Signaling Behaviors of Intrinsically Disordered Proteins

Lara Clemens

Make a title page and better title

Contents

1 Abstract

Disordered proteins are present in at least 40% of human proteins, including signaling molecules. What and how disordered proteins contribute to a signaling network is not well understood. I don't like this sentence Here we explore how disordered proteins impact singular and multi-site ligand binding through a variety of disorder-specific phenomenons. In addition to basic properties of disordered protein interactions, we investigate disorder-to-order transitions, electrostatic membrane association, simultaneous ligand binding and effects of surface presence. We find that disordered proteins may create positive or negative cooperativity and intrinsic sequential binding. These effects are influenced by the length of the disordered protein, size of the binding ligand, location of the binding sites along the polymer and presence or absence of a surface. Intrinsically disordered proteins themselves may therefore act as signaling modules that contribute complex signaling behavior to a network.

2 Introduction/Background/Motivation

Traditionally, studies of protein function have gone hand-in-hand with studies of protein structure. Proteins such as hemoglobin exhibit complicated behavior, such as cooperativity, by modifying their structure. The cooperative transition of hemoglobin from a closed to open state is well studied with the aid of crystallography and other structure elucidation tools. CITE

Of more recent interest is the behavior of molecules that cannot be crystallized. These proteins lack a defined structure and instead are capable of assuming many different conformations. Although examples of intrinsically disordered proteins (IDPs) have been reported since the 1970s, it was only in the past two decades that they became a focus of major research. [4]

As our understanding of disordered proteins develops, so too will our understanding of a variety of cellular behaviors. These studies will elucidate aspects of signaling, cytoskeleton formation, clustered reactions, and really. I need a fourth function.. Investigations into how disordered proteins mediate each of these processes may lead to new drug targets or introduce new directions for synthetic biology.

2.1 Intrinsically Disordered Proteins

Studies have shown that disordered proteins or disordered domains are present in at least 40% of human proteins, including those involved with signal propagation.[10] General functions of IDPs include as tethers between two globular domains cite, receptor subunits in signaling pathways cite, tethers to the membrane cite, and facilitators to actin polymerization cite Bryant2017 or someone else older?. Given the ubiquitous nature of IDPs, many questions arise: How does their existence influence cellular functions such as biochemical reactions, signaling networks, or cytoskeletal structure? Are there any benefits to being disordered over structured? and Can IDPs exhibit the same complicated behavior as structured proteins?

Paragraph focusing on disordered proteins in signaling pathways. E.g. CD28, CD3zeta, etc

Examples: B Cells, NK cells, Fc, etc cite - Xu et al 2008, who cites Reth1989, Weiss and Littman 1994

2.2 T Cell Receptor Zeta Chain

One example of an intrinsically disordered protein is the CD3 ζ chain, one of six disordered chains composing the T Cell Receptor (TCR) intracellular region. This molecule facilitates signal propagation in the TCR network in the immune system. An antigen binding to the extracellular regions of the TCR creates a signal

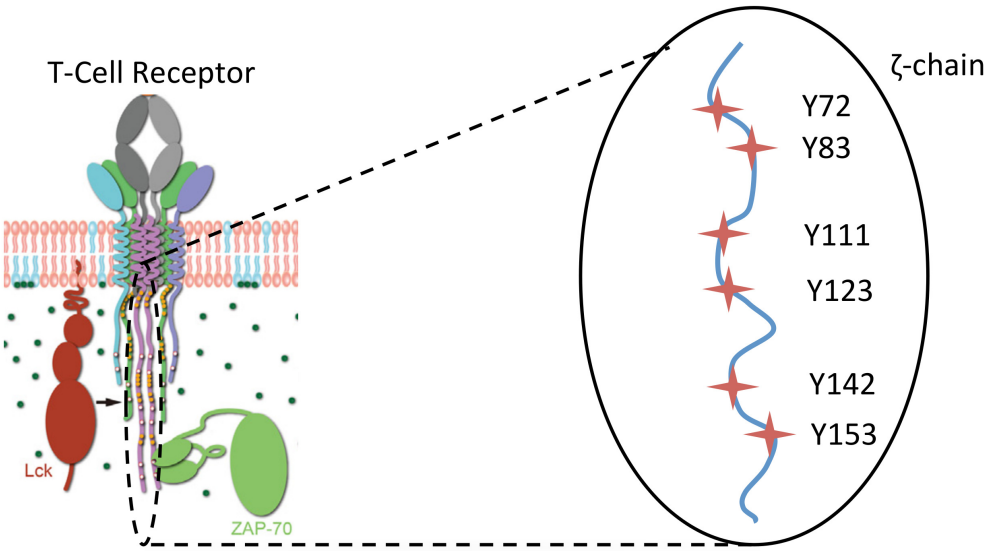


Figure 1: Cartoon of T Cell Receptor network (left) from [12]. Location of tyrosines in single ζ chain (right).

transmitted via a chain of events into the cell to the intracellular components of the TCR including the CD3ζ chain. CD3ζ undergoes multiple phosphorylation by kinase Lck before another molecule, ZAP-70 can attach and propagate the signal downstream. This pathway ultimately regulates T-cell cell fate decisions through cytokines production, (i.e. interleukin-2). Cite - Cell Signal.com?? Presumably I need to cite all of this somewhere?

Experiments with only a reconstituted mouse CD3ζ dimer, simplified extracellular receptor, kinase, and phosphatase have shown that the number of tyrosines impacts the potency and maximum phosphorylation but not the switch-like response reword [8]. Initial differential equation models indicate to achieve these characteristics, there would need to be a phosphorylation-dependent enhancement of more than 100-fold. That is, the sixth phosphorylation event would be at least 100-fold faster than the first phosphorylation event. [8]

Questions:

Their ODE model is single chain, right?

Is a single zeta chain sufficient for signalling?

Mukhop. et al says a single ITAM phos is technically sufficient, but has it been shown that it will?

2.3 Multisite Phosphorylation

Stuff about switch-like responses

Multisite phosphorylation specifically is a well-studied post-translational modification. This phenomenon occurs on both structured and unstructured proteins in many cell systems. CITE CITE CITE In signaling pathways, multisite phosphorylation can create ultrasensitivity. cite, anything else? Ultrasensitivity creates a strong response from intermediate signals and minimal none? response from low signals which reduces the impact of noise. CD3ζ - does give ultrasensitivity and we know it or it doesn't and we know it or we

don't know? We want to explore if and how multisite phosphorylation of disordered proteins conveys similar signaling functions.

2.4 Previous Modeling Attempts - FJC as Model for IDPs

Disordered proteins are commonly represented with models from polymer physics.[11],[9] The distribution of end-to-end distances for disordered domains matches a worm-like chain (WLC) model with persistence length 3.04Å. This suggests $l_k = 0.6$nm....which is not what we use...not sure how to explain that one.Seriously - even in Mukhopadhyay2016 cite{Zhou2001}Models of disordered proteins freely-jointed chain (FJC) and WLC converge in the thermodynamic limit, with the persistence length for the WLC as half the Kuhn length used for the FJC. cite Something

Could explain why those models were developed - what are they good at describing

Alternative models for multisite phosphorylation of IDPs include molecular dynamic, ordinary differential equations, or particle based models. Cite times when each of these has been used However, FJC or other 'mesoscale' approaches reach timescales on the order of microseconds to seconds, which are computationally out of reach for traditional atomic scale MD. Cite time scales? This approach also allows us to capture the steric effects of a disordered chain, which are missed by coarser models. Do I need to cite stuff here?

Representations of disordered proteins as freely-jointed chains have already been used to elucidate properties of IDPs. The disordered molecule, formin, captures profilin-actin monomers and delivers them to the growing end of actin, increasing the actin polymerization rate. In experimental studies of formin, exerting force on the distal end of formin counterintuitively enhances the polymerization rate two-fold over relaxed formins. [**Jegou2013**] Models of formin as a freely-jointed chain offer an explanation for this phenomenon. A force exerted on a freely-jointed chain extends the polymer, increasing capture of profilin-actin by increasing the availabiltiy of binding sites. This increase of capture rate balances the reduction in delivery rate to have a net positive impact on the actin polymerization rate. [3]

2.5 The Big Question

Since IDPs lack a consistent, rigid structure, they must fluctuate between multiple conformations. If we assume IDPs have no structure at all, then the protein may be in any conformation at any time and in fact would sample all of these conformations. This high degree of structure variability makes ligand binding to a specific binding site more challenging since it is now possible for the protein to transiently block its own binding site. That is, there will be conformations where the region around the binding site is occupied by other segments of the protein, preventing a ligand from occupying that space. In order to explore if and how disordered proteins can provide complex signaling behavior to a network, we create a model of a simple disordered protein using principles from polymer physics.

3 Model Development

3.1 Freely Jointed Chain

We create a generic model of a disordered protein using a simplified θ-solvent freely-jointed chain (FJC) from polymer physics. This model requires only specifying a number of rods (N) and a length per rod (Kuhn length, δ). The FJC consists of N rigid rods of length δ which are allowed to perform a random walk where the only constraints are the chain length and connections. In our simulation, the FJC is allowed to explore its configuration space through randomized movements. The FJC model tracks where each joint

is located, making steric interactions with the ligand easy to simulate compared to using the continuous WLC model. The ligand is simulated as an idealized sphere which may interact with the FJC. We compute quasi-equilibrium statistics of the chain and its bound or unbound ligands using a Monte Carlo (Metropolis) Algorithm.

We calculate how often a ligand is able to bind to an oriented sphere tangentially attached to the polymer, where 'able to bind' refers to the specified sphere being empty of both other polymer segments and half-space barriers weird wording?. To determine if a site is occluded in a given conformation, we check if any of the segment end points are located within the sphere of interest. We also check if the sphere crosses below the half-space surface designated at $z = 0$. Given that the binding sphere is large compared to the Kuhn length, we assume the probability of tangential occlusion (where a segment has end points outside the sphere but part of the segment lies within) is negligible compared to end point occlusion. Should we produce a plot for this someday? Or create a diagram to make it more obvious what all the types of occlusion are?

3.2 Theoretical results

From detailed balance, which is what and we can use it why? the dissociation constants $K_D \equiv k_{\text{off}}/k_{\text{on}}$ of the floppy (native) state and the rigid (partially phosphorylated) state are

$$\frac{k_{on}}{k_{off}} = \exp\left(\frac{-\Delta G}{k_B T}\right) \tag{1}$$

$$\frac{K_{D_F}}{K_{D_R}} = \frac{K_F}{K_R} = \frac{\frac{1}{\exp\left(\frac{-\Delta G_F}{k_B T}\right)}}{\frac{1}{\exp\left(\frac{-\Delta G_R}{k_B T}\right)}} \tag{2}$$

$$\frac{K_F}{K_R} = \frac{\exp\left(\frac{-\Delta G_R}{k_B T}\right)}{\exp\left(\frac{-\Delta G_F}{k_B T}\right)} \tag{3}$$

$$= \exp\left(\frac{\Delta G_F - \Delta G_R}{k_B T}\right) \tag{4}$$

$$= \exp\left(\frac{(E_F - TS_F) - (E_R - TS_R)}{k_B T}\right) \tag{5}$$

$$= \exp\left(\frac{S_R - S_F}{k_B}\right) \tag{6}$$

$$= \text{ex}\left(\frac{(S^R_{on} - S^R_{off}) - (S^F_{on} - S^F_{off})}{k_B}\right) \tag{7}$$

$$= \exp\left(\frac{(k_B \ln(\frac{\Omega_R}{\Omega_R})) - (k_B \ln(\frac{\Omega_F P_F}{\Omega_F})}{k_B}\right) \tag{8}$$

$$= \frac{1}{P_F} \tag{9}$$

where $G_j = E_j - TS_j$ is the free energy of binding in the free or rigid state, $S_j = k_B \ln W$ is the entropy of binding, where W is the number of microstates and P_j is the probability in the canonical ensemble that the configuration allows for binding. We let Ω be the total number of microstates, with $\Omega_F * P_F$ being the microstates available when the ligand is bound and Ω_R is the number of microstates available when the polymer is rigid. We also note $E_R = E_F$ since we are working at equilibrium. Since the perfectly rigid state always allows binding, $P_R = 1$, it is sufficient to compute P_F. We define P_{occ} as the probability that the region of space needed by the kinase domain is occupied by some of the polymer. Thus, $P_{occ} = 1 - P_F$. Then

$$\frac{K_F}{K_R} = \frac{1}{(1 - P_{occ})}. \tag{10}$$

Our problem of interest is now reduced to computing the occlusion probability.

3.3 Code Validation

There are theoretical solutions for many aspects of the freely-jointed chain. This provides a basis with which to verify our code.

3.3.1 Root-mean-square end-to-end distance

First, we look at the average end-to-end distance of the polymer (RMS). In our simulations, we normalize by the Kuhn length, so all simulations assume $\delta = 1$ and record all other parameters in units of Kuhn lengths. We know from polymer physics that the RMS should increase as $\delta\sqrt{N}$, which given $\delta = 1$, is just $\sqrt{N}$.

Average end-to-end distance (Root-mean-square end-to-end distance): [9]

$$\sqrt{\langle r_{ee}^2 \rangle} = \sqrt{N\delta^2} = \delta\sqrt{N}$$

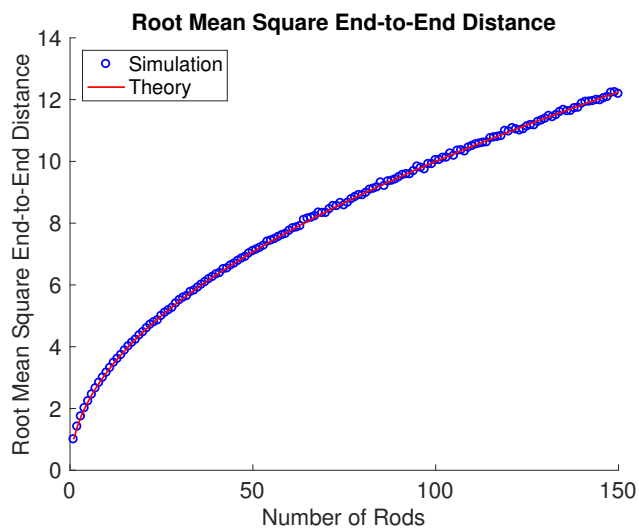


Figure 2: Theoretical root mean square end-to-end distance (red line) against simulated values (blue circles).

3.3.2 R_{ee} Distribution

End-to-end distribution: [11], [9]

$$P(r_{ee}) = 4\pi r^2 \left(\frac{3}{2\pi N\delta^2}\right)^{\frac{3}{2}} exp\left(\frac{-3r^2}{2N\delta^2}\right)$$

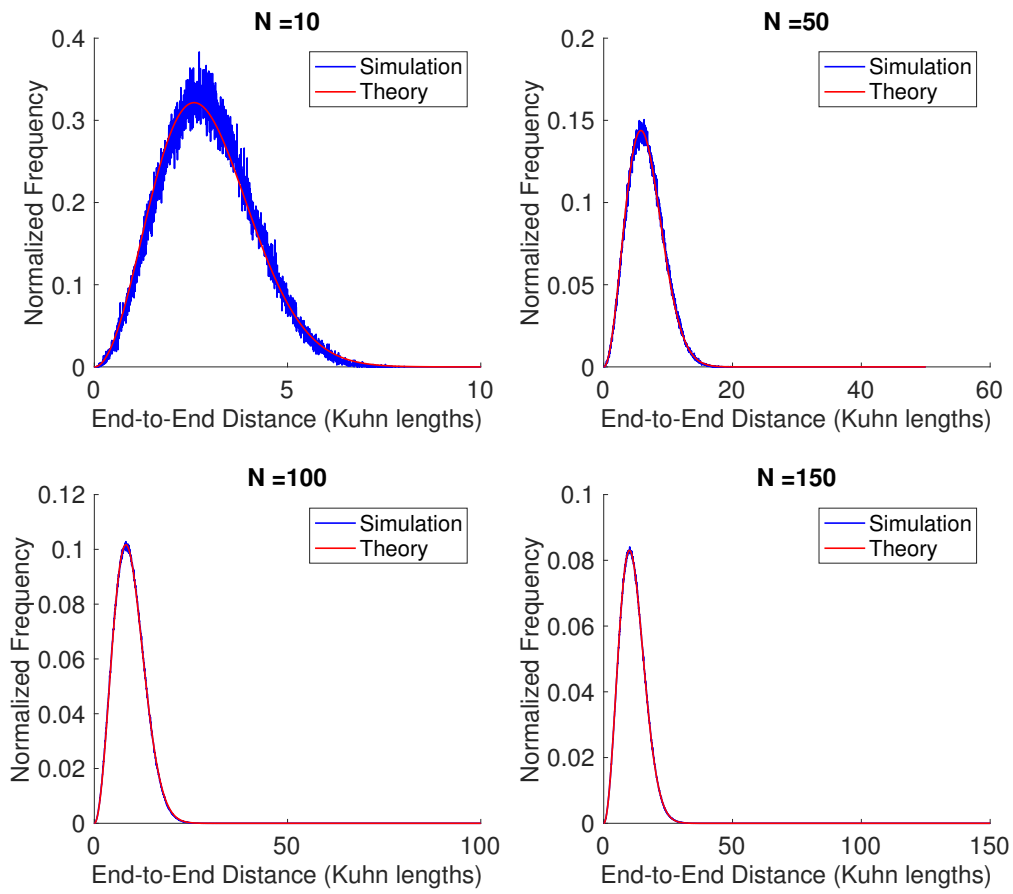


Figure 3: Simulated end-to-end distance distribution (blue) against theoretical distribution (red) for multiple polymer lengths (N).

and show derivation (?) of distribution formula?

