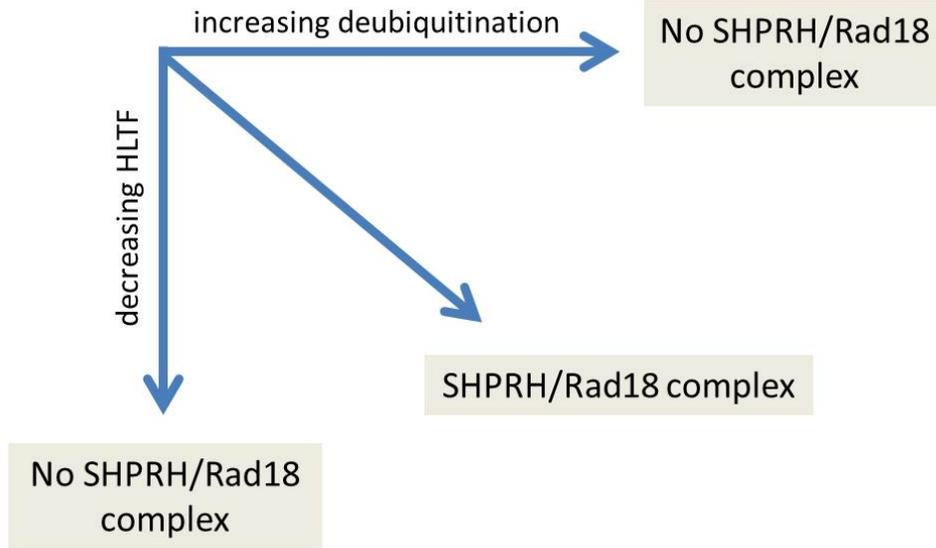
Figure 6.3: Method of analysis.

different competitive binding events. We search over a range of dissociation constants for the desired behavior of the system. All parameters used in the results presented are listed with the figures.

As one final consideration, Zeman et al. report that under normal conditions they observe 25% of Rad18 is ubiquitinated [51]. Therefore, $0.25R_t = Ru + RuR$. Using this fact, we can calculate possible starting positions in our parameter space for the B and H_t values before MMS treatment. This gives us starting positions from which to look for the behavior depicted in Figure 6.3.

We calculate the B value necessary to meet this condition for each model.

In Model 1, we compute

$$B = \frac{R}{0.25R_t}.$$ (6.51)

In Model 2, we compute

$$B = \frac{K_1R + R^2}{0.25dR_tK_1}.$$ (6.52)

In Model 3, we compute

$$B_2 = \frac{2K_1R^2 + 2R^3}{0.25R_tK_1K_4}.$$ (6.53)

We use these values of B in the equations for H_t in order to compute the steady state of the system. The possible starting positions are shown in black on each plot of SHPRH in our results.

6.4   Results

6.4.1   Necessity of K_d Measurements

Figure 6.4 gives plots of S_R as functions of H_t and B with all other parameters fixed as specified. In each model, we observe the desired behavior that both HLTF degradation and increasing deubiquitination are necessary but neither is sufficient on its own to induce formation of SHPRH/Rad18 complexes as shown in Figure 6.4. However, in each model, there are different sets of dissociation constants for which we observe these experimental results. The calculated possible starting positions are shown in black on each plot.

Our results suggest that measurements of the dissociation constants are necessary in order to assess the validity of the current proposed models. As we observe the desired behavior in all three models, we cannot yet discuss the importance of the Rad18/Rad18u complex or autoubiquitination in this system. Knowing the true dissociation constants for this system would allow us to define one parameter space for each model which may or may not exhibit the desired behavior.

Further, our results suggest that it is not enough to simply consider the comparative strengths of each binding event. Figure 6.5 shows plots with equal dissociation constants for SHPRH/Rad18 complexes and HLTF/Rad18 complexes in Model 1. For $K_2 = K_3 = 150$ we observe the desired behavior while for $K_2 = K_3 = 1$ we do not. In the latter case, there is no possible starting position which demonstrates the desired behavior. These results further argue for the necessity of measuring the dissociation constants for each Rad18 binding event in the system.

6.4.2   Deubiquitination Rate Prediction

To observe the desired behavior in all three models, the starting condition suggests almost no deubiquitination, $B \sim 0$, under normal circumstances. As this result is consistent among all three models and all parameter ranges we considered, it suggests that prior to MMS treatment essentially no DUBs are functioning in this system. This can be seen from the possible starting positions shown in Figures 6.4 and 6.5.