3.3.3 Occlusion Probability of End Site VS Analytical Result

We consider analytical solutions to the simplest case, where the ligand is attempting to bind at the end of the freely-jointed chain in free-space. The probability that a random walk over time 0 to $\delta * N$, beginning at the edge of a sphere of radius, R, will cross into the sphere at any time is analagous to the probability the ligand is occluded by the freely jointed chain. This is a confusing description In order to solve this analytically, we must assume the random walk begins ϵ away from the sphere or it will always count as a failure. Then we may formulate as follows for the probability of survival, $p(\vec{x}, t)$ when the random walk begins at a starting point, $r = R + \epsilon$:

I got confused.

$$\begin{cases} \frac{\partial p}{\partial t} = D\nabla^2 p, & p = 0 \quad \text{at} \quad r = R \\ p(\vec{x}, 0) = \delta(\vec{x} - (R + \epsilon)) \end{cases} \tag{11}$$

In order to get a finite estimate to compare, we consider the solution when $N \to \infty$.

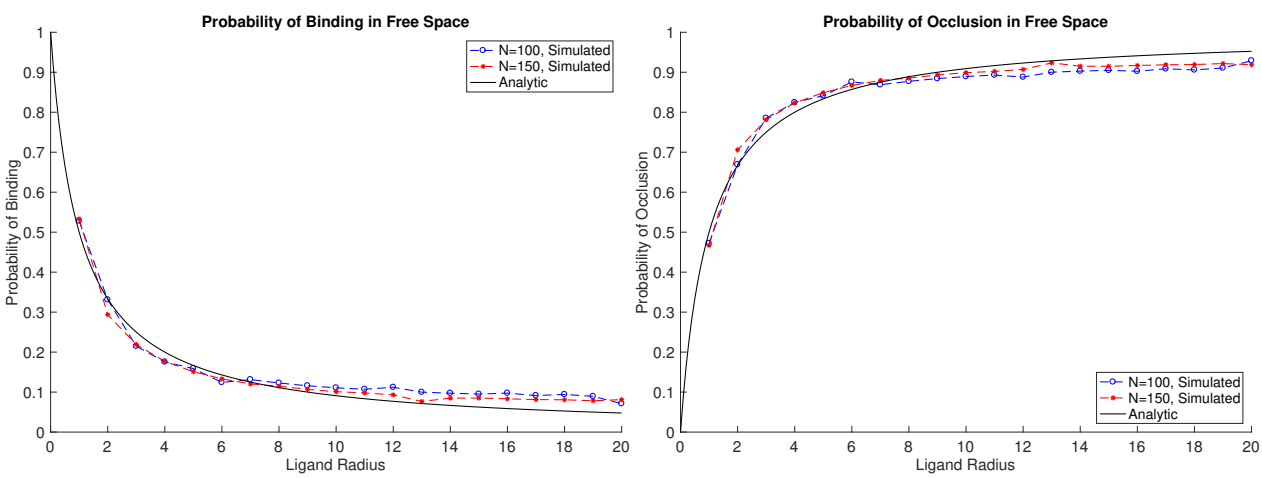
$$\begin{cases} q(r) = \mathbb{P}(\text{not hit sphere}|x(0) = r) \\ 0 = \frac{2}{r}q'(r) + q''(r) \\ q(R) = 0 \\ q(r) \to 1 \quad \text{as} \quad r \to \infty \end{cases} \tag{12}$$

We assume $\epsilon = 1$ Kuhn length and R measured in Kuhn lengths:

$$q(r) = 1 - \frac{R}{r}q(r) = 1 - \frac{R}{R+1} \tag{13}$$

We see that when we compare our simulated binding probability to the analytic solution, we get good agreement. This tells us both our code is working as desired and that N=100 is approximately equivalent to $N \to \infty$.

pick one of these - don't need both



(a) (b)

Figure 4: Analytic Solution

3.4 Case Study: T Cell Receptor Zeta Chain

For the following study, we will focus on the mouse TCR CD3 ζ chain. The TCR CD3 ζ chain is a subunit of CD3 consisting of 164 amino acids. Of these, 21 are included in the signal peptide region of the protein. The remaining 143 amino acids make up the extracellular, transmembrane and cytoplasmic regions of the CD3 ζ chain. The cytoplasmic tail is an intrinsically disordered chain of 113 amino acids containing multiple phosphorylation sites, called ITAMs (immunoreceptor tyrosine-based activation motif). There are three ITAMs on the ζ chain, each containing two tyrosines. The tyrosine kinase Lck phosphorylates each tyrosine in the ζ chain and it is thought that phosphorylation of one site influences the binding of the kinase to other sites.

In mouse CD3 ζ, the cytoplasmic tail spans residues 52-164 and the tyrosines are located at residues 72, 83, 111, 123, 142, 153. Therefore, if we were to renumber to begin at the beginning of the cytoplasmic tail, the region would be $N = 113$ amino acids long, with tyrosines located at $i = 21, 32, 60, 72, 91, 102$. (UniProt, entry P24161) Given an assumption of 0.3 nmper Kuhn length (i.e. one Kuhn length is equivalent to one

Cite UniProt? Or find paper?

amino acid), then the tyrosines are similarly located along the 113 segments of the FJC.

Mouse Lck is composed of an SH3, SH2 and protein kinase domain connected by small loops. The domains are 61, 98, and 254 amino acids respectively. (UniProt entry P06240) Using a protein molecular mass calculator (http://www.bioinformatics.org/sms/prot_mw.html) , we calculate that the kinase domain is 29.08 kDa. If we assume a protein density of 1.41 g / cm^3 then we can estimate the volume of the kinase domain [5]:

Cite UniProt? Or find paper?

Cite molecular mass calculator?

$$(29 \times 1000\text{Da}) * (1.66 \times 10^{-27}\text{kg / Da}) * (1000\text{g / kg})/(1.41\text{g / cm}^3) = 34\,\text{nm}^3.$$

If we then approximate the kinase domain as a sphere, then we can estimate a radius:

$$\begin{aligned} V &= \frac{4}{3}\pi r^3 \\ 34\,\text{nm}^3 &= \frac{4}{3}\pi r^3 \\ r &\approx 2\,\text{nm} \end{aligned}$$

Measurements from the crystal structure of Lck suggest a volume of 45 nm^3 for the kinase domain. This estimate is on the same order of magnitude as our previous estimate. Note also, if we recalculate the kinase radius using a volume of 45 nm^3, we estimate a radius of 2.2 nm. Given an assumption of 0.3nm per Kuhn length, we represent Lck with a radius of seven Kuhn lengths.

Cite or show PyMol for Lck

3.5 Possibly unnecessary stuff

where does this stuff go?

In this model, the FJC is allowed to pass through itself. However, this turns out to be unimportant. Although the disordered protein would have a nonzero bond width, this size would be small in comparison to the kinase volume and volume of exploration space. If you consider two thin strings in 3-dimensional space, they will almost never intersect. Therefore, we model our polymer with infinitely thin width. Since it is infinitely thin, the time when it would take on a conformation where it overlaps itself is negligible compared to the number of legal conformations it assumes.

Additionally, we are only interested in the ensemble of configurations, not the dynamics. Therefore, the way the polymer achieves its conformations is unimportant to the results and any instances of self-intersection are negligible compared to the larger ensemble. Allowing the FJC to pass through itself, while unphysical, does not change the validity of our simulation.

4 Local structuring of disordered signaling proteins gives rise to cooperativity and sequential binding

4.1 Introduction

Intrinsically disordered proteins sterically occlude ligands by transiently sequestering the binding site.

There is evidence that indicates some IDPs undergo a disordered-to-ordered transition upon post-translational modification, i.e. phosphorylation. For example, phosphorylation of PAGE4 restricts the number of conformations it may explore, increasing rigidity of the backbone.[6] Post-translational modifications may also cause a conformational change, altering the availability of the binding site. The neural protein 4E-BP2 undergoes multisite phosphorylation which leads to a disorder-to-order transition, forming β sheets which hide the eIF4E binding site. [1] Any other examples to include? tau protein! Go read that paper.

talk about initial models of TCR - aka your part in Mukhopadhyay 2016?

From our simple model of a disordered protein, we may answer some basic questions. What is the probability that a ligand will be able to bind a specific site on a cytosolic disordered protein? How does this probability change when the protein is anchored to the membrane? How do these probabilities change with the contour length of the protein, size of the ligand, and where the binding site is located along the length of the protein?

We are interested in how cooperative effects may arise in reactions with a disordered protein. One model we explore is disorder-to-order transitions as a result of post-translational modifications. With regards to our case study, CD3ζ, we assume that each phosphorylation event causes a fraction of the chain to undergo a disorder-to-order transition. Partial structuring will reduce the number of conformations available to the protein. We predict this will increase the number of conformations where a ligand may access its binding site due to a forced linearization of part of the chain. We investigate how this local structuring phenomenon would impact subsequent phosphorylations by Lck along TCR CD3ζ. Specifically, we want to know if local structuring of a disordered protein upon multi-site phosphorylation can induce enhancement of subsequent binding events. This is repetitive.

Include something about dephosphorylation? order-to-disorder transitions? Reversible phosphorylation?

4.2 Model and Methods

We model the TCR CD3ζ chain as a freely jointed chain, where each amino acid is represented as a segment as described in Sec. **??**. The unbound ligand is simulated as an idealized ghost sphere tangentially located at a single segment. Is it clear what this means? We calculate quasi-equilibrium statistics for the FJC and unbound ligand in both free-space and half-space. For the CD3ζ chain, we simulate only the cytoplasmic region by an FJC with 113 segments. In our simulation, all lengths are normalized by the Kuhn length, so each rod has length 1 and the radius of the ligand is also measured in Kuhn lengths. We consider a spherical estimate for the kinase Lck, with a radius of 2.1 nm(7 Kuhn lengths) as calculated in **??**.

We calculate how often a ligand is able to bind to an oriented sphere tangentially attached to the polymer, where 'able to bind' refers to the specified sphere being empty of both other polymer segments and half-space barriers weird wording?.

We include local structuring as a limitation on which joints of the freely jointed chain are allowed to rotate. For each modified residue, some number of joints are 'frozen' in a straight conformation. (Fig. **??**) The section of frozen segments acts as a single long segment, still rotating when other neighboring segments rotate. This method of local structuring leaves the original binding site in a primarily available configuration, however, we are interested in how it impacts the neighboring modification targets (i.e. tyrosines to be phosphorylated). We explore how the total number of structured residues per modification impacts the result.

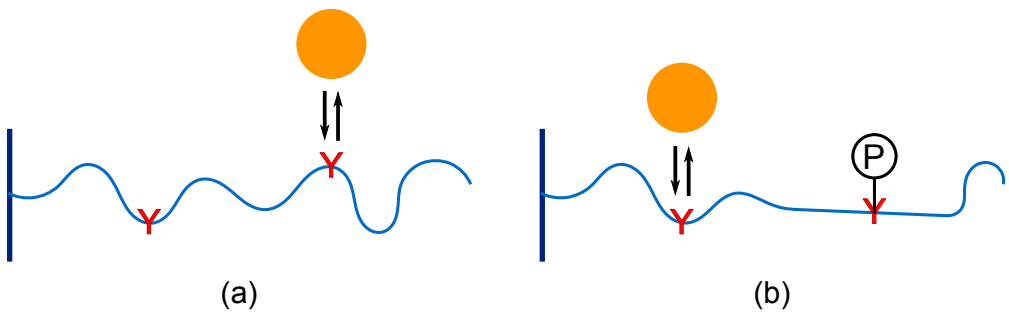


Figure 5: Cartoon of partial stiffening model. Entirely floppy FJC prior to modification (a) transitions to partially rigid chain after modification (b).

With respect to TCR CD3ζ, phosphorylation events are modeled as creating local structure as described above. We model dephosphorylation of CD3ζ, using the same parameters for the FJC. We also assume the phosphatase is the same size as the ligand Lck. However, dephosphorylation we model as 'unfreezing' nearby segments. Under the premise of eventually modeling reversible phosphorylation, we consider the case where the initial polymer state is only as stiff as it would be from the local stiffening by full phosphorylation. That is, if phosphorylation would locally structure the nearest eleven segments, then the fully phosphorylated form of CD3ζ would have 66 residues frozen. This is, the state the chain would then begin dephosphorylating from, where each dephosphorylation would good word for this? unstructure 11 of the initial 66 frozen residues.

4.3 Results

4.3.1 Surface presence influences polymer configuration

We first explore how a freely-jointed chain is affected by the presence of a surface or membrane. Generally, FJCs are most likely to be in a 'hairball' conformation due to entropic forcesinclude image of simulation in hairball?. When we compare the end-to-end distances of a FJC in free space to a FJC in half space, we find a shift in the distribution. The half-space polymer tends to have a larger end-to-end distance than the free-space polymer. (Fig. **??**) In half space, the polymer has fewer conformations it can take on since it can only occupy half of the region. This prevents many more 'hairball' conformations than it prevents straightened conformations. Overall, the presence of a membrane straightens the polymer. I don't really like these last couple sentences

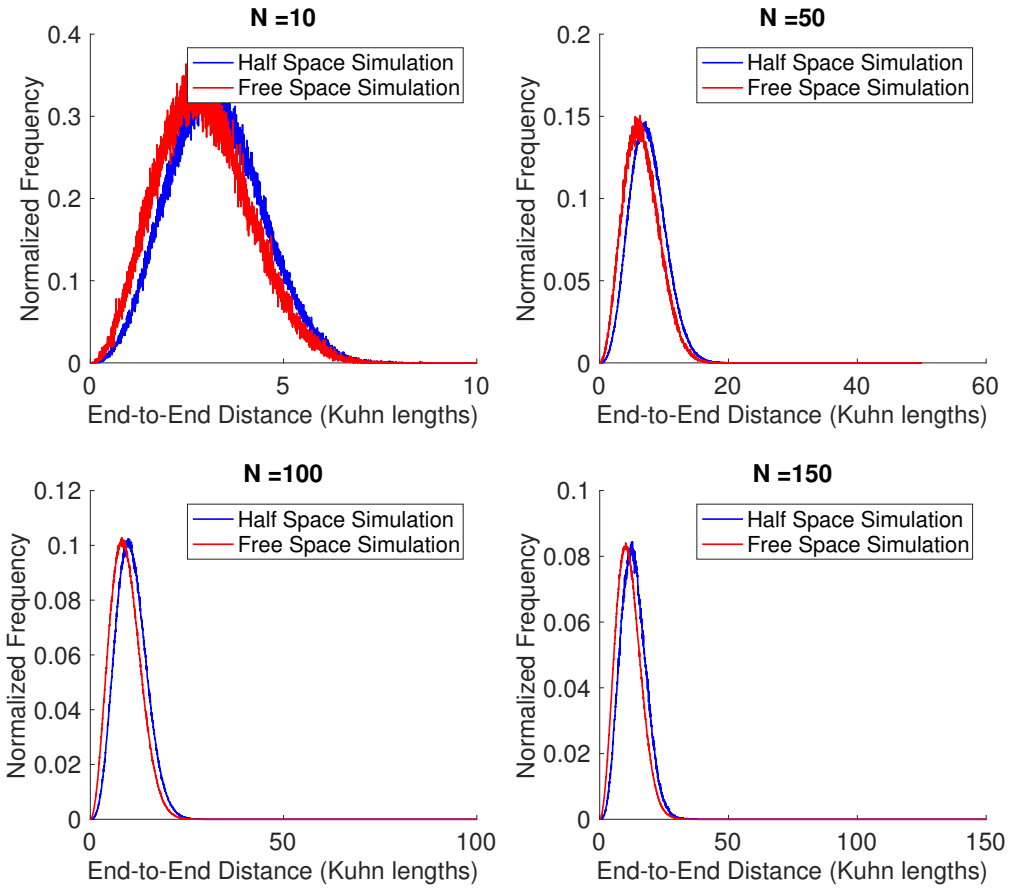


Figure 6: Distributions of polymer end-to-end distance in free space and half space for various polymer lengths (N). For each length, the end-to-end distance of the polymer in half-space is overall straighter than in free space.

4.3.2 Disorder inhibits ligand binding

We calculate how often a ligand may bind to a FJC, under varying characteristics for the FJC. We can see that in free space, the ligand is more occluded when there are more rods and/or the ligand is larger, for a specific binding site. (Fig. **??**) These are both intuitive results. A longer chain creates more conformations where any location on the chain will be occluded by other segments. A shorter or less flexible chain will not be able to bend enough to bury a binding site. A larger ligand will require more space near the binding site to be open, which is naturally less likely. Explain better?

A binding site in the middle of the polymer will also be less available to ligand binding since it is more likely sheltered by the rest of the 'hairball' polymer.

We compare how FJC parameters influence binding in free-space to binding in half-space. The membrane increases the likelihood of straighter conformations, which would be likely to increase the binding probability. However, the membrane itself may also occlude the ligand since we include an orientation of the binding site in our simulation. Therefore if the binding site is oriented down towards the membrane, then even if the polymer is not occluding the ligand, the membrane still might. When we include a hard plane membrane in our simulation, we see that the presence of a membrane generally decreases the probability of ligand binding. When the ligand is small enough, the difference in binding probabilities between free space and half space binding is minimal. (Fig. **??**)

Re-bin histograms Include image of simulation? Usually look bad....but might be a worthwhile visual. This is really ugly. Accidentally omitted last column? Color bars need to be fixed to be consistent somehow Color scales the same?

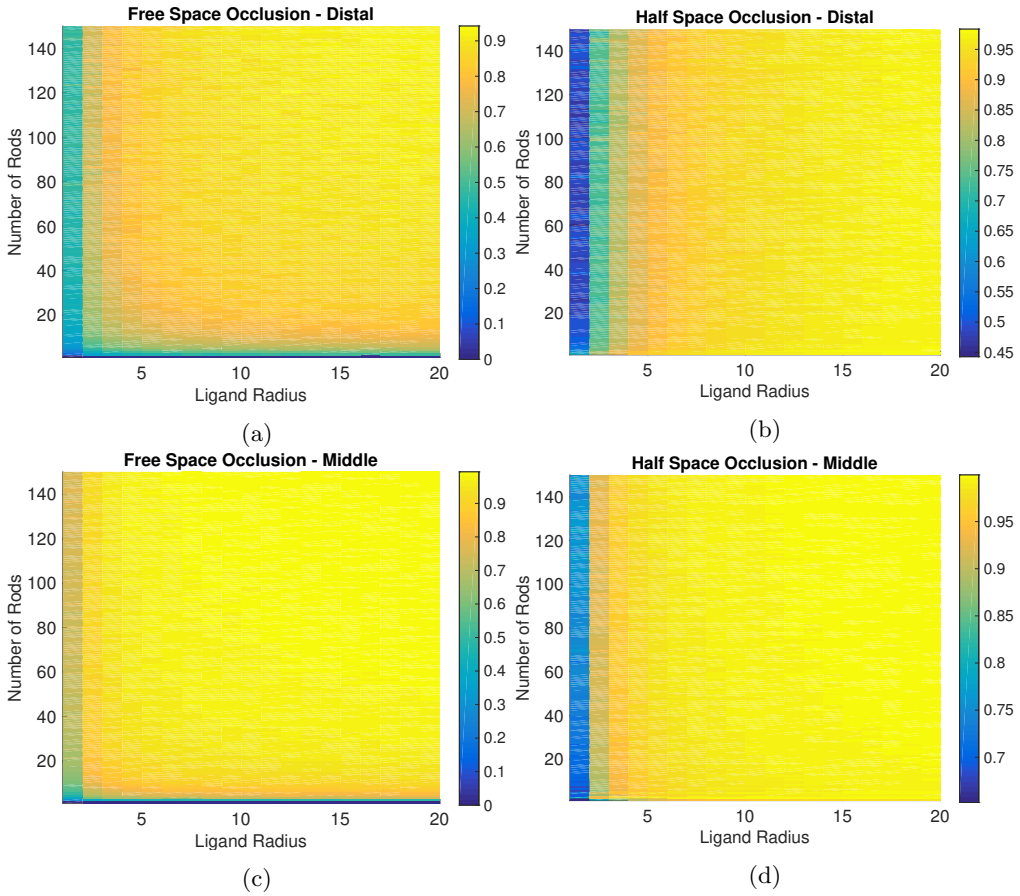


Figure 7: Probability of occluding a ligand of variable radius from a polymer of variable contour length when the binding site is at the distal end or middle of the chain in free space or half space. (a) Free space polymer with binding site at distal end. (b) Half space polymer with binding site at distal end. (c) Free space polymer with binding site in the middle. (d) Half space polymer with binding site in the middle.

4.3.3 Cooperativity may arise from local structuring in intrinsically disordered proteins

We now simulate the kinase Lck binding to the CD3ζ chain, assuming that each phosphorylation event will create a disorder-to-order transition. If we assume that phosphorylation of tyrosines on the CD3ζ chain induces local structuring, then neighboring tyrosines will feel a reduction of entropic occlusion. This will cause them to be more available to binding by a ligand. The magnitude of this effect will vary based on how much structuring occurs, e.g. if it affects the nearest two residues or the nearest ten residues.

We see from these simulations that the average binding rate of a ligand to the unphosphorylated tyrosines increases as the number of phosphorylated tyrosines (and therefore the number of structured residues) increases. From the first to the sixth phosphorylation event, there is a 3-10 fold increase in the average binding rate for moderate (1/12 of chain length) to severe (1/6 of chain length) local stiffening per phosphorylation. (Fig. **??**)

We have 10 fold change. Where does Omer's 100 fold change come in?

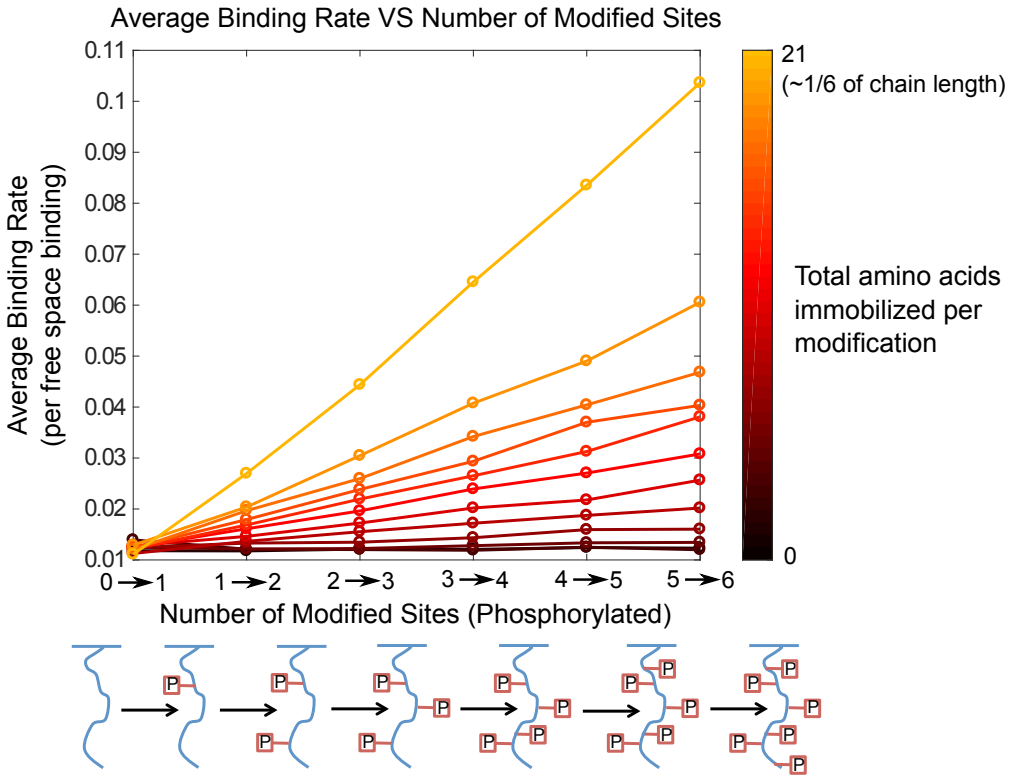


Figure 8: (Above) Simulated average binding rates of Lck to CD3ζ for different levels of phosphorylation given a specified degree of local structuring (none $\rightarrow$ black, low $\rightarrow$ red, high $\rightarrow$ orange). (Below) Cartoon representing a possible phosphorylation state series of CD3ζ.

4.3.4 Sequential binding naturally arises from polymer characteristics

Review this section

We hypothesized that without any modifications due to binding, then with only the presence of a membrane, we would see preferential binding in an outwards to inwards (6-5-4-3-2-1) manner. In regards to our model system, that is to say we expected the tyrosine furthest from the membrane to be phosphorylated first and the tyrosine closest to the membrane to be phosphorylated last. We surmised this from the idea that tyrosine's too close to the membrane would see extra kinase inhibition from membrane occlusion, whereas the tail of the disordered protein would not experience this same effect.

When considering the partial stiffening model, we predict that this effect would be enhanced. In an unphosphorylated state, if the tyrosine furthest from the membrane is most likely to be bound, then partial stiffening would give increased preference to the next furthest from the membrane (here called the fifth tyrosine) as the second to be phosphorylated. This comes from the reduction of total configurations available to the protein. Since the stiffening would occur around the sixth tyrosine, the nearest unphosphorylated tyrosine would be the most affected. The segment of polymer near the fifth tyrosine would have fewer available conformations and therefore proportionately fewer would be occluding the fifth tyrosine.

Explain the above better. Not totally convinced that's the best argument. What's the correct term for 'unoccluded'? Available? Open?

Do we introduce this as a purposeful topic or as a neat thing that falls out? Include Dushek paper as benefit or as intro?

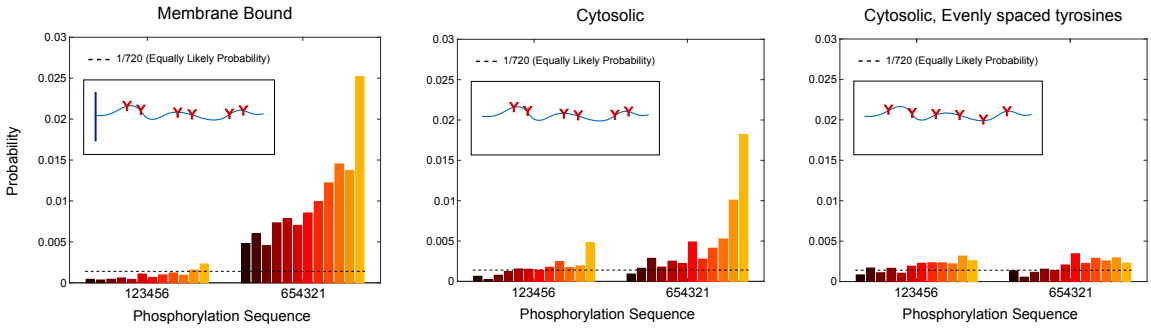
We obtained relative probabilities of binding for each tyrosine along the chain at each phosphorylation state. With this information, we are able to explore whether or not there is a dominant sequence of phosphorylation events. We ran a Gillespie algorithm with six events, one for each phosphorylation. From each run, we record which sequence or 'path' of phosphorylation. When we compile the path frequency data, it becomes clear that the first phosphorylation event is dominant. That is to say, any path beginning by phosphorylating the membrane distal tyrosine (6th tyrosine), will occur more often than any path beginning with the membrane proximal tyrosine.

For simplicity, we consider only the paths membrane proximal to membrane distal (123456) and distal to proximal (654321). We see that when a membrane is present, the probability of phosphorylating distal to proximal is much higher than that of phosphorylating proximal to distal. It is also more probable than if all paths were equally likely. This phenomenon is easily explained by the presence of a membrane. The membrane proximal tyrosine has a smaller range of space it may occupy than the distal tyrosine. Therefore, it is more likely to be close to the membrane in a configuration. Since the ligand cannot penetrate the membrane, tyrosines closer to the membrane have a higher probability of being effectively sheltered from ligand binding. This makes the distal to proximal path much more likely, regardless of local stiffening effects. This path preference is enhanced with local structuring since the distal tyrosine is more accessible and local structuring makes neighboring tyrosines even more available.

One would expect then, that a cytoplasmic disordered protein would not experience a path preference between N-terminal to C-terminal or vice versa. When we simulate CD3ζ without a membrane, there is still a marked preference for C-terminal to N-terminal phosphorylation. This effect arises from the spacing of the tyrosines along the polymer's length. As we noted above, the location of a tyrosine impacts how likely a ligand is to bind there. Since the tyrosines in CD3ζ are not equally spaced along the length of the protein, there is a bias for the tyrosine closest to one end, in this case the tyrosine closest to the C-terminal. When we simulate a polymer with the same length as CD3ζ but equally spaced tyrosines, we see that there is no longer a preference between the two paths.

We are soon going to introduce a ranking system for the paths in order to do a full analysis of all possible paths. This will help analyze the likelihood of any of the non-sequential paths.

Better to make one figure and label a,b,c in Inkscape?



(a) (b) (c)

Figure 9: Probability of phosphorylating membrane proximal to distal (123456) compared with distal to proximal (654321). Equally likely probability of all paths indicated with black, dotted line. (a) CD3ζ parameters, with membrane. (b) CD3ζ parameters, without membrane. (c) CD3ζ length, evenly spaced tyrosine locations, without membrane.

Presence of a membrane inherently leads to sequential binding, a feature that plays a key role in creating or enhancing ultrasensitivity.cite

4.3.5 Negative cooperativity may arise from local disordering of proteins

We investigate how dephosphorylation of a locally structured protein would impact future dephosphorylation events. Beginning with a model of CD3ζ where each tyrosine has been locally structured, we investigate how dephosphorylation and consequently, local unstructuring, influence the rate of subsequent dephosphorylations. We find that as the degree of local de/structuring increases, future dephosphorylations become less likely. For moderate de/structuring, this can create up to a 2-fold negative cooperative effect on the dephosphorylation rates from the first to the sixth event. (Fig. **??**)

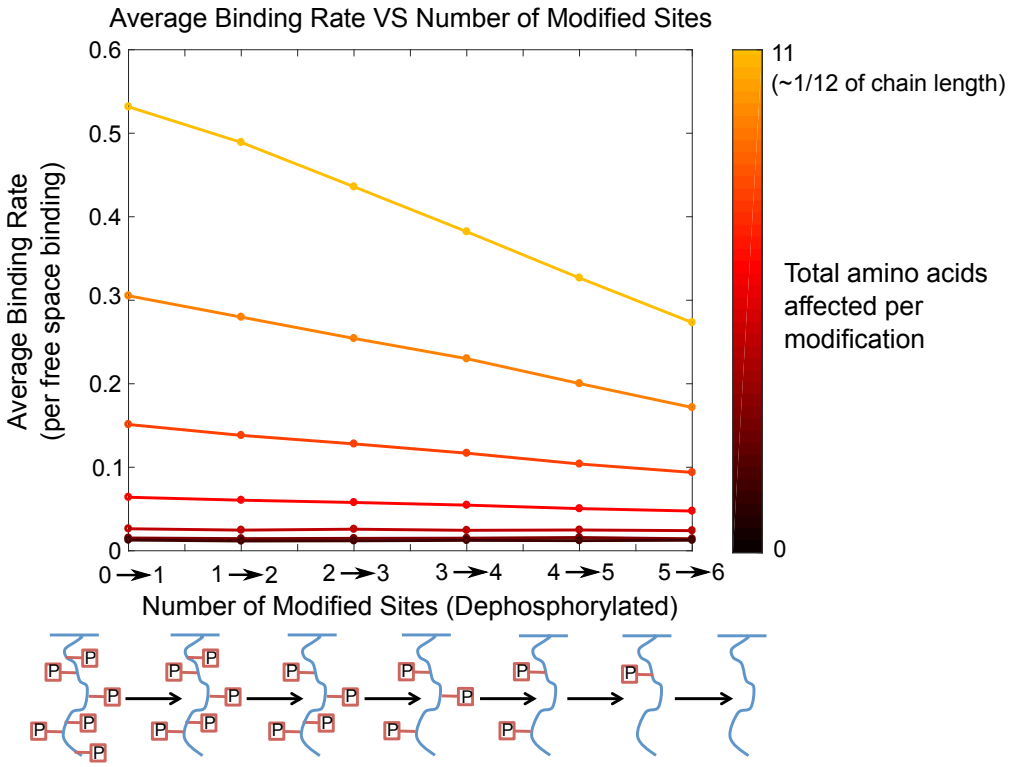


Figure 10: (Above) Simulated average binding rates of Lck to CD3ζ for different levels of dephosphorylation given a specified degree of local unstructuring (none $\rightarrow$ black, low $\rightarrow$ red, high $\rightarrow$ orange).

4.3.6 Future Work: Quantifying cooperativity from reversible phoshphorylation with Hill coefficients

We will create Hill curves of phosphorylation, dephosphorylation, and reversible phosphorylation. This will allow us to better understand and communicate our results with the rest of the biochemistry commmunity. In particular, the Hill curve of reversible phosphorylation will indicate if ultrasensitivity is achievable through this mechanism. true?

4.3.7 Future Work: Exploring sequence of binding from dephosphorylation and reversible phosphorylation

We will continue to investigate if phosphorylation, dephosphorylation or reversible phosphorylation naturally impose a sequence of binding. This will elucidate whether sequential binding occurs when both phosphorylation and dephosphorylation are impacting the polymer simultaneously.

Does 'why' we're doing this go here or in discussion?

4.4 Discussion

What is the definition of cooperativity? How does this compare to the 'not modifying the switch-like response? vs potency? vs ultrasensitivity?