6.5   Discussion

The goal of these models was to gain insight into the function of the Rad18/Rad18u complex as well as to better understand the role of autoubiquitination. Unfortunately, we are not able to address these issues with our current model. In order to decide on the validity of the current functional models, our results suggest that experimental measurements of the dissociation constants for Rad18 binding events are necessary. With a given set of measured
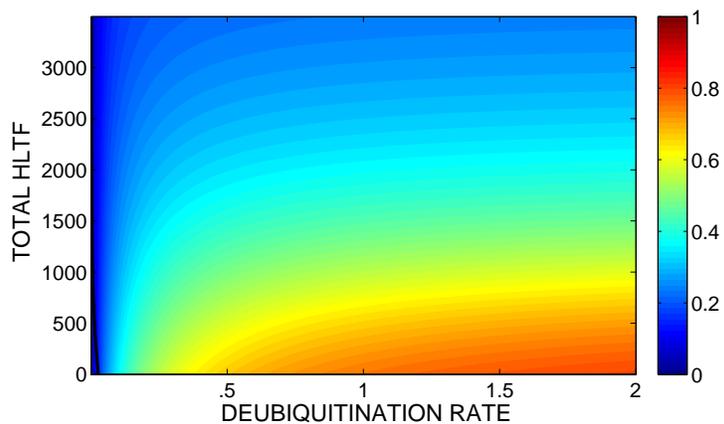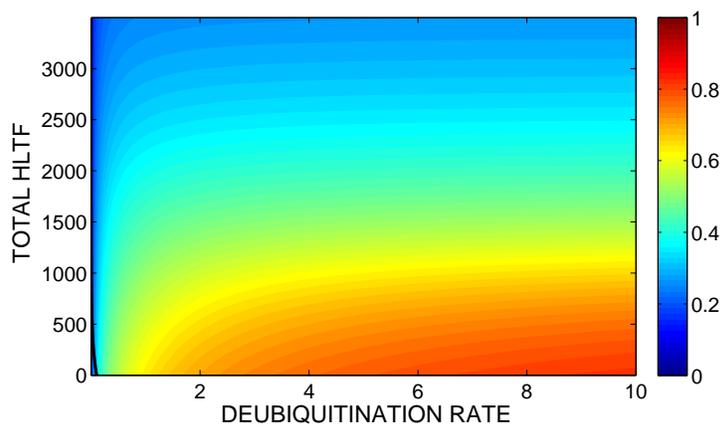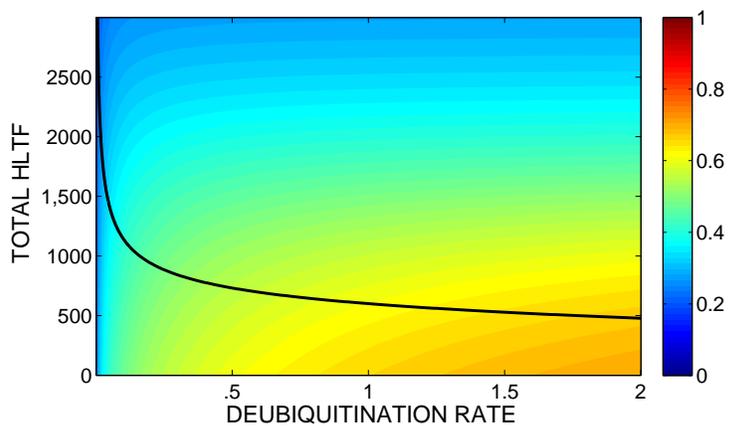
Figure 6.4: Proportions of total SHPRH in complex with Rad18. (a) Model 1 with $S_t = 100$, $R_t = 1000$, $K_2 = K_3 = 150$. (b) Model 2 with $S_t = 100$, $R_t = 1000$, $K_1 = K_2 = K_3 = 100$. (c) Model 3 with $S_t = 100$, $R_t = 1000$, $K_1 = 10$, $K_2 = 25$, $K_3 = 30$, $K_4 = 100$.
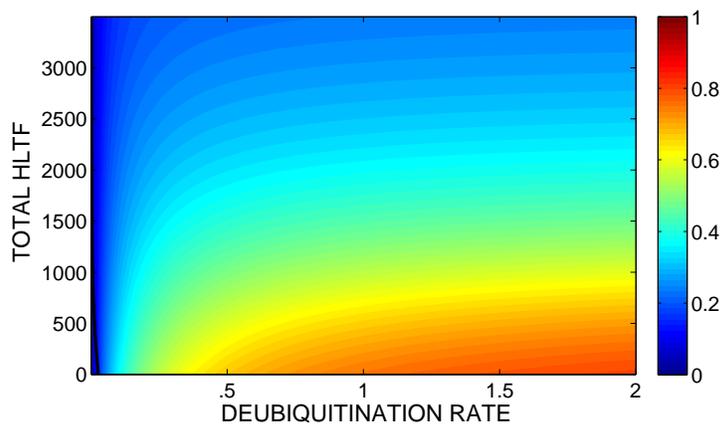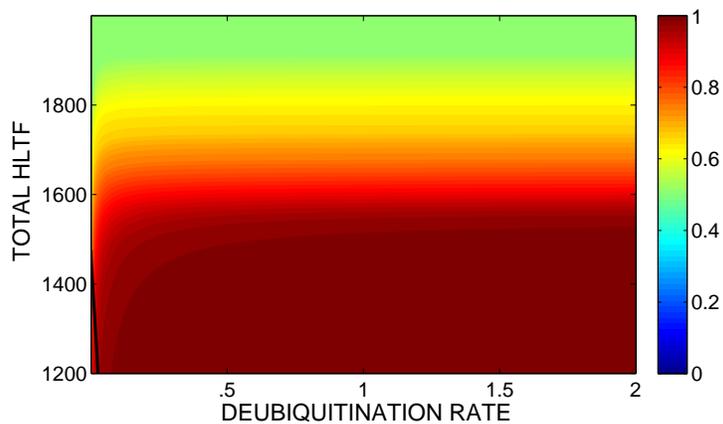
Figure 6.5: Proportions of total SHPRH in complex with Rad18 in Model 1 with equal dissociation constants. (a) Model 1 with $S_t = 100$, $R_t = 1000$, $K_2 = K_3 = 150$. (b) Model 1 with $S_t = 100$, $R_t = 1000$, $K_2 = K_3 = 1$.

dissociation constants, we will be able to better distinguish between our three models and make claims about the importance of Rad18/Rad18u complex and autoubiquitination. Our results suggest that knowing the dissociation constants is crucial to understanding the system. Our results also suggest that under normal circumstances there are essentially no DUBs acting on Rad18u. It is of note that in this system ubiquitin is regulating Rad18 without relying on degradation, unlike the regulation of Fur4 which relies on ubiquitin dependent degradation. This system is also demonstrates ubiquitin dependent regulation of soluble proteins which is necessary for a specific response to MMS treatment.

CHAPTER 7

CONCLUSION

Proper regulation of proteins is at the core of cellular physiology. As such, lack of regulation leads to pathophysiology and, therefore, disease. Ubiquitin is one of the key players in protein regulation in most tissues and acts as both a post-translational protein modification and as a binding partner to achieve regulation. In this dissertation, the primary goal is to better understand some of ubiquitin's roles. Specifically, we considered ubiquitin-dependent regulation of Fur4, ubiquitin-dependent protein sorting by ESCRTs, and ubiquitin-dependent regulation of Rad18 in the DNA damage tolerance pathway.

We first explored the regulation of the yeast uracil permease, Fur4. We modeled the mechanism by which Fur4 is tagged with ubiquitin, Fur4 trafficking between the plasma membrane and the early endosome, and endosome maturation leading to deletion of Fur4. We wanted to understand the role of the Rsp5/Ubp2 complex which is localized to the endosome. Our results suggest that the Rsp5/Ubp2 complex, which we refer to as the Endosomal Quality Control and Retention System (EQRS), is crucial for proper regulation of Fur4. Without the deubiquitination and retention functions of the EQRS, we do not observe the marked downregulation of Fur4 upon increasing extracellular uracil. Further, our results suggest that the pool of endosomal Fur4 remains nearly constant over a range of extracellular uracil concentrations. We believe this should also be observed experimentally. This pool allows for rapid localization of Fur4 to the plasma membrane upon a shift to starvation conditions. In general, our results suggest that ubiquitination and deubiquitination are equally important. In this specific system, we observe that the localization of ubiquitinating enzymes and deficiency of deubiquitinating enzymes at the plasma membrane allow for quick removal of proteins which are misfolded or not needed. Subsequently, the EQRS, with both ubiquitinating enzymes and deubiquitinating enzymes, provides a method by which the cell can "double check" whether or not it should degrade a given protein. As ubiquitin is the degradation signal in this pathway, it is through its ubiquitin tag that Fur4 is retained on the endosome by interacting with the EQRS. This leads to its inclusion in

ILVs and subsequent degradation. In this system, we observe that ubiquitin acts as both a post-translational modification and a binding partner to acheive regulation. This is a common theme from our work.

We next explored the mechanism by which ubiquitinated proteins (cargoes) are sorted from protein lacking a ubiquitin tag on the endosome. This sorting process results in the inclusion of cargo into ILVs leading to their degradation. This sorting is achieved with the help of the ESCRT proteins. ESCRTs interact with protein through their UBDs and also interact with other ESCRTs. Recent experimental results led to our current hypothesis that ESCRTs and cargo form flexible networks prior to ILV formation. We first developed a cellular automata model which suggests that a flexible network of ESCRT and cargo is enough to achieve high efficiency sorting. In addition, this cellular automata model importantly led us to realize the need for a better understanding of how dimensionality changes protein interactions. All of the reported ESCRT dissociation constants are measured in solution as ESCRTs are soluble proteins. It was initially unclear how to transform a three-dimensional dissociation constant into a two-dimensional dissociation constant in order to model ESCRT-cargo networks on membranes.

To address the role of dimensionality in protein interactions, we took a stochastic approach to solving this problem with the simple assertions that in order to bind two proteins must find one another and then be oriented correctly to bind. Using exit time calculations in both discrete and continuous space, we derived reaction rates for protein interactions in one, two, and three dimensions. Our results suggest that dimensionality has a great affect on protein interactions. Surprisingly, we observe that dissociation constants in one and two dimension are dependent on the size of the domain on which they are measured. This differs from the traditional three-dimensional dissociation constants which does not depend on the size of the domain. Using our results, we are able to formulate a transformation equation to transform a dissociation constant from three dimensions to two dimensions. This transformation is very involved and requires much to be known about the structure of the proteins.

Using our general results on dimensional dependence of protein interactions, we again turned to ESCRT sorting. We transformed a three-dimensional dissociation constant to a two-dimensional dissociation constant and our results illustrated that being membrane bound is enough to enhance the affinity of the ESCRT UBDs. Therefore, ESCRT cargo interactions are likely more stable than previously thought. With this result and using our derived two-dimensional reaction rates, we explored the time it takes to sort cargo into

networks of a critical size prior to ILV formation. Our results suggest that sorting can be achieved on the order of seconds. As more is understood about how ESCRT-III interacts with the early ESCRTs to form vesicles, it will be interesting to include this in our model to assess the time it takes to form MVBs. As two-dimensional reaction rates depend on the size of the domain, we predict that sorting will speed up as vesicles form since the amount of endosomal membrane decreases. This work again shows that both functions of ubiquitin are important. Cargoes are ubiquitin tagged and cargoes also interact with ESCRT through ubiquitin interaction mediated by UBDs.