We find that phosphorylation-induced local disordered-to-ordered transitions of a disordered protein with multiple phosphorylation sites will enhance the rate of future phosphorylation events. Local disordered-to-ordered transitions are sufficient to create positive cooperativity of PTMs. Question: is this 'cooperativity', how does that influence the network? Positive cooperativity can increase the sensitivity of a response to changes in ligand concentration and increase the potency of signaling responses. We find that for a simulated TCR CD3ζ chain, disordered-to-ordered transitions can give a 3 to 8 fold enhancement of phosphorylation rate from the first to the last event for certain ranges of local structuring. This provides one method of achieving nonlinear signaling behavior from a protein with no native structure, assuming high kinase to phosphatase concentration ratio.

Conversely, if we assume a low kinase to phosphatase ratio, we find that dephosphorylation-induced local unstructuring can create a mild reduction of rates of future dephosphorylations. is this rest of the stuff relevant? : This can confer negative cooperativity to the system, endowing signaling systems with two features: high turnover of ligands even in high ligand concentration (which might be advantageous if ligands are involved in other reactions), and constant signaling activity in low ligand concentration. Similarly, this might reduce the impact of inhibitors, requiring much higher concentrations of inhibitor to completely turn off signaling. A negative cooperative dephosphorylation step could bias the system towards phosphorylation, since dephosphorylation events hinder future dephosphorylations, but phosphorylation is positively cooperative. However, it could decrease the maximum average? phosphorylation state, since as more tyrosines are phosphorylated, it is easier for both the kinase and phosphatase to bind.

Not sure if any of this matters Disordered-to-ordered transitions create these signaling characteristics when kinase and phosphatase concentrations are unbalanced. Future work will examine how the ratio of their concentrations impacts the cooperative effects of the system. However, evidence for positive cooperativity suggests a regime of high kinase to phosphatase post-TCR triggering. This is further supported by evidence of spatial segregation of phosphatase CD148 after TCR triggering. cite

We also find preferential sequential phosphorylation of a membrane-anchored disordered protein occurs naturally. Local structuring is not required for preferential sequential binding, but does enhance the magnitude of the preference. Does it? This seems like an important question. Sequential binding can create or enhance ultrasensitivity.which does what? cite Omer?

Both disordered-to-ordered transitions and membrane-anchoring of disordered proteins can create nonlinear cooperative effects in the signaling network.

Tie back to other systems, greater community.....

5 Electrostatics

5.1 Introduction

Subunits of many receptors have been shown to associate with the membrane prior to signaling. [**Xu2008**],[**Shi2013**],[**Zhang2**
Others? Dobbins et al, Chenqi paper Close association with the membrane inhibits ligands from binding to

their target site. Several of these chains have been shown to dissociate from the membrane after the extracellular receptor receives a stimulus. cite cite This mechanism may help prevent unwanted signaling. wording here is mediocre and possibly incorrect?

In particular, (CITE CITE CITE)CD3ζ associates with the membrane prior to phosphorylation, sequestering the tyrosines in the bilayer. [**Aivazian2000**],[**Zhang2011**],[**Shi2013**] Do I cite them here? Do they actually show it's associated? I guess? Post-phosphorylation, CD3ζ is no longer associated with the membrane, but remains anchored by its transmembrane region. [**Aivazian2000**]

Stimulating the TCR with an antigen will cause the CD3ζ chain to dissociate from the membrane. However, mutation of the tyrosines to phenylalanine to prevent phosphorylation abrogates this effect. Thus, phosphorylation is required to for ζ chain dissociation from the membrane even with TCR triggering. [**Zhang2011**]

Basic residue regions in the CD3ζ and ϵ chains play a major role in maintaining the association between the polymer and the membrane. Studies show that mutation of the basic residue regions is sufficient to cause the protein to dissociate from the membrane.[**Zhang2011**] Fully phosphorylated tyrosines are also sufficient to pull the protein away from the membrane, even when basic residue regions are present. cite

whoa whoa. Zhang et al claims mutating BRSs actually inhibits phos??? And only associated in 25%?

Calcium influx causes the CD3ζ chain to dissociate from the membrane prior to phosphorylation. An initial signal causes a calcium dump, creating a positive feedback onto the signaling pathway as more ζ-chains dissociate from the membrane and become phosphorylated, enhancing the signal. [**Shi2013**] Prior to the calcium influx, the ζ-chain tyrosines remain sequestered in the membrane, preventing signaling. Where the initial signal comes from to instigate the feedback loop remains mysterious.

There are two hypotheses which could explain the initial signaling event. One hypothesis is that a first chain will fall out of the membrane due to steric occlusion created by TCR clustering. When TCR triggering occurs, TCRs cluster. CITE With all six disordered subunits associated with the membrane, when the TCRs cluster, the subunits may become too crowded on the membrane and induce one or more to dissociate and become accessible to phosphorylation.

Independently or in conjunction with the above hypothesis, we formulate a second hypothesis. Although the tyrosines will spend much of their time sequestered in the membrane, there is some probability of entropic forces ehhh transiently pulling them to the cytoplasm. hence distribution below... When this occurs, Lck would be able to phosphorylate the tyrosine, creating a large negative charge on the chain. This phosphotyrosine could repel the negative polar heads of the membrane, making it more favorable for that section of the polymer to pull out of the membrane.

We first explore this second hypothesis. We hypothesize that entropic fluctuations of the tyrosines in and out of the membrane are sufficient to lead to an initial phosphorylation event. We also hypothesis that following this initial event, there would be a cooperative enhancement of the accessibility of neighboring tyrosines.

5.2 Model and Methods

A molecular dynamics simulation of CD3ϵ indicates an approximately Gaussian distribution of the tyrosines, centered at the phospholipid heads of the membrane. [7]

We model the polymer-membrane association as a potential acting on each rod in a freely-jointed chain in half space. The potential only acts in the direction of the half space plane (in this case, z-direction). To develop the model, we need to explore parameter space to create potentials which display the same behaviors seen experimentally and through molecular dynamics. In particular, we want to match three conditions:

1. The polymer should display locational distributions consistent with Lopez et al. 2015. [7]

2. When basic residues are mutated, the polymer dissociates from the membrane. [**Zhang2011**]
3. When all tyrosines are phosphorylated, the polymer should dissociate from the membrane.

Based on these goals, we group residues into four distinct categories: tyrosines, phosphotyrosines, basic residues, and remaining residues. Based on these groups, we develop a set of possible relationships to explore (Table).

Table 1: Notation for electrostatic potentials applied to different groups of residues.

Residue Group	Electrostatic Potential Abbreviation
Tyrosines	E_Y
Phosphotyrosines	E_P
Basic Residues	E_B
Remaining Residies	E_R

Table 2: Electrostatic potential relationships to explore for behavior matching experimental results.

Group Number	Electrostatic Potential Relationship
1	$E_Y = E_B < 0$ $E_P = E_R = 0$
2	$E_Y \neq E_B$ $E_Y < 0$ $E_B < 0$ $E_P = E_R = 0$
3	$E_Y \neq E_B$ $E_Y < 0$ $E_B < 0$ $E_P > 0$ $E_R = 0$
4	$E_Y = E_B < 0$ $E_P > 0$ $E_R = 0$
5	$E_B < 0$ $E_Y = E_R = 0$ $E_P > 0$

Table 3: Cytoplasmic amino acid sequence of CD3ζ and ϵ chains. Basic residues (arginine, lysine, histidine) and tyrosines are highlighted separately in sequence. Relative numeric location in cytoplasmic sequence and fraction of residues from that group are noted.

Chain	Group	Location in Cytoplasmic Sequence	Numeric Location	# / Total
CD3ζ	Basic Residues	RAKFSRSAETAANLQDPNQLYNE LNLGRREEYDVLEKKRARDPEMG GKQQRRRNPQEGVYNALQKDKM AEAYSEIGTKGERRRGKGHDGLY QGLSTATKDTYDALHMQTLAPR	1,3,5,28,29,37,38, 39,41,48,51,52,53,60,65, 67,72,78,81,82,83,85,91, 99,102,106,113	29/113
	Tyrosines	RAKFSRSAETAANLQDPNQLYNE LNLGRREEYDVLEKKRARDPEMG GKQQRRRNPQEGVYNALQKDKM AEAYSEIGTKGERRRGKGHDGLY QGLSTATKDTYDALHMQTLAPR	21,32,60,72,91,102	6/113
CD3ϵ	Basics Residues	WSKNRKAKAKPVTRGAGAGGRQ RGQNKERPPPVPNPDYEPIRKGQ RDLYSGLNQRRI	3,5,6,8,10,14,21,23,27, 29,42,43,46,55,56	15/57
	Tyrosines	WSKNRKAKAKPVTRGAGAGGRQ RGQNKERPPPVPNPDYEPIRKGQ RDLYSGLNQRRI	38,49	2/57

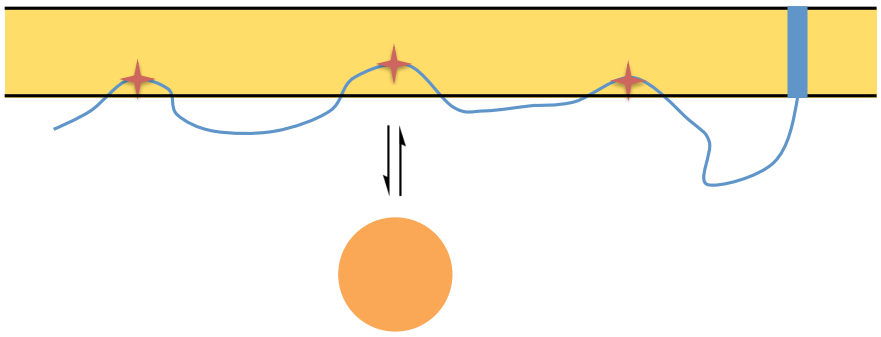


Figure 11: Cartoon of electrostatics model. COULD HAVE POTENTIAL IN CARTOON?. FJC associates with membrane, burying modification sites within membrane rendering them inaccessible to ligands.

5.2.1 Electrostatic Potentials

The model where basic residues and tyrosines have similar energetics (Table **??**, Group 1) is the simplest case to explore. There are many types of electric potential we could explore for the correct behavior. Given that one of the phenomenons we wish to match is an approximately Gaussian distribution of both the tyrosines and polymer along the membrane edge, we use a parabolic-constant piecewise potential for our basic residues and tyrosines, instead of a Lennard-Jones potential.

Parabolic-constant piecewise potential:

Piecewise potential (aka, parabola-constant), (PC):

$$\begin{cases} width * z^2 - depth & z < \sqrt{\frac{depth}{width}} \\ 0 & z \geq \sqrt{\frac{depth}{width}} \end{cases} \tag{14}$$

We have two possible potentials for the remaining amino acids. In general, the peptide should not be able to penetrate deep into the membrane. Therefore, we can implement a hardwall or softwall constraint. The hardwall will prevent the remaining amino acids from passing the beginning of the membrane. A softwall constraint will allow amino acids to enter the membrane, but will incur an energetic penalty. For this initial model, phosphotyrosines will follow this potential also.

Hardwall:

$$\begin{cases} \infty & z < 0 \\ 0 & z \geq 0 \end{cases} \tag{15}$$

Softwall:

$$\begin{cases} k * z^2 & z < 0 \\ 0 & z \geq 0 \end{cases} \tag{16}$$

We later develop a more complicated but possibly more motivated model with a repulsive force on the phosphotyrosines. There are many more basic residues compared to tyrosines in both the CD3ζ and ϵ chains. Therefore, in order to account for tyrosine phosphorylation being sufficient to dissociate the polymer from the membrane, we would expect a repulsive force from the phosphorylated tyrosines. We therefore shift future focus to models of Groups 3,4, and 5. We first consider the simple case where tyrosines do not experience their own potential but experience the potential of either the basic residues or experience the same potential as the rest of the amino acids. Since tyrosines are not positively charged, it seems less likely that they would experience the same force as the basic residues. Although their aromatic ring structure may help to anchor the polymer to the membrane once associated with the membrane, it would not drive the polymer to the membrane to begin with. [7] Therefore we focus future efforts on Group 5.

In this new model, we maintain the same potentials as above for the basic residues and the general amino acids, but now the tyrosines behave as generic amino acids and the phosphotyrosines feel a repulsive force. This feels overly wordy For the phosphorylated tyrosines, we will include an exponential distribution where the repulsive effect of the membrane drops off quickly as the phosphotyrosines move further away from the membrane.

Exponential decay:

go find actual equation to put here

$$\begin{cases} \infty & z \leq 0 \\ e^{-z} & z > 0 \end{cases} \tag{17}$$

5.3 Results

5.3.1 Model in which basic residues and tyrosines have similar energetics reveals puzzle in existing data

We first explore the simplest model, where basic residues and tyrosines feel the same piecewise parabolic-constant potential and the phosphotyrosines feel the same softwall (or hardwall) potential as the rest of the residues. (Table **??**, Group 1) We first explore parameter space to find a set of parameters where the distribution of tyrosines is roughly gaussian centered at the membrane and extending about two Kuhn lengths in each direction, as seen in Lopez et al. 2015. (Condition 1) From this sweep, we see there are multiple parameter sets which achieve this distribution. Fig. **??** For the following exploration of how phosphorylation impacts the distribution, we show the minimal parameters of those tested which give the desired distribution. We establish the following potentials:

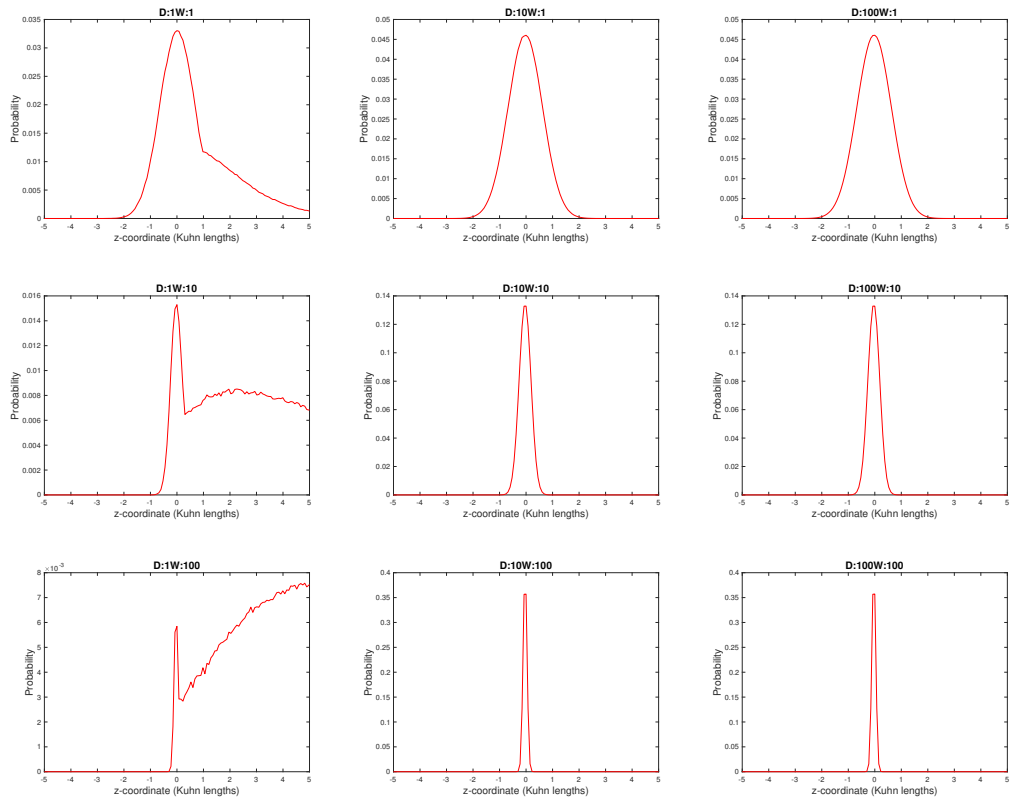


Figure 12: **??**

Parabolic-constant piecewise potential:

Piecewise potential (aka, parabola-constant), (PC):

$$\begin{cases} 1 * z^2 - 10 & z < \sqrt{\frac{10}{1}} \\ 0 & z \geq \sqrt{\frac{10}{1}} \end{cases} \tag{18}$$

Softwall:

$$\begin{cases} 0.01 * z^2 & z < 0 \\ 0 & z \geq 0 \end{cases} \tag{19}$$

We explore how phosphorylation, or removing tyrosines from the parabolic-constant potential and adding them to the softwall potential, impacts the distribution of the tyrosines and full chain. We find that under this regime, phosphorylating individual tyrosines does not impact the distribution of neighboring tyrosines. Fig. **??** Additionally, phosphorylating two tyrosines does not influence the distribution of a tyrosines located between the two. Fig. **??** We also find that even when all tyrosines are phosphorylated, the polymer does not fully dissociate from the membrane. This is represented by the distribution of the (basic) polymer tip residue distribution, which remains closely associated with the membrane. A non-basic, non-tyrosine residue located centrally is only minorly affected by the change in phosphorylation state. This result indicates these electric potentials may be insufficient to exhibit the characteristics displayed experimentally.