The final model in this dissertation aims to better understand ubiquitin-dependent regulation of Rad18 in the DNA damage tolerance pathway. In order to respond appropriately to DNA damage caused by MMS, Rad18 must interact with SHPRH. Under normal conditions, this interaction does not take place; rather, Rad18 forms complexes with HLTF. Upon MMS treatment, HLTF is degraded and an increase in deubiquitination is observed. Both HLTF degradation and deubiquitination appear to be necessary while neither is sufficient to induce Rad18/SHPRH complexes. As SHPRH, HLTF, and ubiquitin interact with the Znf domain of Rad18, this becomes a competitive binding problem. With differential equation models, we sought to address the importance of Rad18/Rad18u complexes and Rad18 autoubiquitination in this regulation. Unfortunately, we were not able to do so. However, our results suggest that a simple measurement of dissociation constants for Rad18 binding events will allow us to distinguish between these models and assess the importance of Rad18/Rad18u complexes and Rad18 autoubiquitination. While these results are somewhat preliminary without the necessary experimental measurements, we include them in this dissertation as an example of ubiquitin dependent regulation which does not rely on degradation. This is also an example of ubiquitin-dependent regulation of a strictly soluble protein. We again observe that ubiquitin acts in this system as both a post-translational modification and as a binding partner to achieve regulation.

As a whole, this dissertation gives three examples of ubiquitin-dependent regulatory systems and makes predictions about the mechanisms involved in each system. In all three examples, ubiquitin functions in a regulatory role through not only being added to proteins by enzymes but also binding to proteins through UBDs. Importantly, this dissertation suggests that dimensionality plays a significant role in protein regulation as protein interaction affinities, including UBD affinities, are affected by the dimension in which the interactions take place.

APPENDIX A

SUPPLEMENT TO CHAPTER 2

In order to compute the flip rate, r, from the reported V_{max} for uracil uptake in [45] we use the following simplified model for flux of uracil across the plasma membrane for a fixed amount of Fur4, P_{tot}:

$$\frac{dP^e}{dt} = -kS_eP^e + jPb^e - rP^e + rP^i, \tag{A.1}$$

$$\frac{dPb^e}{dt} = kS_eP^e - jPb^e - rPb^e + rPb^i, \tag{A.2}$$

$$\frac{dP^i}{dt} = -\frac{k}{w}SP^i + jPb^i - rP^i + rP^e, \tag{A.3}$$

$$\frac{dPb^i}{dt} = \frac{k}{w}SP^i - jPb^i - rPb^i + rPb^e. \tag{A.4}$$

$$\tag{A.5}$$

We calculate the inward flux across the plasma membrane at steady state as

$$F_{mem} = jPb^i - \frac{k}{w}SP^i. \tag{A.6}$$

$$\tag{A.7}$$

The maximal flux is given by

$$V_{max} = \lim_{S_e \to \infty} F_{mem}(S = 0) \tag{A.8}$$

$$= \frac{jrP_{tot}}{2(j+r)}. \tag{A.9}$$

We can then solve for r:

$$r = \frac{2V_{max}j}{-2V_{max} + jP_{tot}}. \tag{A.10}$$

The reported V_{max} is $\sim 8.8(\text{nmol} \cdot min^{-1}10^7\text{cells}^{-1})$. For our model, we change the units and use $V_{max} = 8.8 \cdot 10^3(\text{molecules/cell/ms})$. This V_{max} is reported with enriched plasma membrane Fur4 so we assume a high amount of Fur4, $P_{tot} = 1000$. We estimate j to be 10^3 in our model giving $r = 17.915$ for the flip rate.

APPENDIX B

SUPPLEMENTS TO CHAPTER 4

B.1 Discrete Space Exit Time

We now present the general exit time calculation for a discrete state space following the matrix differential equation,

$$
\begin{pmatrix} \frac{d\mathbf{q}}{dt} \\ \frac{d\mathbf{p}}{dt} \end{pmatrix} = \begin{pmatrix} Z_1 & B \\ Z_2 & W \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \tag{B.1}
$$

with $\mathbf{q}$ being the vector of probabilities of being in the absorbing states and $\mathbf{p}$ being the vector of probabilities of being in the nonabsorbing states. Z_1, Z_2, B, and W are matrices of the appropriate size depending on the sizes of $\mathbf{q}$ and $\mathbf{p}$. Note that since $\mathbf{q}$ represents the probabilities of being in absorbing states, Z_1 and Z_2 are matrices of all zeros. The above matrix differential equation can be split into two differential equations as follows:

$$
\frac{d\mathbf{q}}{dt} = B\mathbf{p}, \tag{B.2}
$$

$$
\frac{d\mathbf{p}}{dt} = W\mathbf{p}. \tag{B.3}
$$

As $\mathbf{q}$ represents the probabilities of being in the absorbing states it follows that the CDF for exit time is $\mathbf{1}^T\mathbf{q}(t)$, the sum of the absorbing states at any given time. Therefore, the expected time of exit is given by

$$
T = \int_0^\infty t(\mathbf{1}^T \frac{d\mathbf{q}}{dt})dt. \tag{B.4}
$$

Using (B.2) and (B.3), we can compute this exit time as

$$
\begin{aligned}
T &= \int_0^\infty t(\mathbf{1}^T \frac{d\mathbf{q}}{dt})dt \tag{B.5} \\
&= \mathbf{1}^T B \int_0^\infty t\mathbf{p}\,dt \\
&= \mathbf{1}^T B W^{-1} \int_0^\infty t\frac{d\mathbf{p}}{dt}dt. \tag{B.6}
\end{aligned}
$$

Integrating by parts,

$$= \mathbf{1}^T B W^{-1} \left[\mathbf{p}(t)|_0^\infty - \int_0^\infty \mathbf{p} dt \right] \tag{B.7}$$

$$= -\mathbf{1}^T B W^{-1} \int_0^\infty \mathbf{p} dt \tag{B.8}$$

$$= -\mathbf{1}^T B W^{-2} \int_0^\infty \frac{d\mathbf{p}}{dt} dt \tag{B.9}$$

$$= -\mathbf{1}^T B W^{-2} \left[\mathbf{p}(\infty) - \mathbf{p}(0) \right] \tag{B.10}$$

$$= \mathbf{1}^T B W^{-2} \mathbf{p}(0). \tag{B.11}$$

In the specific case of interest in Chapter 4, this reduces to

$$T = -\mathbf{1}^T W^{-1} \mathbf{p}(0). \tag{B.12}$$

We see this from computing the splitting probability, $\mathbf{q}(\infty)$. Assuming that $\mathbf{q}(0) = 0$, i.e., nothing is initially in an absorbing state, we see that

$$\int_0^\infty \frac{d\mathbf{q}}{dt} dt = \mathbf{q}(\infty). \tag{B.13}$$

We can now compute the splitting probability to be

$$\int_0^\infty \frac{d\mathbf{q}}{dt} dt = B \int_0^\infty \mathbf{p} dt \tag{B.14}$$

$$= B W^{-1} \int_0^\infty \frac{d\mathbf{p}}{dt} dt \tag{B.15}$$

$$= B W^{-1} \left[\mathbf{p}(\infty) - \mathbf{p}(0) \right]. \tag{B.16}$$

$$\tag{B.17}$$

Since $\mathbf{p}(0) = 0$ the splitting probability is given by

$$\mathbf{q}(\infty) = -B W^{-1} \mathbf{p}(0). \tag{B.18}$$

In the case of only one absorbing state, $-B W^{-1} \mathbf{p}(0) = 1$. Assuming that nothing is initially in the absorbing state, we also know that $\mathbf{1}^T \mathbf{p}(0) = 1$. Therefore, in this case, $-B W^{-1} \mathbf{p}(0) = \mathbf{1}^T \mathbf{p}(0)$ and $-B W^{-1} = \mathbf{1}^T$. It follows that in the case of only one absorbing state, the exit time is given by

$$T = -\mathbf{1}^T W^{-1} \mathbf{p}(0). \tag{B.19}$$

The problem in Chapter 4 is a case in which there is only one absorbing state.

B.2 Proof of $W^{-1} = \frac{A}{\delta} - \frac{\Pi}{k}$

For a matrix W which can be written as $W = \delta L - kE$ where L is a matrix with zero column sum and

$$E = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{B.20}$$

we can prove that $W^{-1}\frac{A}{\delta} - \frac{\Pi}{k}$ where A and Π are independent of both δ and k, and $\Pi = \phi \mathbf{1}^T$ where ϕ is an element of the null space of L normalized such that $\phi(1) = 1$. The proof is as follows. Let $E = ee^T$ where

$$e = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{B.21}$$

It follows that

$$\frac{dW}{dk} = -E \tag{B.22}$$

$$\frac{dW}{d\delta} = L. \tag{B.23}$$

We observe

$$0 = \frac{d}{dk}(WW^{-1}) \tag{B.24}$$

$$= \frac{dW}{dk}W^{-1} + W\frac{dW^{-1}}{dk} \tag{B.25}$$

$$= -Ew^{-1} + W\frac{dW_{-1}}{dk}. \tag{B.26}$$

Therefore,

$$W\frac{dW^{-1}}{dk} = EW^{-1}. \tag{B.27}$$

Now consider the matrix $W(\frac{dW_{-1}}{dk} - \frac{\Pi}{k^2})W$.

$$W(\frac{dW^{-1}}{dk} - \frac{\Pi}{k^2})W = W\frac{dW^{-1}}{dk}W - \frac{W\Pi W}{k^2} \tag{B.28}$$

$$= EW^{-1}W - \frac{W\Pi W}{k^2} \tag{B.29}$$

$$= E - \frac{(\delta L - kE)\phi \mathbf{1}^T (\delta L - kE)}{k^2}. \tag{B.30}$$

Remember that $L\phi = 0$ and that $\mathbf{1}^T L = 0$. Therefore,

$$E - \frac{(\delta L - kE)\phi\mathbf{1}^T(\delta L - kE)}{k^2} \;=\; E - \frac{k^2 E\phi\mathbf{1}^T E}{k^2} \tag{B.31}$$

$$= \;E - ee^T\phi\mathbf{1}^T ee^T \tag{B.32}$$

$$= \;E - E \tag{B.33}$$

$$= \;0. \tag{B.34}$$

Therefore,

$$W\left(\frac{dW_{-1}}{dk} - \frac{\Pi}{k^2}\right)W = 0 \tag{B.35}$$

and since W is nonsingular

$$\frac{dW^{-1}}{dk} - \frac{\Pi}{k^2} = 0 \tag{B.36}$$

This can be rewritten as

$$\frac{d}{dk}\left(W^{-1} + \frac{\Pi}{k}\right) = 0. \tag{B.37}$$

which leads to the conclusion that $W^{-1} + \frac{\Pi}{k}$ is independent of k. Therefore,

$$W^{-1} + \frac{\Pi}{k} = f(\delta). \tag{B.38}$$

It now remains to prove that A is independent of δ. To do this, we first show that $EA = AE = 0$.