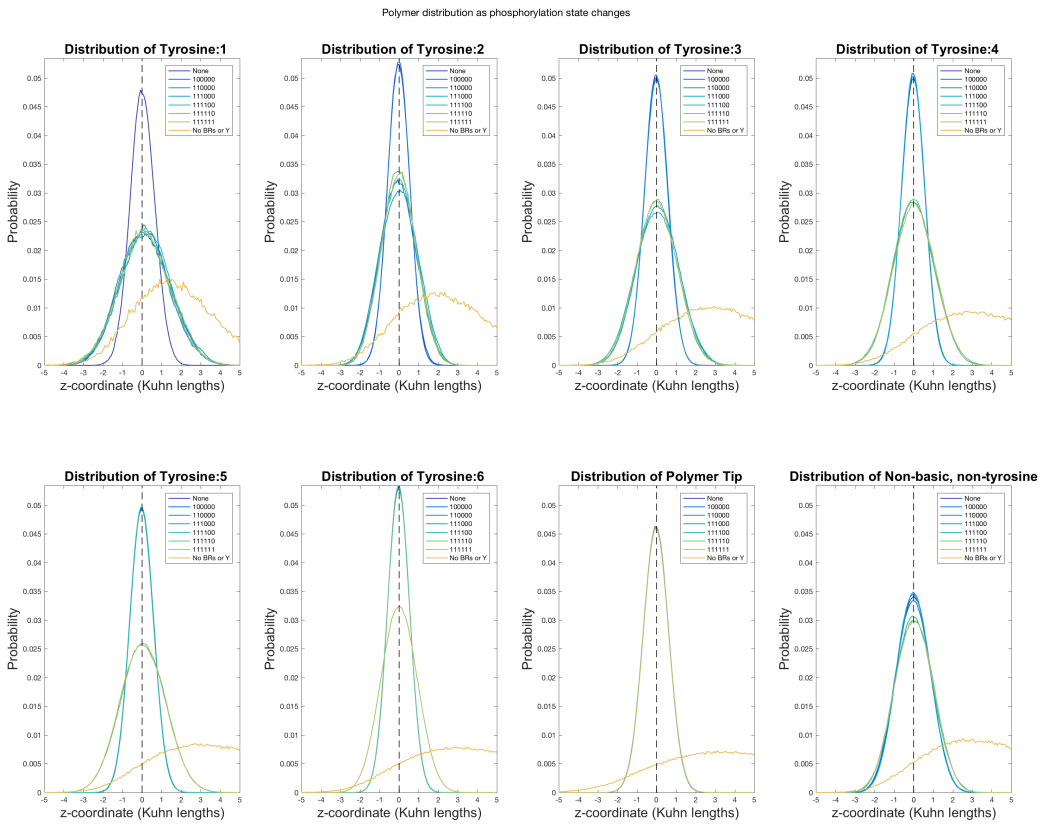


Figure 13

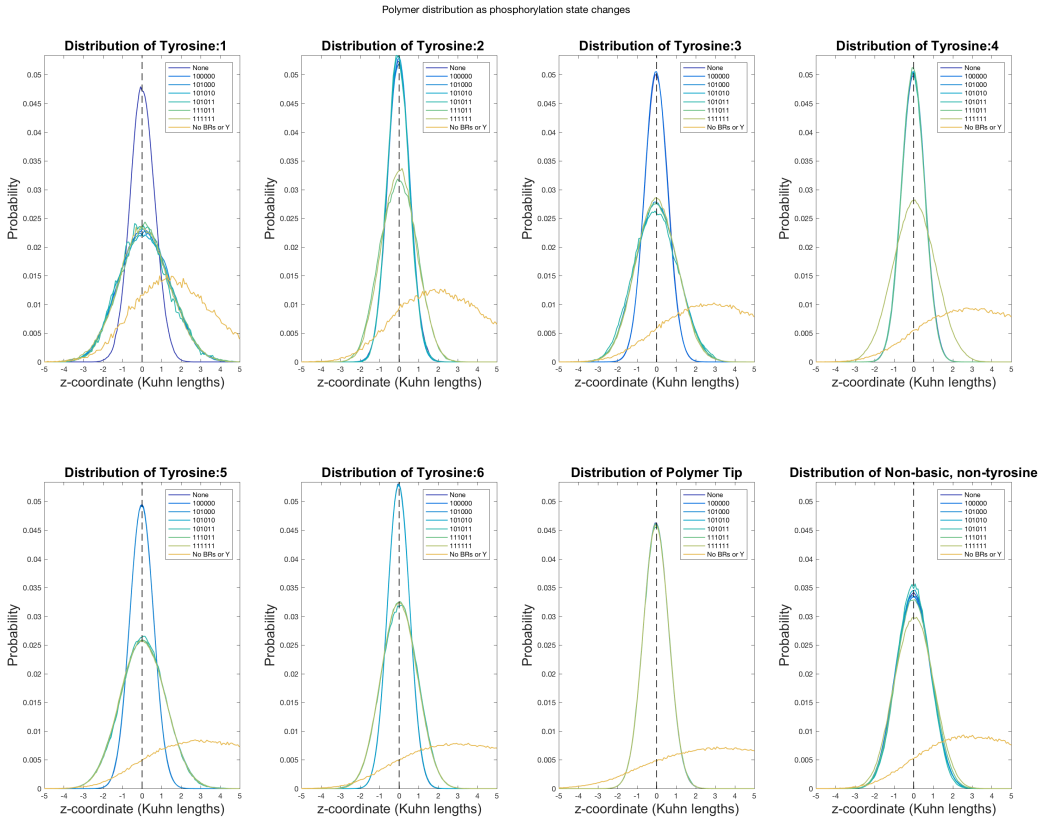


Figure 14

5.3.2 Initial parameter exploration of more general energetic model yields fit to polymer distribution

We first want to match the distribution of the polymer to the molecular dynamics simulations from Lopez et al. We sweep through parameters for the basic residue potential and the softwall potential. The softwall potential permits a Gaussian curve for the tyrosine distributions. The hardwall condition (not shown) prevents the distributions from spreading below the zero axis where we have defined the membrane to be. When we explore these parameters, multiple possibilities for parameter sets emerge which meet the distribution conditions. Below are examples of distributions arising from our parameter exploration for both a tyrosine and a basic residue. Fig. **??**, Fig. **??**rephrase this next sentence, because duh We note that the potentials affect the distribution of the basic residues more strongly than they affect the distributions of the tyrosines. From these parameter sets, we may refine our parameter search and begin exploring how the distributions are affected when all tyrosines are phosphorylated.

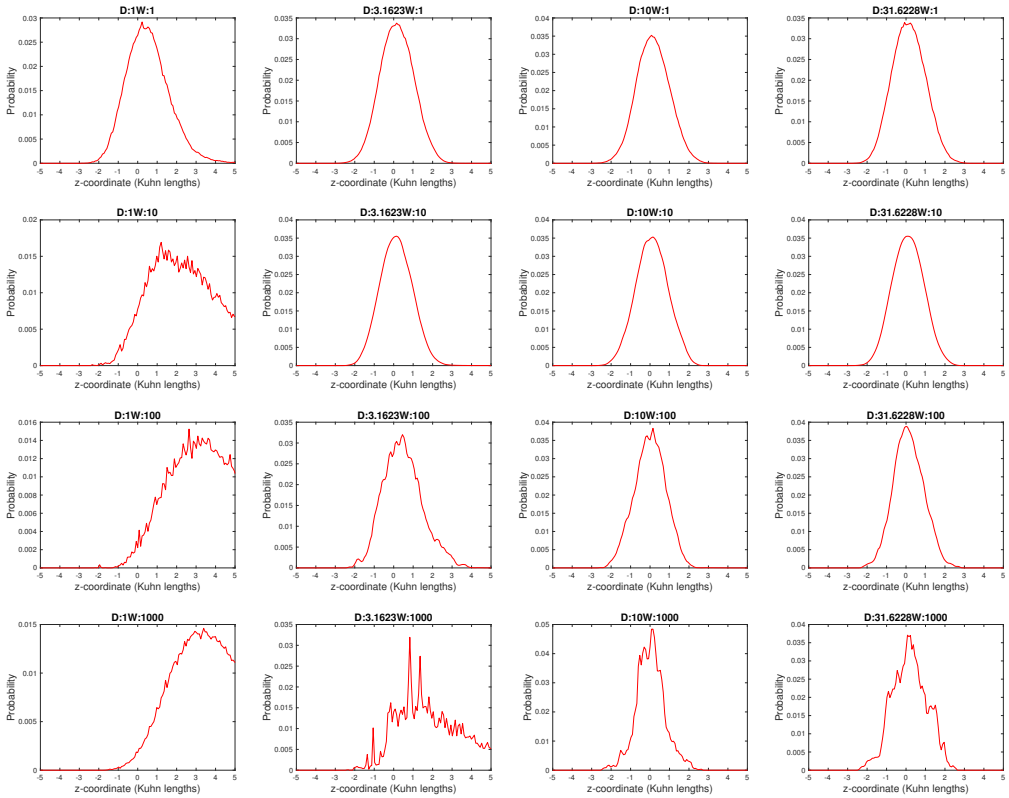


Figure 15: Tyrosine distribution resulting from parameter exploration of electric potentials. For this set, softwall potential is set at k=0.1. Basic residue potential depth ranges from 10^0 to $10^{1.5}$.

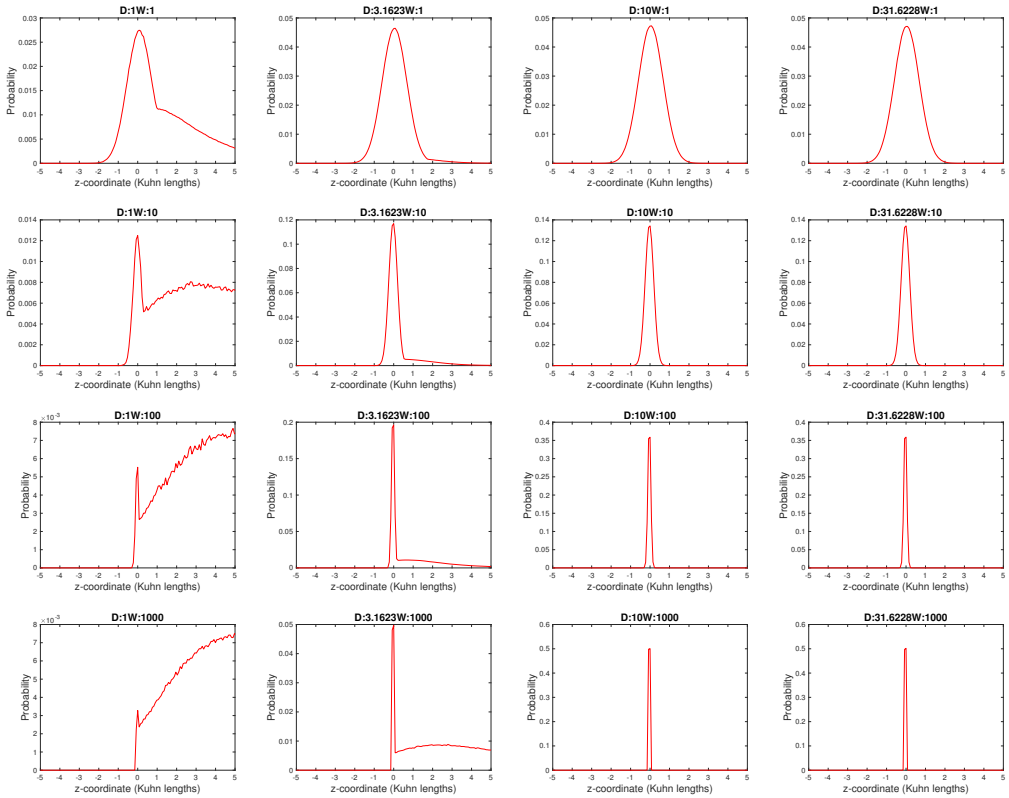


Figure 16: Basic residue distribution resulting from parameter exploration of electric potentials. For this set, softwall potential is set at k=0.1. Basic residue potential depth ranges from 10^0 to $10^{1.5}$.

5.3.3 Future Work

First, we want to establish parameters (if they exist) which make our model exhibit the same behaviors as experimental and MD results. In particular, we now need parameters allowing the polymer to be free from the membrane when all tyrosines are phosphorylated. Either there will be only one reasonable parameter set which we can explore or there will be many parameter sets which display the desired properties. If there are multiple parameter sets then we may try to find other experimental results to narrow our parameter regime. Alternatively, we may explore two or three major parameter regimes: weakest, strongest, and median potentials. The results from a wide spread of parameters will give us an indication of how much the parameters matter to the qualitative results.

Second, with our parameter sets, we will now explore the effect of phosphorylation on the accessibility of successive tyrosines to a kinase. We wish to see if there is a cooperative enhancement of the binding rates based on phosphorylation of previous residues. Cooperative phosphorylation could contribute to the first signaling event, explaining how an initial calcium release feeds back on the system, causing more CD3ζ chains to dissociate from the membrane and create a stronger signal. cite

Additionally, we will use this data to explore if there is a natural sequential binding sequence that arises. One would imagine that once a single tyrosine is phosphorylated then its nearest neighbor would be the most likely next target. However it is unclear if there is a first tyrosine that will be most likely to be phosphorylated and whether this will depend dominantly on basic residue distribution or on proximity to the transmembrane region.

Should I include this? I don't know if we're actually going to do it. We will also explore how clustering

of TCRs influences the association of the disordered subunits to the membrane. By modeling multiple disordered chains, we will attempt to estimate how many TCRs must be in a cluster and how tightly clustered they must be for this model to explain the initial triggering event.

paragraph about relating back to something useful or experimental?

5.4 Discussion?

Maybe hybrid future work and discussion? or give if/then results sections? (I.e. if we see coop, then this?)

Our initial model attempts show it is nontrivial to create a simple model where phosphotyrosines act to pull the rest of the polymer away from the membrane.

6 Results: Simultaneous Binding

6.1 Introduction

Find more examples of multiple binding

Here we consider the case where ligands are able to bind to a polymer and remain bound. For instance, the SH2 domains of ZAP-70 bind to two tyrosines in an ITAM on CD3ζ and remain bound to phosphorylate downstream. cite!! This leads to the possibility that the polymer conformations are limited by the presence of a bound ligand, or even multiple bound ligands. We expect that as more ligands are added to the polymer, binding site occlusion will increase. The ligand may help to straighten out the polymer by limiting the number of conformations it can take on, but it will also occlude the binding site in many conformations. Depending on the size, number, and relative location of the ligand to the binding site, this effect may be stronger than the decreased occlusion via straightening.

Once TCR CD3ζ is phosphorylated, the kinase ZAP-70 binds and phosphorylates the transmembrane protein LAT. CITE EVERYTHING! ZAP-70 is composed of two SH2 domains and a kinase domain.cite - - i.e. Wang et al 2010 or alternate These tandem SH2 domains each bind a phosphotyrosine on CD3ζ. Since minimally two SH2 domains must be able to bind to CD3ζ, it is natural to wonder how many domains could fit on a disordered chain and how each successive domain impacts the binding of another.

Describe different scenarios - pose slightly different questions. Two parameters to sweep over. We do diagonal and one T-shape.

6.2 Model and Methods

literally copy pasted from ref(sec: LocalStruct), is that a problem?

We model the TCR CD3ζ chain as a freely jointed chain, where each amino acid is represented as a segment as described in Sec. **??**. The unbound, incoming ligand is simulated as an idealized ghost sphere tangentially located at a single segment. Is it clear what this means? We calculate quasi-equilibrium statistics for the FJC and unbound ligand in both free-space and half-space. For the CD3ζ chain, we simulate only the cytoplasmic region by an FJC with 113 segments. In our simulation, all lengths are normalized by the Kuhn length, so each rod has length 1 and the radius of the ligand is also measured in Kuhn lengths. We consider a spherical estimate for the kinase Lck, with a radius of 2.1 nm(7 Kuhn lengths) as calculated in **??**.

To simulate simultaneous binding, each bound ligand is included as an idealized sphere tangentially bound to the polymer. The bound ligands rotate with the polymer, but maintain their tangent orientation to their binding site. The polymer may not occupy the same space as the bound ligands, nor may bound ligands occupy the same space as other bound ligands. When a membrane is present, the bound ligand cannot penetrate the membrane. There are separate parameters for the radius of the bound and unbound ligands.

We calculate the probability a ligand is able to bind to the unbound ligand site, where 'able to bind' refers to the specified sphere being empty of both other polymer segments, half-space barriers, and bound ligands. weird wording?.

would it be helpful to have a dotted sphere in cartoons to show unbound ligand?

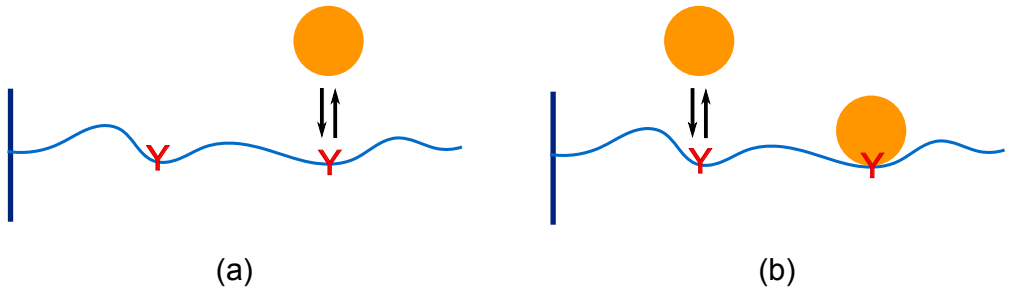


Figure 17: Cartoon of simultaneous binding model. Surface bound FJC prior to ligand binding (a) transitions to FJC with single bound ligand and a second ligand attempting to bind (b).