Assume that $W^{-1} = \frac{A}{\delta} - \frac{\Pi}{k}$. Since $WW^{-1} = I$, it follows that

$$I \;=\; (\delta L - kE)\left(\frac{A}{\delta} - \frac{\Pi}{k}\right) \tag{B.39}$$

$$= \;\frac{\delta LA}{\delta} - \frac{\delta L\Pi}{k} - \frac{kEA}{\delta} + \frac{kE\Pi}{k} \tag{B.40}$$

$$= \;\frac{WA}{\delta} + E\Pi. \tag{B.41}$$

Rearranging we get,

$$\frac{WA}{\delta} \;=\; I - E\Pi. \tag{B.42}$$

Notice that $I - E\Pi$ has the following structure

$$\begin{pmatrix} 0 & -1 & -1 & \cdots & -1 \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \tag{B.43}$$

We now observe that $(I - E\Pi)E$ is a matrix of all zeros.

$$
\begin{pmatrix}
0 & -1 & -1 & \cdots & -1 \\
& 1 & & & \\
& & 1 & & \\
& & & \ddots & \\
& & & & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}
=
\begin{pmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}
\tag{B.44}
$$

Therefore, $\frac{WAE}{\delta} = (I - E\Pi)E = 0$. Since W is nonsingular, $AE = 0$. We can similarly show that $EA = 0$. We now show that A is independent of δ.

$$
\begin{aligned}
0 &= \frac{d}{d\delta}(WW^{-1}) \tag{B.45} \\
&= \frac{dW}{d\delta}W^{-1} + w\frac{dW^{-1}}{d\delta} \tag{B.46} \\
&= L\left(\frac{A}{\delta} - \frac{\Pi}{k}\right) + (\delta L - kE)\left(\frac{dA}{d\delta}\frac{1}{\delta} - \frac{A}{\delta^2}\right) \tag{B.47} \\
&= \frac{LA}{\delta} - \frac{L\Pi}{k} + \frac{W}{\delta}\frac{dA}{d\delta} - \frac{\delta LA}{\delta^2} + \frac{kEA}{\delta^2} \tag{B.48} \\
&= \frac{W}{\delta}\frac{dA}{d\delta}. \tag{B.49}
\end{aligned}
$$

Therefore, $\frac{dA}{d\delta} = 0$ and A is indeed independent of δ which proves that $W^{-1} = \frac{A}{\delta} - \frac{\Pi}{k}$ with A independent of δ and k.

REFERENCES

[1] O. S. ANDERSEN AND R. E. KOEPPE, *Bilayer thickness and membrane protein function: an energetic perspective*, Annu. Rev. Biophys. Biomol. Struct., 36 (2007), pp. 107–130.

[2] M. BABST, D. J. KATZMANN, E. J. ESTEPA-SABAL, T. MEERLOO, AND S. D. EMR, *ESCRT-III: an endosome-associated heterooligomeric protein complex required forMVB sorting*, Developmental Cell, 3 (2002), pp. 271–282.

[3] G. I. BELL, M. DEMBO, AND P. BONGRAND, *Cell adhesion. competition between nonspecific repulsion and specific bonding.*, Biophysical Journal, 45 (1984), p. 1051.

[4] O. BERG, *Orientation constraints in diffusion-limited macromolecular association. the role of surface diffusion as a rate-enhancing mechanism.*, Biophysical Journal, 47 (1985), p. 1.

[5] O. G. BERG AND P. H. VON HIPPEL, *Diffusion-controlled macromolecular interactions*, Annual Review of Biophysics and Biophysical Chemistry, 14 (1985), pp. 131–158.

[6] M.-O. BLONDEL, J. MORVAN, S. DUPRÉ, D. URBAN-GRIMAL, R. HAGUENAUER-TSAPIS, AND C. VOLLAND, *Direct sorting of the yeast uracil permease to the endosomal system is controlled by uracil binding and Rsp5p-dependent ubiquitylation*, Molecular Biology of the Cell, 15 (2004), pp. 883–895.

[7] E. BOURA, B. RÓŻYCKI, H. S. CHUNG, D. Z. HERRICK, B. CANAGARAJAH, D. S. CAFISO, W. A. EATON, G. HUMMER, AND J. H. HURLEY, *Solution structure of the ESCRT-I and -II supercomplex: implications for membrane budding and scission*, Structure, 20 (2012), pp. 874–886.

[8] G. BULTYNCK, V. L. HEATH, A. P. MAJEED, J.-M. GALAN, R. HAGUENAUER-TSAPIS, AND M. S. CYERT, *Slm1 and slm2 are novel substrates of the calcineurin phosphatase required for heat stress-induced endocytosis of the yeast uracil permease*, Molecular and Cellular Biology, 26 (2006), pp. 4729–4745.