6.3 Results

6.3.1 Negative cooperativity arises from simultaneous binding of ligands to an IDP

Do we want to explore free space here? Did that have much different results? Probably not?

We first explore how two ligand species binding to the same domain may influence each other's binding. We explore a simplification of ZAP-70 and Lck molecules binding to CD3ζ. We assume that each of the bound ligands is approximately the size of an SH2 domain (i.e. ZAP-70 is bound to the polymer) and we calculate the probability of a separate species binding to one of the remaining binding sites. We investigate this for many sizes of incoming ligand, including our estimates for Lck and ZAP-70. We find simultaneous binding of multiple SH2 sized domains will cause mild to severe negative cooperativity with the binding of another ligand. Larger ligands have a harder time binding when there are multiple SH2 domains already bound. For a 4.2nm radius incoming ligand, there would be a 30 fold decrease in the probability of binding from no other bound ligands to five SH2 domains bound to the same disordered chain. Only very small ligands would be unhampered by the presence of other bound ligands. An Lck molecule, only slightly larger than an SH2 domain, would experience an approximately eight fold decrease in its ability to bind once five SH2 domains are already bound to the disordered chain. (Fig. **??**)

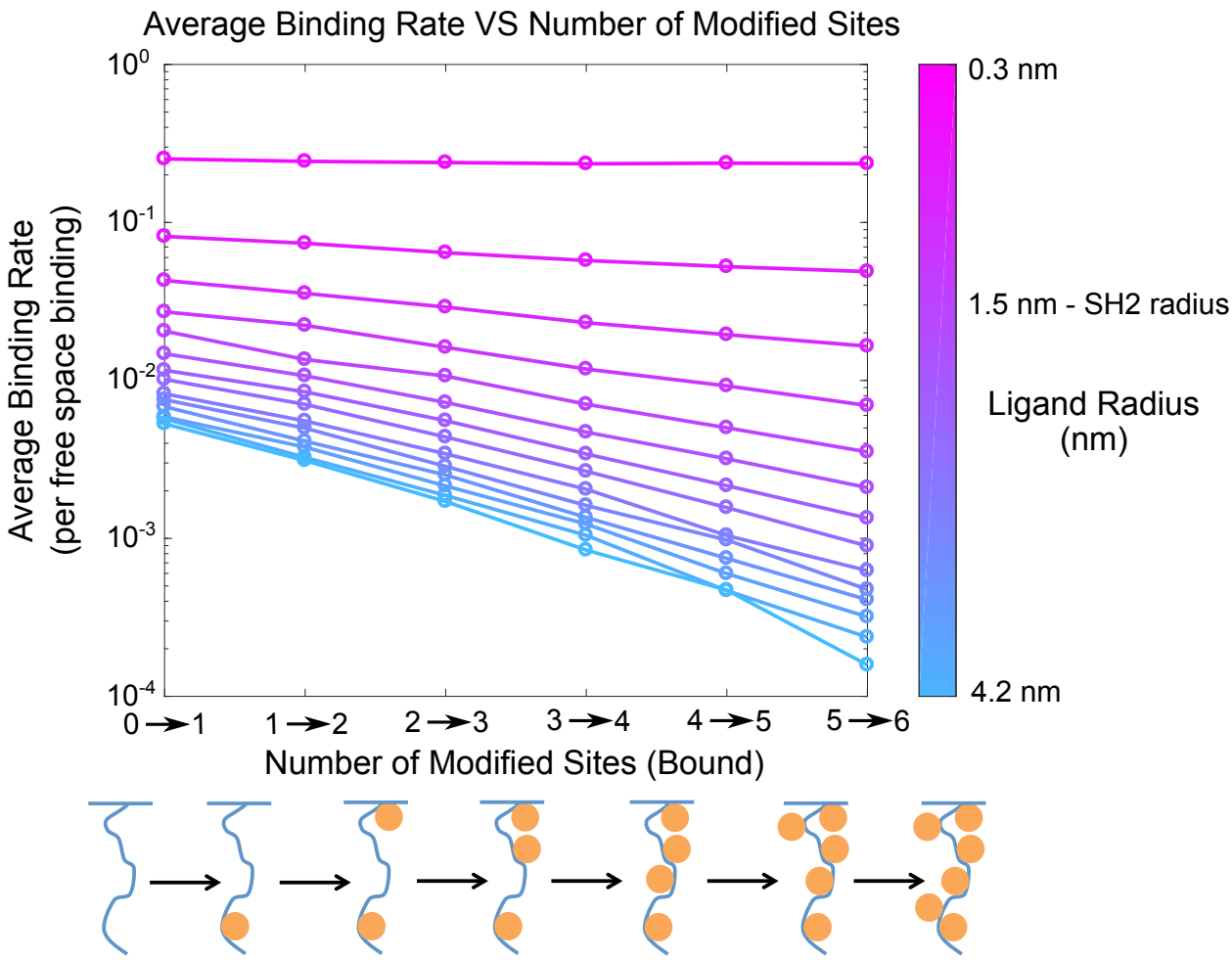


Figure 18: (Above) Simulated average binding rates of Lck to CD3ζ for specified number of SH2 domain sized (1.5nm radius) ligands previously bound and varying unbound ligand sizes (0.3 nm $\rightarrow$ pink, 4.2 nm $\rightarrow$ blue). (Below) Cartoon representing a possible bound configuration series of CD3ζ.

Should this case go first?

The second case of simultaneous binding we explore is when only one ligand species attaches to the disordered chain. We investigate how multiple bound ligands of a given size impacts another ligand of the same size binding. We see the same trend of behavior as above, but more dramatic. Here, if we consider a ligand with radius 2.1nm, we already see a 30 fold decrease between no ligands bound and five bound. This matches intuition, since although a large bound ligand will straighten the polymer more, it will also create a much larger occluded volume making binding site accessibility much less likely. Note that similarly, if the bound ligand size is less than that of an SH2 domain, then we actually see less negative cooperativity than when all bound ligands were SH2 domains. (Fig. **??**)

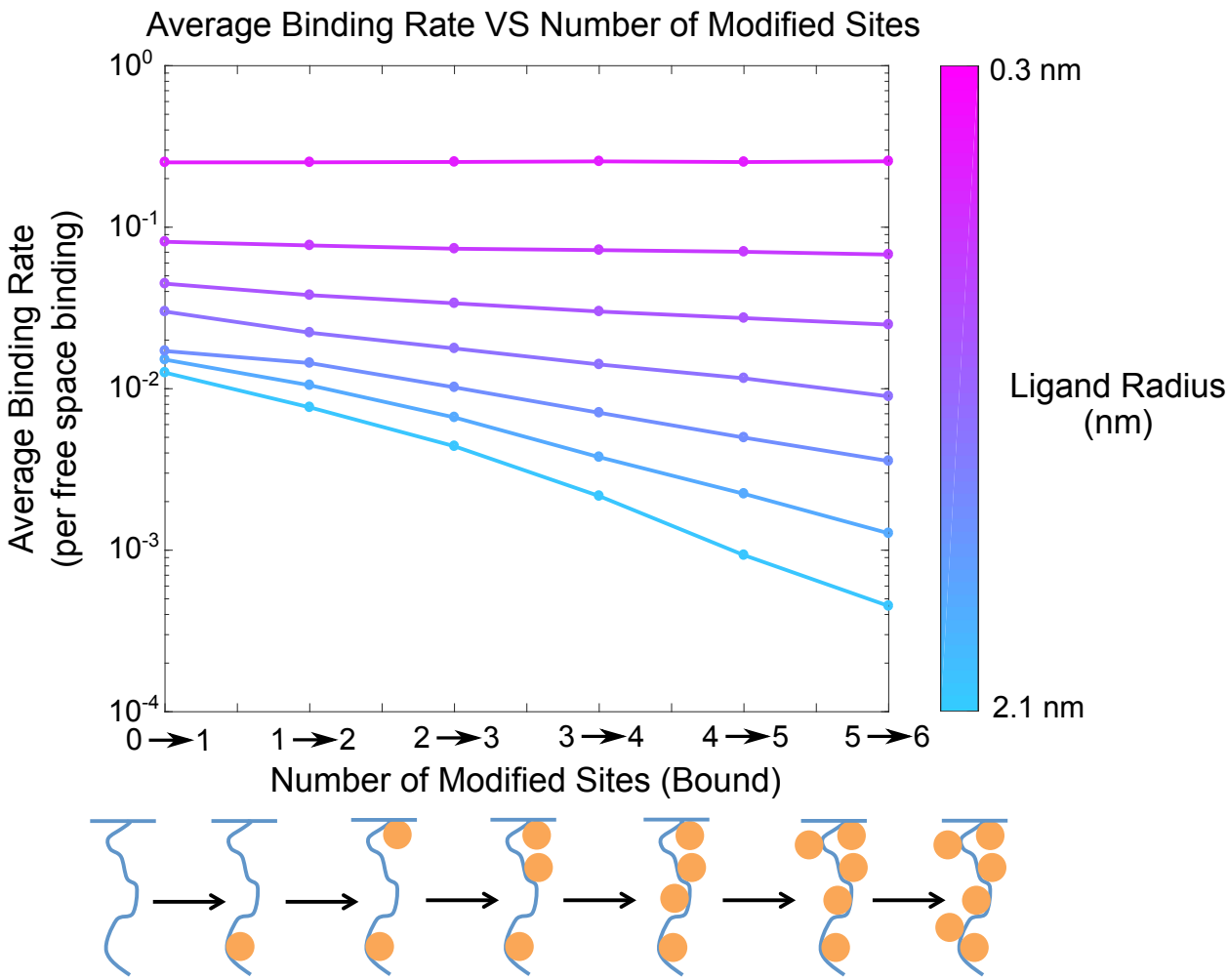


Figure 19: (Above) Simulated average binding rates of Lck to CD3ζ for specified number of ligands previously bound and varying ligand sizes (0.3 nm $\rightarrow$ pink, 4.2 nm $\rightarrow$ blue). (Below) Cartoon representing a possible bound configuration series of CD3ζ.

6.3.2 Future Work: Sequence of binding multiple ligands

rewrite this for simultaneous binding

We obtained relative probabilities of binding for each tyrosine along the chain at each bound state. With this information, we are able to explore whether or not there is a dominant sequence of binding events. We ran a Gillespie algorithm with six events, one for each ligand binding. From each run, we record which sequence or 'path' of binding. When we compile the path frequency data, it becomes clear that the first binding event is dominant. That is to say, any path beginning by binding to the membrane distal tyrosine (6th tyrosine), will occur more often than any path beginning with the membrane proximal tyrosine.

For simplicity, we consider only the paths membrane proximal to membrane distal (123456) and distal to proximal (654321). We see that when a membrane is present, the probability of binding distal to proximal is much higher than that of phosphorylating proximal to distal. This phenomenon is easily explained by the presence of a membrane. The membrane proximal tyrosine has a smaller range of space it may occupy than the distal tyrosine. Therefore, it is more likely to be close to the membrane in a configuration. Since the

ligand cannot penetrate the membrane, tyrosines closer to the membrane have a higher probability of being effectively sheltered from ligand binding. This makes the distal to proximal path much more likely.

The preference for distal to proximal binding is enhanced by increasing ligand size. Larger ligands remain able to bind the membrane distal tyrosine with similar probability but will have a lower probability of binding tyrosines close to the membrane. When we simulate CD3ζ without a membrane, this preference is eliminated. Note that unlike with the local structuring model, a strict comparison of distal to proximal vs proximal to distal binding does not indicate a preference based on uneven spacing of tyrosines and in fact both of these paths are less likely than if all paths were equally probable. Although the most likely tyrosine to be bound is still the distal tyrosine, there is not a strong preference for the distal to proximal sequence because perfectly sequential binding is unlikely. Each bound ligand is more likely to sterically occlude a neighboring tyrosine than one further away. Therefore, if the distal tyrosine is the first bound, the next most likely to be bound is not its neighbor.

We are soon going to introduce a ranking system for the paths in order to do a full analysis of all possible paths. We look to answer if, for simultaneous binding, there a single path or group of paths that stand out as most likely.

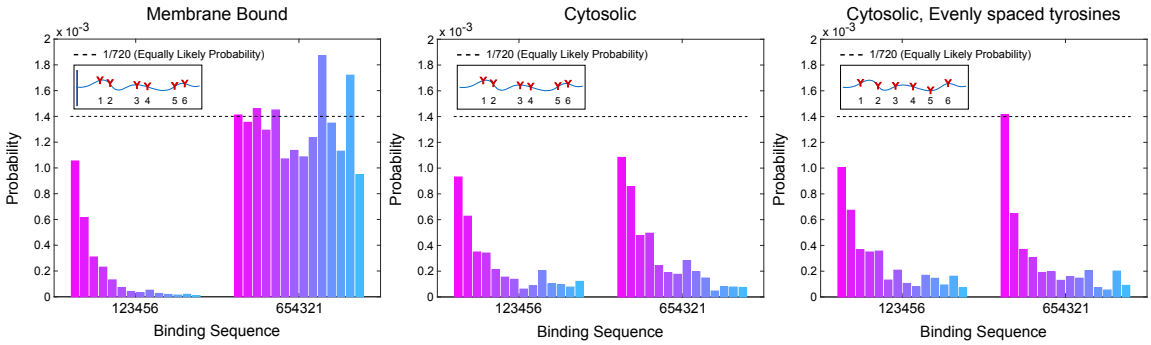


Figure 20: Probability of ligand binding membrane proximal to distal (123456) compared with distal to proximal (654321). Equally likely probability of all paths indicated with black, dotted line. (a) CD3ζ parameters, with membrane. (b) CD3ζ parameters, without membrane. (c) CD3ζ length, evenly spaced tyrosine locations, without membrane.

6.3.3 Future Work: Multiple chains towards a full model of a T cell receptor signalosome

There is work showing that six ZAP-70 molecules bind to the disordered proteins of the T Cell Receptor on average. cite.... We want to eventually report an equilibrium value for the number of molecules able to bind to a CD3ζ chain. We will also expand our model to include all of the domains of the receptor. This will allow us to test if steric effects are sufficient to explain why only six ZAP-70 molecules would bind to the receptor simultaneously. When all of the chains are included in the simulation, each chain will be able to explicitly occlude a binding site on the opposite chain. With this, we will be able to answer questions of if or how disordered dimers impede binding.

Eventually, we would like to include co-receptors of TCR in our model. This will help expand our understanding of the TCR network, in particular elucidating why co-receptors are necessary for full T cell signaling.CITE Talk about CD28? There is evidence that Lck is constitutively? some fraction? associated with membrane-bound proteins CD4 and CD8 [2]. There is also evidence that the main CD3ζ phosphatase, CD148, is also membrane bound. Cite These changes to the model could impact the dynamics of reversible phosphorylation of the CD3ζ chain.This needs to be better.

6.4 Discussion

An Lck molecule, only slightly larger than an SH2 domain, would experience an approximately eight fold decrease in its ability to bind once five SH2 domains are already bound to the disordered chain. This suggests the CD3ζ chain is mostly phosphorylated before ZAP-70 begins to attach or that phosphorylation by Lck slows down when ZAP-70 attaches. In the second case, this could lead to phosphorylation needing to occur via another mechanism, for instance via the kinase domain of ZAP-70 itself.

Simultaneous binding of ligands leads to negative cooperativity. This endows signaling systems with two features: high turnover of ligands even in high ligand concentration (which might be advantageous if ligands are involved in other reactions), and constant signaling activity in low ligand concentration. Similarly, this might reduce the impact of inhibitors, requiring much higher concentrations of inhibitor to completely turn off signaling.

Connect to other things

7 Role of the surface on tether-tether reactions within a signaling cluster

7.1 Introduction

7.1.1 Tethering

Disordered proteins aid many reactions by acting as biological tethers. They may be permanent tethers tying two globular domains together or a domain to a surface example, example, cite cite. They may also be transient tethers, catching diffusing particles and tethering them to another molecule, i.e. formin. cite By tethering molecules or surfaces together, disordered domains effectively increase the local concentration of a particle. This increases reaction rates by improving the probability of two molecules bumping into each other through restriction of their diffusive domain.

Given that a surface reduces the exploration space of a molecule, it will naturally increase the local concentration without a tether. However, it will also reduce the accessibility of a molecule by sometimes occluding, for example, an enzymatic domain from reaching its substrate in the correct orientation. Given that the surface has both positive and negative effects on reaction rates, how does being tethered to a surface impact a reaction compared to being tethered to a smaller domain?

This type of interaction arises frequently with clustered reactions. For instance, the phosphorylation of a membrane tethered disordered domain by a membrane bound kinase. cite Experimental studies of such reactions are more easily performed on a sparse matrix such as dextran than on a hard surface. Therefore, it is important to understand the effect tethering to a surface has over tethering to a matrix. We explore how the presence of a surface influences the reach of a tethered domain and the overall local concentration.

7.1.2 SHP-1

Paragraph about SHP1/PD-1/ITIM?

reread SPR paper

read SHP-1/PD-1 review paper

Table 4: Structure and number of simulated segments in each explored tether.