[9] M. L. DUSTIN, S. K. BROMLEY, M. M. DAVIS, AND C. ZHU, *Identification of self through two-dimensional chemistry and synapses*, Annual Review of Cell and Developmental Biology, 17 (2001), pp. 133–157.

[10] M. L. DUSTIN, L. M. FERGUSON, P.-Y. CHAN, T. A. SPRINGER, AND D. E. GOLAN, *Visualization of CD2 interaction with LFA-3 and determination of the two-dimensional dissociation constant for adhesion receptors in a contact area.*, The Journal of Cell Biology, 132 (1996), pp. 465–474.

[11] J. M. GALAN, V. MOREAU, B. ANDRE, C. VOLLAND, AND R. HAGUENAUER-TSAPIS, *Ubiquitination mediated by the npi1p/rsp5p ubiquitin-protein ligase is required for*

endocytosis of the yeast uracil permease, Journal of Biological Chemistry, 271 (1996), pp. 10946–10952.

[12] C. GARDINER, *Stochastic methods*, Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1985.

[13] J. GILES, *Chemistry nobel for trio who revealed molecular death-tag*, Nature, 431 (2004), pp. 729–729.

[14] D. J. GILL, H. TEO, J. SUN, O. PERISIC, D. B. VEPRINTSEV, S. D. EMR, AND R. L. WILLIAMS, *Structural insight into the ESCRT-I/-II link and its role in MVB trafficking*, The EMBO Journal, 26 (2007), pp. 600–612.

[15] G. GOLDSTEIN, M. SCHEID, U. HAMMERLING, D. SCHLESINGER, H. NIALL, AND E. BOYSE, *Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells*, Proceedings of the National Academy of Sciences, 72 (1975), pp. 11–15.

[16] G. GRUENERT, B. IBRAHIM, T. LENSER, M. LOHEL, T. HINZE, AND P. DITTRICH, *Rule-based spatial modeling with diffusing, geometrically constrained molecules*, BMC Bioinformatics, 11 (2010), p. 307.

[17] G. D. HOLMAN AND I. V. SANDOVAL, *Moving the insulin-regulated glucose transporter GLUT4 into and out of storage*, Trends in Cell Biology, 11 (2001), pp. 173–179.

[18] J. HURLEY, S. LEE, AND G. PRAG, *Ubiquitin-binding domains*, Biochem. J, 399 (2006), pp. 361–372.

[19] J. H. HURLEY, *The ESCRT complexes*, Critical Reviews in Biochemistry and Molecular Biology, 45 (2010), pp. 463–487.

[20] J. JANIN, *The kinetics of protein-protein recognition*, Proteins-Structure Function and Genetics, 28 (1997), pp. 153–161.

[21] Y. KEE, N. LYON, AND J. M. HUIBREGTSE, *The Rsp5 ubiquitin ligase is coupled to and antagonized by the Ubp2 deubiquitinating enzyme*, The EMBO Journal, 24 (2005), pp. 2414–2424.

[22] J. M. KEENER AND M. BABST, *Quality control and substrate-dependent downregulation of the nutrient transporter Fur4*, Traffic, 14 (2013), pp. 412–427.

[23] L. KERN, J. DE MONTIGNY, R. JUND, AND F. LACROUTE, *The FUR1 gene of saccharomyces cerevisiae: cloning, structure and expression of wild-type and mutant alleles*, Gene, 88 (1990), pp. 149–157.

[24] M. S. KOSTELANSKY, C. SCHLUTER, Y. Y. C. TAM, S. LEE, R. GHIRLANDO, B. BEACH, E. CONIBEAR, AND J. H. HURLEY, *Molecular architecture and functional model of the complete yeast ESCRT-I heterotetramer*, Cell, 129 (2007), pp. 485–498.

[25] M. H. LAM, D. URBAN-GRIMAL, A. BUGNICOURT, J. F. GREENBLATT, R. HAGUENAUER-TSAPIS, AND A. EMILI, *Interaction of the deubiquitinating enzyme Ubp2 and the E3 ligase Rsp5 is required for transporter/receptor sorting in the multivesicular body pathway*, PloS One, 4 (2009), p. e4259.

[26] J.-R. Lin, M. K. Zeman, J.-Y. Chen, M.-C. Yee, and K. A. Cimprich, *SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis*, Molecular Cell, 42 (2011), pp. 237–249.

[27] C. MacDonald, N. J. Buchkovich, D. K. Stringer, S. D. Emr, and R. C. Piper, *Cargo ubiquitination is essential for multivesicular body intralumenal vesicle formation*, EMBO Reports, 13 (2012), pp. 331–338.

[28] S. K. Mageswaran, M. G. Dixon, M. Curtiss, J. P. Keener, and M. Babst, *Binding to any ESCRT can mediate ubiquitin-independent cargo sorting*, Traffic, 15 (2014), pp. 212–229.

[29] M. Marsh, G. Griffiths, G. E. Dean, I. Mellman, and A. Helenius, *Three-dimensional structure of endosomes in BHK-21 cells*, Proceedings of the National Academy of Sciences, 83 (1986), pp. 2899–2903.