Tether Name	Structure	N
PD-1	(37392 B2) Bio-SRAARGTIGARRTGQPLKEDPSAVPVFSV DYGELDFQWREKTPEPPVPSVPEQTEY*ATIVFPSG	64
PEG-3	Bio-(PEG)3-DLQEVTY*IQLDHH	16
PEG-28	Bio-(PEG)28-DLQEVTY*IQLDHH	41

7.2 Model and Methods

local vs effective concentration?

We first investigate the effects of a surface on a single tether without a bound ligand in free-space and half-space. We model the tether as a freely-jointed chain and the unbound ligand as an idealized ghost sphere tangentially located at a single segment. Is it clear what this means? To explore the effects of a surface on a single tether, we vary the number of segments and the size of the incoming ligand. We calculate quasi-equilibrium statistics for the FJC and unbound ligand in both free-space and half-space. We calculate how often a ligand is able to bind to an oriented sphere tangentially attached to the polymer, where 'able to bind' refers to the specified sphere being empty of both other polymer segments and half-space barriers weird wording?.

We next explore the effects of a surface on three specific tethers: PD-1, PEG-3, and PEG-28. PEG-3 and PEG-28 are tethers created by attaching a single immunoreceptor tyrosine inhibitory motif (ITIM) to a series of PEG linkers. Each tether is modeled as a freely-jointed chain characterized by their lengths as measured in Kuhn lengths. For each tether, we assume a Kuhn length, δ, of 0.3nm, as supported by CITATION HERE. Given that these tethers are made up of PEG linkers and amino acids, each of which is approximately the size of a Kuhn length, the lengths of each tether may be approximated as N Kuhn lengths, where N is the total number of PEG linkers and amino acids. Note that we consider the size of biotin to be negligible for our simulation. (Table **??**)

We investigate SHP-1 as a ligand bound to a tether. In order to simulate SHP-1 as a simple sphere, we estimate its size using three different methods. better word than methods? We first estimate the volume of SHP-1 based on its molecular weight, $67561Da$, as listed in UniProt entry P29350, assuming an average protein density of $1.41g/cm^3$. [5]

$$(67.5 \times 1000\text{Da}) * (1.66 \times 10^{-27}\text{kg / Da}) * (1000\text{g / kg})/(1.41\text{g / cm}^3) = 79.468\,\text{nm}^3.$$

Approximating the phosphatase domain as a sphere, we can estimate a radius:

$$\begin{aligned} V &= \frac{4}{3}\pi r^3 \\ 79.468\,\text{nm}^3 &= \frac{4}{3}\pi r^3 \\ r &\approx 2.7\,\text{nm} \end{aligned}$$

For comparison, we also estimate an upper bound on the radius of SHP-1 from PyMol measurements of the SHP-1 structure (PDB 3PS5). Rounding up for measurement error, the longest part of the structure is 86 Å. (Fig. **??**) From this we calculate a maximum radius of $4.3nm$.

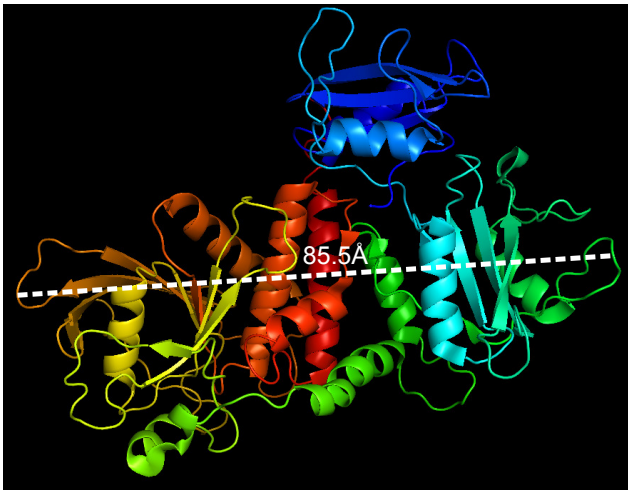


Figure 21: Measurement of maximum length of SHP-1 (PDB 3PS5).

Alternatively, we can estimate the volume of SHP-1 as a rectangular prism and then create a spherical approximation of equal volume. We estimate the length, width and depth of SHP-1 to be 76Å, 36Å, and 58Å respectively, giving a volume of 158688Å^3. (Fig. **??**) From this, we calculate the radius for a spherical approximation to be 3.4nm.

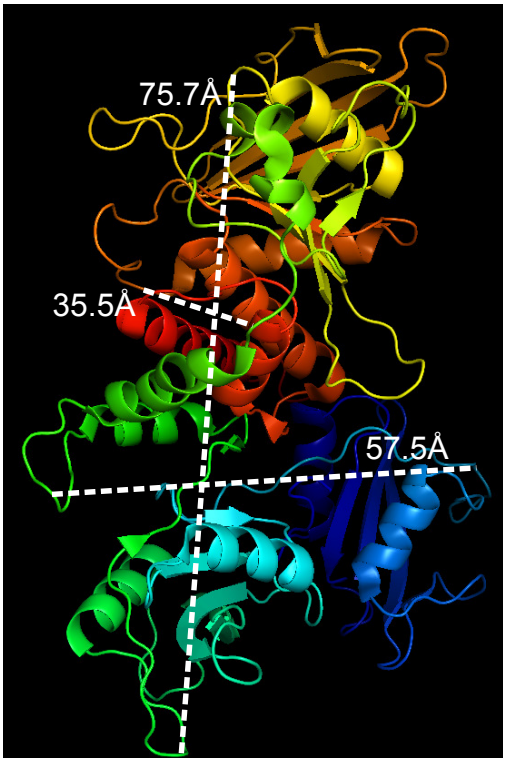


Figure 22: Measurement of rectangular dimensions of SHP-1 (PDB 3PS5).

Should I specify at any point C vs matlab?

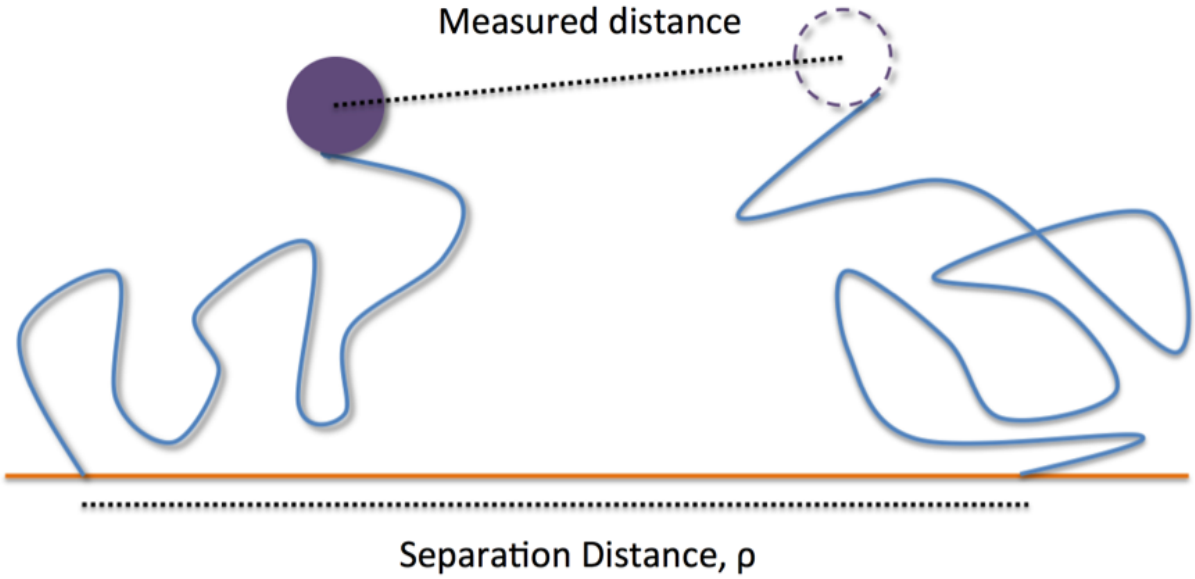


Figure 23: Cartoon of local concentration simulation in half space.

Using the parameters estimated above, we simulate two equivalent tethers. The idealized SHP1 sphere, or ligand, is bound to the end of one of the tethers. The bound ligand is simulated as an idealized sphere which rotates with the tether. The bound ligand may not occupy the same space as the tether nor may it penetrate the surface if present. The second tether has a ghost ligand attached in the same location. (Fig. **??**) We allow each tether to explore conformation space, at each conformation determining if the ghost ligand is occluded by the rest of the tether, or in half-space, by the surface. For the tether with the ghost ligand, we record both the location of the center of the ligand and whether or not it was occluded at each conformation. For the tether with the bound ligand, we only record the location. For these simulations, we assume the two tethers are independent and do not sterically influence each other.

We compute the local concentration as the probability that the bound ligand is near the binding site (where near is defined by within some small radius) and the binding site is not occluded by the rest of the polymer. To calculate this, we separate the the tethers from each other by a distance, ρ, measured in Kuhn lengths by adding ρ to the x-coordinate of the ghost ligand. We then measure the distance between the centers of the bound and ghost ligands. If the centers are within a cutoff distance from each other AND the ghost sphere is not occluded, then the bound ligand is able to bind. The local concentration is calculated by dividing the probability the ligand can bind by the volume of the sphere defined by the cutoff radius.

We fit the local concentration curves to a half-gaussian curve, $\sigma_{fit}(\rho) = a * e^{-\rho^2/c^2}$, as a function of the separation distance between the two tethers. To fit the simulated data, we use Matlab's fminsearch function which uses the Nelder-Mead simplex algorithm. The parameters found by the Nelder-Mead algorithm consistently give a smaller sum of least squares than parameters found using Trust-Region-Reflective Least Squares Algorithm or those found by heuristically sweeping parameter space to minimize the sum of least squares.

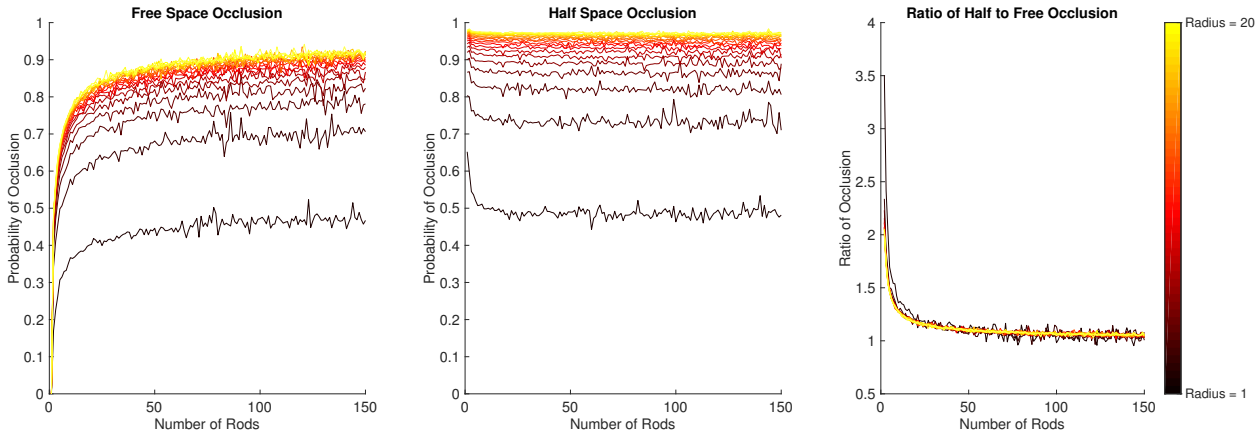
Reference Lagarias et al. ?

7.3 Results

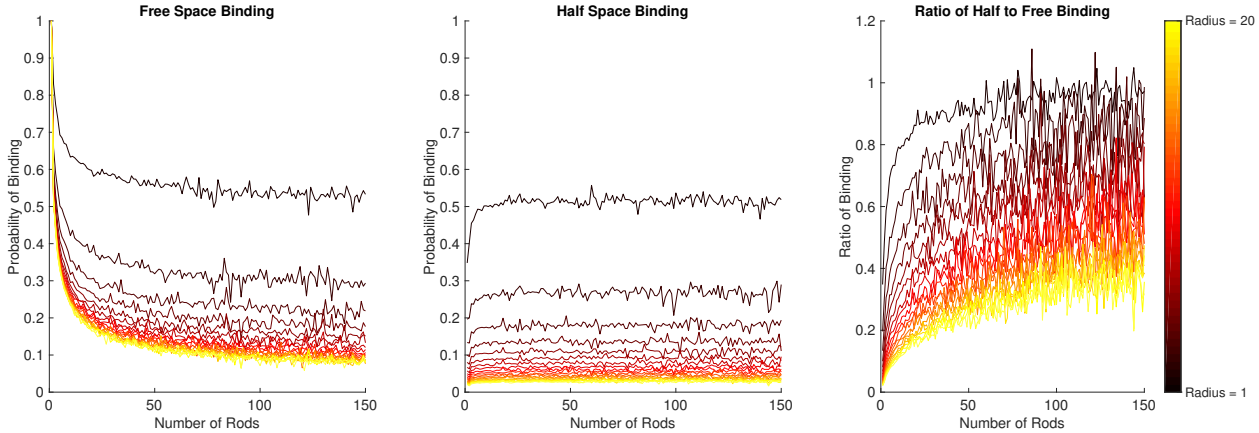
7.3.1 Accessibility of a reaction site is reduced by the presence of a surface

We first investigate how the presence of a surface impacts binding of a ligand to a single tether compared to in free-space. We find that a surface consistently net decreases the probability a ligand can bind to a tether, irrespective of tether length or ligand size. For long tethers, the surface reduction is minimal, in some parameters almost eliminating the difference. A larger ligand creates a larger fold decrease in the ability of the kinase to bind in half-space. (Fig. **??**)

Pick one of these? make one figure in inkscape to control spacing make figure labels match description or vice versa



(a)



(b)

Figure 24: Probability of ligand being occluded, p_{occ}, (top row) and probability of ligand binding, $1 - p_{occ}$, (bottom row) as a function of tether length for various ligand radii. Free-space (left column) half-space (middle), and the ratio of probabilities half-space to free-space (right).

7.3.2 Reactions between tethered reactants are net enhanced by a surface

We explore how effective concentration is impacted by ligand size and polymer length, where effective concentration is a description of how often a ligand bound to the end of one polymer encounters a binding

site on the end of a second polymer, with the two polymers separated by some distance. When ligand size is zero, or equivalent to no ligand attached to the polymer, then we see the free space polymers experience an effective concentration predicted by the worm-like chain model. Note that variability from the theoretical WLC curve is mitigated by decreasing the cutoff distance for the binding encounter. weird wording Naturally, when we introduce a surface or ligand of nonzero size, the worm-like chain model no longer describes the effective concentration. We see that the surface enhances the effective concentration over the free space values but that a larger ligand size reduces the effective concentration. (Fig. **??**)

Should really use SHP1 parameters for this figure.

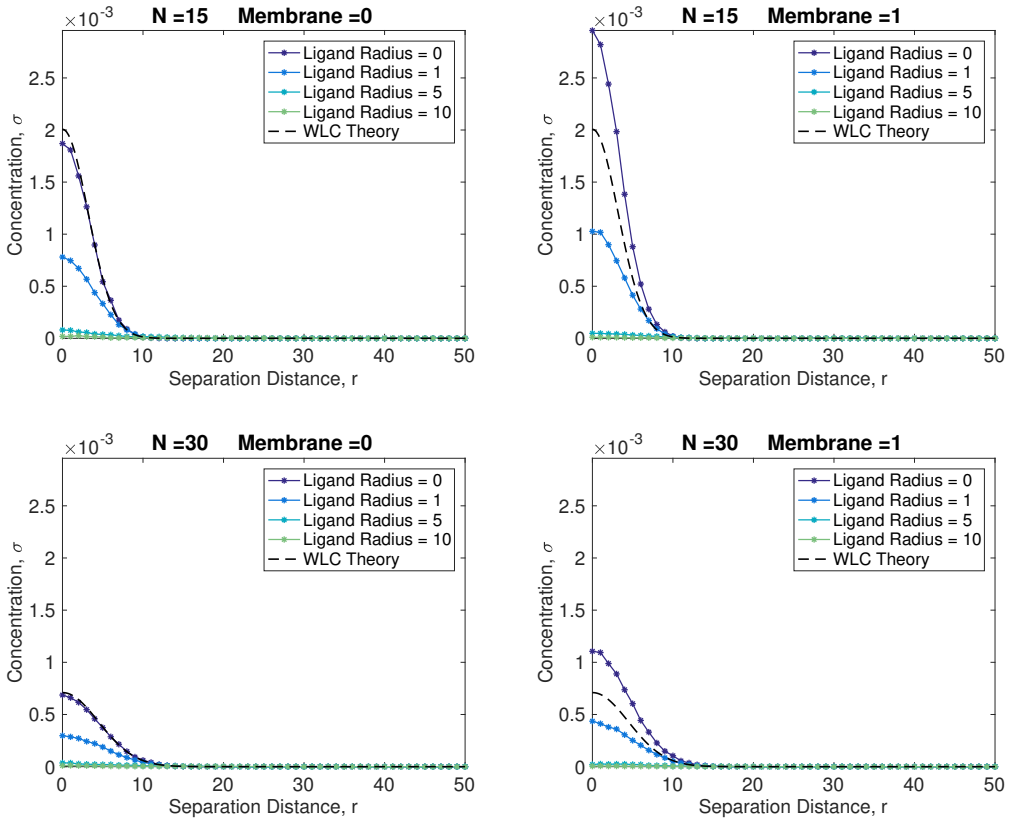


Figure 25: Local concentration as a function of tether separation distance.

7.3.3 Reach parameter is unaffected by presence of surface

We fit a half-gaussian curve to the simulated local concentration data and compare the 'reach parameter', c, in free-space to half-space. Our data shown earlier indicates that the presence of a surface will extend a single polymer compared to in free space. We would expect this straightening to also increase the reach parameter for the effective concentration. However, we find that the reach parameter, c, is unaffected by the presence of a surface. This parameter changes based on polymer length and ligand size, but not presence of a surface.

include curve fitting figure?

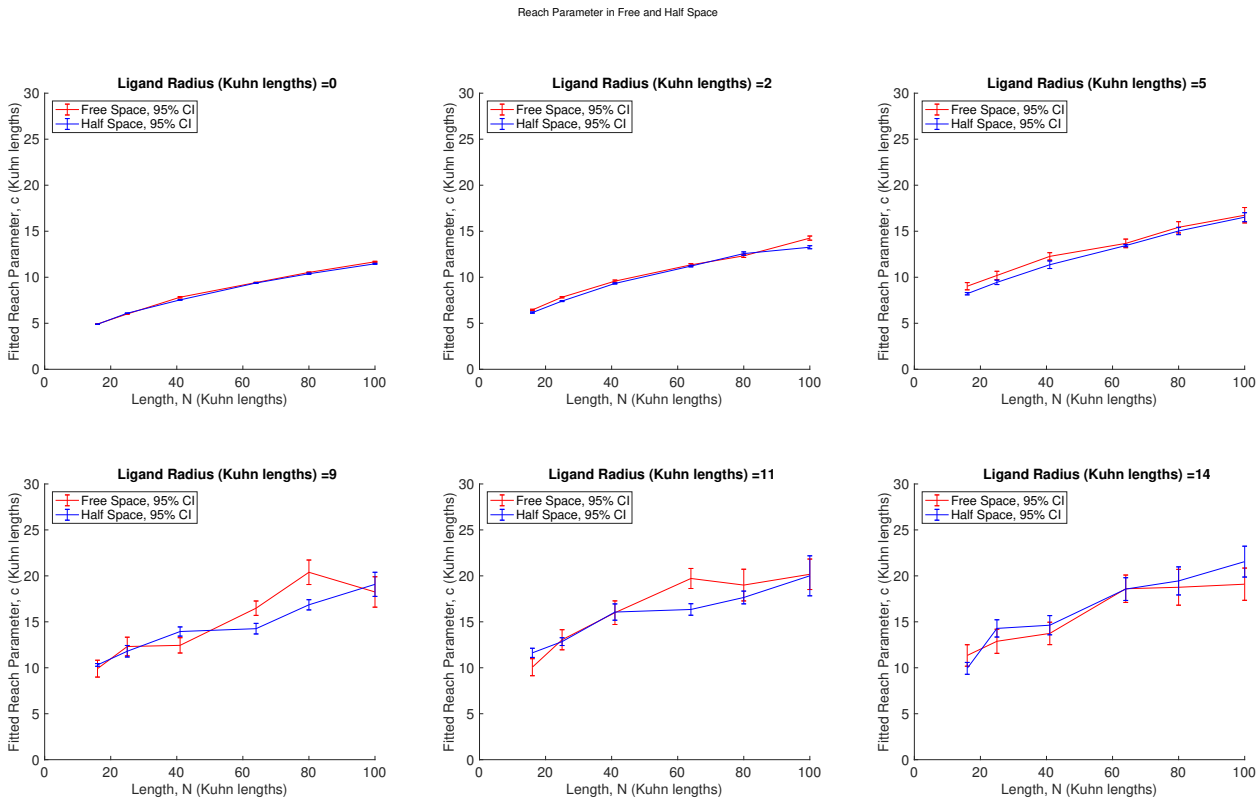


Figure 26: Reach parameter of half-gaussian fit to local concentration data plotted as a function of tether length, N for various ligand radii in free-space (red) and half-space (blue) with 95% confidence intervals for the fit.

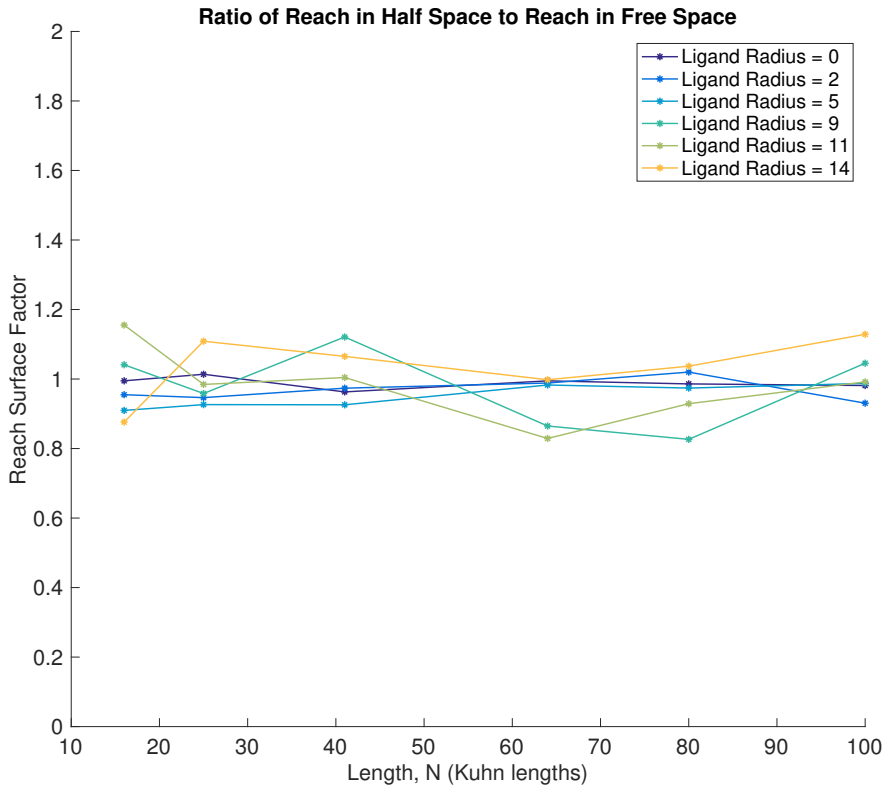


Figure 27: Ratio of half-space fitted reach parameter to free-space fitted reach parameter as a function of tether length, N (Kuhn lengths), for various ligand radii.

7.3.4 Matrix-bound versus surface-bound experiments confirm model predictions

7.3.5 Future Work: Catalysis surface factor $\kappa_{cat}^{(teth2teth)}$ versus l_C and/or versus r_{ligand}

Experimental measurements of effective concentration do not distinguish between the effective concentration and the catalytic rate. Since these values are lumped into a single number, it is important to know how the surface impacts the apparent catalytic rate. This might be flat out incorrect. include calculation of how to find catalytic ratio? We will plot the ratio of apparent catalytic rates for various polymer lengths and ligand sizes. This will indicate if there is a relationship between the catalysis factor and the reaction parameters or if there is a single value describing the impact of the surface. The catalysis surface factor may then be used to determine how experiments on a surface differ from on a matrix and possibly give a conversion factor between the two.

7.4 Discussion?

8 Timeline

Table 5: Timeline of Coursework. Courses in italics do not count towards degree course requirements.

Fall 2014	**Winter 2015**	**Spring2015**	**Summer 2015**
Dev Bio 212 Math 227A Dev Bio 203A Physics 230A	Dev Bio 212 Math 227B Dev Bio 232	Dev Bio 212 Math 227C Dev Bio 203C M&MG 250	Research
Fall 2015	**Winter 2016**	**Spring2016**	**Summer 2016**
Math 205A *Math 220A* *Math 230A* *Math 298A*	*Math 205B* *Math 220B* *Math 230B* *Math 298B*	*Math 205C* *Math 220C* *Math 230C* *Math 298C*	Research
Fall 2016	**Winter 2017**	**Spring2017**	**Summer 2017**
Dev Bio 212 Math 225A MMB 204	Dev Bio 212 Math 290B	Dev Bio 212 CS 284C Physics 230B	Research Preliminary Exam Advancement
Fall 2017	**Winter 2018**	**Spring2018**	**Summer 2018**
Dev Bio 212	Dev Bio 212	Dev Bio 212	Research
Fall 2018	**Winter 2019**	**Spring2019**	**Summer 2019**
Dev Bio 212	Dev Bio 212	Dev Bio 212	Defense

9 Appendix

9.1 Model Modifications

9.1.1 Stiffening

To include this model assumption, we modify the simulation in the following way:

1. Include binary sequence to indicate which sites are "stiff", sequence specified by user.
2. Create distance metric to indicate how far on either side of the site the effect extends, specified by user.
3. Include function to read binary sequence and create list of "stiff" joints, based on distance metric.
4. Prohibit joints on list from being chosen when simulation chooses joint to move.

9.1.2 Electrostatics

To include this model assumption, we modify the simulation in the following way:

1. Create binary sequence to indicate which sites are "phosphorylated".

2. Create distance metric to indicate how far on either side of the site the effect extends, specified by user.
3. Create function to read binary sequence and create list of which sites to ignore energy, based on distance metric.
4. Calculate energy of system based on Lennard-Jones or Piecewise potential curve plus hard or soft wall constraints.

   Polymer energy is the sum of the energy for each joint in polymer.

   Energy of sites on phosphorylated list are not included in the sum.

   Depth, width and wall/center location are specified by user.
5. Accept conformations that produce a lower energy than the previous conformation.
6. Reject with normal probability any conformations with higher energy than previous conformation.
7. Rejection keeps polymer in previous conformation, but goes on to next time step. (Metropolis step.)

9.1.3 LJ Parameter Optimization

First potential:

Lennard-Jones potential, (V_{LJ}):

$$V_{LJ} = 4\epsilon \left(\left(\frac{l_D}{r - r_W} \right)^{12} - \left(\frac{l_D}{r - r_W} \right)^{6} \right)$$

where ϵ is the well depth, l_D is the Debye length, r_W is the potential wall location (vertical asymptote), and r is the location of the joint in question. For our purposes, r is the z-coordinate of the joint in question.

We want to look at what Debye length causes the electrostatic potential minima to shift from before $r = 0$ to after $r = 0$ (before, after membrane). We therefore solve the derivative of V_{LJ} (given below) for l_D when $r = 0$ and $dV_{LJ}/dr = 0$.

The derivative of V_{LJ}:

$$\frac{dV_{LJ}}{dr} = 4\epsilon \left(l_D^{12} \frac{-12}{(r - r_W)^{13}} + l_D^{6} \frac{6}{(r - r_W)^{7}} \right)$$

Input $dV_{LJ}/dr = 0$, $r = 0$:

$$
\begin{aligned}
0 &= 4\epsilon\left(l_D^{12}\frac{-12}{(-r_W)^{13}} + l_D^{6}\frac{6}{(-r_W)^{7}}\right) \\
0 &= \left(l_D^{12}\frac{-12}{(-r_W)^{13}} + l_D^{6}\frac{6}{(-r_W)^{7}}\right) \\
l_D^{12}\frac{12}{(-r_W)^{13}} &= l_D^{6}\frac{6}{(-r_W)^{7}} \\
l_D^{6}\frac{12}{(-r_W)^{13}} &= \frac{6}{(-r_W)^{7}} \\
12 * l_D^6 &= 6 * (-r_W)^6 \\
l_D^6 &= \frac{(-r_W)^6}{2} \\
l_D &= \sqrt[6]{\frac{1}{2}} * |r_W|
\end{aligned}
$$

Note that if the potential wall location is negative, then the Debye length may be either positive or negative. If the wall location is positive, then for our simulation, there is no Debye length that would cause the potential minimum to fall at $r = 0$.

Therefore, we can say that when $l_D < \sqrt[6]{\frac{1}{2}} * |r_W|$ then the polymer will be preferentially located 'under the membrane', or with $r < 0$ for most of its rod locations. Similarly, for $l_D > \sqrt[6]{\frac{1}{2}} * |r_W|$, the polymer will primarily be 'above' the membrane.

9.1.4 Simultaneous Binding

To include this model assumption, we modify the simulation in the following way:

1. Create "bound ligands", with location and radius specified by user. (Location specified by file or binary sequence.)
2. Bound ligands occlude polymer, rejecting conformations that intersect the ligand sphere.
3. Bound ligands occlude each other, rejecting conformations where two ligands intersect each other.
4. If there is a membrane, bound ligands may not fall below membrane, rejecting conformations where this occurs.
5. Rejecting conformations does not step time forward. Continue looking for conformations until one meets constraints.
6. Binding sites are considered 'occluded' if a bound ligand intersects the binding site sphere.
7. Binding sites located where bound ligand is located read as always occluded.

9.2 Assumptions

9.3 Gillespie Algorithm

References

[1] Alaji Bah et al. "Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch." In: *Nature* 519.7541 (2015), pp. 106–9. ISSN: 1476-4687. DOI: 10.1038/nature13999. URL: http://dx.doi.org/10.1038/nature13999%7B%5C%%7D5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2842604%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract%7B%5C%%7D5Cnhttp://www.nature.com/doifinder/10.1038/nature13999%7B%5C%%7D5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/20298971%7B%5C%%7D5Cnhttp://w.

[2] E K Barber et al. "The CD4 and CD8 antigens are coupled to a protein-tyrosine kinase (p56lck) that phosphorylates the CD3 complex." In: *Proceedings of the National Academy of Sciences* 86.9 (1989), pp. 3277–3281. ISSN: 0027-8424. DOI: 10.1073/pnas.86.9.3277. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.86.9.3277.

[3] Derek Bryant, Lara Clemens, and Jun Allard. "Computational simulation of formin-mediated actin polymerization predicts homologue-dependent mechanosensitivity". In: *Cytoskeleton* 74.1 (2017), pp. 29–39. ISSN: 19493592. DOI: 10.1002/cm.21344.

[4] A Keith Dunker et al. "The unfoldomics decade: an update on intrinsically disordered proteins." In: *BMC genomics* 9 Suppl 2.Suppl 2 (2008), S1. ISSN: 1471-2164. DOI: 10.1186/1471-2164-9-S2-S1. URL: http://www.biomedcentral.com/1471-2164/9/S2/S1.

[5] Hannes Fischer, Igor Polikarpov, and Aldo F. Craievich. "Average protein density is a molecular-weight-dependent function". In: *Protein Science* 13.10 (2004), pp. 2825–2828. ISSN: 09618368. DOI: 10.1110/ps.04688204. URL: http://doi.wiley.com/10.1110/ps.04688204.

[6] Yanan He et al. "Phosphorylation-induced conformational ensemble switching in an intrinsically disordered cancer/testis antigen". In: *Journal of Biological Chemistry* 290.41 (2015), pp. 25090–25102. ISSN: 1083351X. DOI: 10.1074/jbc.M115.658583.

[7] Cesar A Lopez et al. "Membrane-Mediated Regulation of the Intrinsically Disordered CD3ϵ Cytoplasmic Tail of the TCR". In: *Biophysical Journal* 108.May (2015), pp. 2481–2491. ISSN: 00063495. DOI: 10.1016/j.bpj.2015.03.059.

[8] Himadri Mukhopadhyay et al. "Multisite Phosphorylation Modulates the T Cell Receptor ??-Chain Potency but not the Switchlike Response". In: *Biophysical Journal* 110.8 (2016), pp. 1896–1906. ISSN: 15420086. DOI: 10.1016/j.bpj.2016.03.024. URL: http://dx.doi.org/10.1016/j.bpj.2016.03.024.

[9] Daniel Reeves, Keith Cheveralls, and Jane Kondev. "Regulation of biochemical reaction rates by flexible tethers". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 84.2 (2011), pp. 1–12. ISSN: 15393755. DOI: 10.1103/PhysRevE.84.021914.

[10] P Tompa and K-H Han. "Intrinsically disordered proteins". In: *Physics Today* 65 (2012), pp. 64–65. ISSN: 0968-0004. DOI: 10.1016/j.tibs.2012.08.004.

[11] David Van Valen, Mikko Haataja, and Rob Phillips. "Biochemistry on a leash: The roles of tether length and geometry in signal integration proteins". In: *Biophysical Journal* 96.4 (2009), pp. 1275–1292. ISSN: 00063495. DOI: 10.1016/j.bpj.2008.10.052. URL: http://dx.doi.org/10.1016/j.bpj.2008.10.052.

[12] Wei Wu et al. "Lipid in T-cell receptor transmembrane signaling". In: *Progress in Biophysics and Molecular Biology* 118.3 (2015), pp. 130–138. ISSN: 0079-6107. DOI: 10.1016/j.pbiomolbio.2015.04.004. URL: http://dx.doi.org/10.1016/j.pbiomolbio.2015.04.004.