[30] J. R. Mayers, I. Fyfe, A. L. Schuh, E. R. Chapman, J. M. Edwardson, and A. Audhya, *ESCRT-0 assembles as a heterotetrameric complex on membranes and binds multiple ubiquitinylated cargoes simultaneously*, Journal of Biological Chemistry, 286 (2011), pp. 9636–9645.

[31] K. Mitra, I. Ubarretxena-Belandia, T. Taguchi, G. Warren, and D. M. Engelman, *Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol*, Proceedings of the National Academy of Sciences, 101 (2004), pp. 4083–4088.

[32] S. Miyase, S. Tateishi, K. Watanabe, K. Tomita, K. Suzuki, H. Inoue, and M. Yamaizumi, *Differential regulation of Rad18 through Rad6-dependent mono-and polyubiquitination*, Journal of Biological Chemistry, 280 (2005), pp. 515–524.

[33] G.-L. Moldovan and A. D. D'Andrea, *DNA damage discrimination at stalled replication forks by the Rad5 homologs HLTF and SHPRH*, Molecular Cell, 42 (2011), pp. 141–143.

[34] D. P. Nickerson, M. R. Russell, and G. Odorizzi, *A concentric circle model of multivesicular body cargo sorting*, EMBO Reports, 8 (2007), pp. 644–650.

[35] C. M. Pickart, *Mechanisms underlying ubiquitination*, Annual Review of Biochemistry, 70 (2001), pp. 503–533.

[36] J. Ren, Y. Kee, J. M. Huibregtse, and R. C. Piper, *Hse1, a component of the yeast Hrs-STAM ubiquitin-sorting complex, associates with ubiquitin peptidases and a ligase to control sorting efficiency into multivesicular bodies*, Molecular Biology of the Cell, 18 (2007), pp. 324–335.

[37] X. Ren, D. P. Kloer, Y. C. Kim, R. Ghirlando, L. F. Saidi, G. Hummer, and J. H. Hurley, *Hybrid structural model of the complete human ESCRT-0 complex*, Structure, 17 (2009), pp. 406–416.

[38] P. Saffman and M. Delbrück, *Brownian motion in biological membranes*, Proceedings of the National Academy of Sciences, 72 (1975), pp. 3111–3113.

[39] K. Séron, M.-O. Blondel, R. Haguenauer-Tsapis, and C. Volland, *Uracil-induced down-regulation of the yeast uracil permease*, Journal of Bacteriology, 181 (1999), pp. 1793–1800.

[40] S. B. Shields, A. J. Oestreich, S. Winistorfer, D. Nguyen, J. A. Payne, D. J. Katzmann, and R. Piper, *ESCRT ubiquitin-binding domains function cooperatively during MVB cargo sorting*, The Journal of Cell Biology, 185 (2009), pp. 213–224.

[41] D. Shoup, G. Lipari, and A. Szabo, *Diffusion-controlled bimolecular reaction rates. the effect of rotational diffusion and orientation constraints.*, Biophysical Journal, 36 (1981), p. 697.

[42] H. Takahashi, J. R. Mayers, L. Wang, J. M. Edwardson, and A. Audhya, *Hrs and STAM function synergistically to bind ubiquitin-modified cargoes in vitro*, Biophysical Journal, 108 (2015), pp. 76–84.

[43] H. Teo, D. J. Gill, J. Sun, O. Perisic, D. B. Veprintsev, Y. Vallis, S. D. Emr, and R. L. Williams, *ESCRT-I core and ESCRT-II GLUE domain structures reveal role for GLUE in linking to ESCRT-I and membranes*, Cell, 125 (2006), pp. 99–111.

[44] D. Torney and H. McConnell, *Diffusion-limited reaction rate theory for two-dimensional systems*, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 387 (1983), pp. 147–170.

[45] D. Urban-Grimal, B. Pinson, J. Chevallier, and R. Haguenauer-Tsapis, *Replacement of Lys by Glu in a transmembrane segment strongly impairs the function of the uracil permease from saccharomyces cerevisiae*, Biochem. J, 308 (1995), pp. 847–851.

[46] M. D. Vahey and D. A. Fletcher, *The biology of boundary conditions: cellular reconstitution in one, two, and three dimensions*, Current Opinion in Cell Biology, 26 (2014), pp. 60–68.

[47] S. Vijay-Kumar, C. E. Bugg, and W. J. Cook, *Structure of ubiquitin refined at 1.8 åresolution*, Journal of Molecular Biology, 194 (1987), pp. 531–544.

[48] R. L. Williams and S. Urbé, *The emerging shape of the ESCRT machinery*, Nature Reviews Molecular Cell Biology, 8 (2007), pp. 355–368.

[49] T. Wollert, D. Yang, X. Ren, H. H. Lee, Y. J. Im, and J. H. Hurley, *The ESCRT machinery at a glance*, Journal of Cell Science, 122 (2009), pp. 2163–2166.

[50] Y. Wu, J. Vendome, L. Shapiro, A. Ben-Shaul, and B. Honig, *Transforming binding affinities from three dimensions to two with application to cadherin clustering*, Nature, 475 (2011), pp. 510–513.

[51] M. K. Zeman, J.-R. Lin, R. Freire, and K. A. Cimprich, *DNA damage-specific deubiquitination regulates Rad18 functions to suppress mutagenesis*, The Journal of Cell Biology, 206 (2014), pp. 183–197.