

Mathematical Models for Multiubiquitination and Polyubiquitination Processes

Guy Needham

Word count: 5332

Abstract

Ubiquitination is the addition of ubiquitin monomers to a substrate protein. Ubiquitination is important for regulation of many cellular functions and is of clinical significance. Ubiquitination can result in multiubiquitination, the attachment of single ubiquitin monomers to multiple sites of the protein, or polyubiquitination, the buildup of long chains of ubiquitin. Polyubiquitination has two modes: distributive polyubiquitination and processive polyubiquitination. Polyubiquitination occurs on a sliding scale of processivity between the two mechanistic extremes of distributive or processive polyubiquitination. Here, we present mathematical models of multiple modes of ubiquitination that were solved using computational methods. Significantly, distributive and processive polyubiquitination were clearly distinct processes. A unified model of polyubiquitination with distributive and processive elements is presented.

Contents

1	Introduction	4
2	Methods and Materials	7
3	Results	8
3.1	Modelling Multiubiquitination	8
3.2	Polyubiquitination	9
3.2.1	Distributive Polyubiquitination	9
3.2.2	Processive Polyubiquitination	9
3.2.3	Comparison to Experimental Results	11
3.2.4	Unified Model of Polyubiquitination	13
3.3	Predicted Western Blot	13
3.4	Sampling the Polyubiquitination Models	16
4	Discussion	19
4.1	Multiubiquitination	19
4.2	Polyubiquitination	20
5	Conclusion	23
6	Appendices	23
1	Multiubiquitination Model	23
2	Distributive Polyubiquitination Model	24
3	Processive Polyubiquitination	25
4	Unified Model of Polyubiquitination	26

1 Introduction

Ubiquitin is a very stable 76 amino acid protein with a very strongly conserved structure through evolution¹ which is present in all eukaryotic cells.² When all ubiquitin-coding genes were deleted in *Saccharomyces cerevisiae*, the result was a complete loss of viability.³ This demonstrates how essential ubiquitin is for eukaryotic life. Ubiquitin has four exposed lysine residues, allowing it to covalently bond with other proteins. Ubiquitination is the covalent attachment of ubiquitin monomers, via the protruding C-terminus, to residues of a substrate protein by a ligase. Polyubiquitin chains form through the sequential addition of ubiquitin monomers to ubiquitins that are already attached to a protein substrate.

Polyubiquitination is a post-translational modification which is vital for regulating protein function and degradation and plays a role in many cellular processes such as regulation of cell-cycle progression, transcription activity, signal transduction, antigen processing, apoptosis, chromatin structure and internalisation of receptors leading to their degradation.⁴ Processes such as the cell cycle are regulated by a series of protein degradation events.^{5,6} Such is the importance of ubiquitination that it has been implicated in multiple diseases. Ubiquitination is important in cancer as it can regulate the cell cycle, the balance between proliferation and apoptosis, and autophagy.⁷ Ubiquitination has also been linked to neurodegenerative diseases as it effects synaptic plasticity and the buildup of protein aggregates.⁸

Ubiquitination is important in plant molecular biology. Auxin is a plant hormone which is involved in most aspects of plant growth and development.⁹ Ubiquitination of transcriptional repressors in response to auxin signalling is a vital mode of action in plant development.¹⁰ Jasmonate is important in regulation of plant development and response to environmental pressures.¹¹ A jasmonate receptor coronatine-insensitive1 is important in the degradation of some transcriptional repressors in response to jasmonate¹² via the ubiquitin pathway.

Ubiquitination involves three enzymes. The ubiquitin-activating enzyme (E1) activates ubiquitin in an ATP-dependent manner. The ubiquitin-conjugating enzyme (E2) accepts the activated ubiquitin from E1. Activated ubiquitin is attached to the substrate in an E3 ligase-mediated fashion.¹³

Ubiquitination can occur in multiple ways. For example, a single lysine of a protein can have a chain of ubiquitin monomers attached to it in polyubiquitination. Attachment of ubiquitin

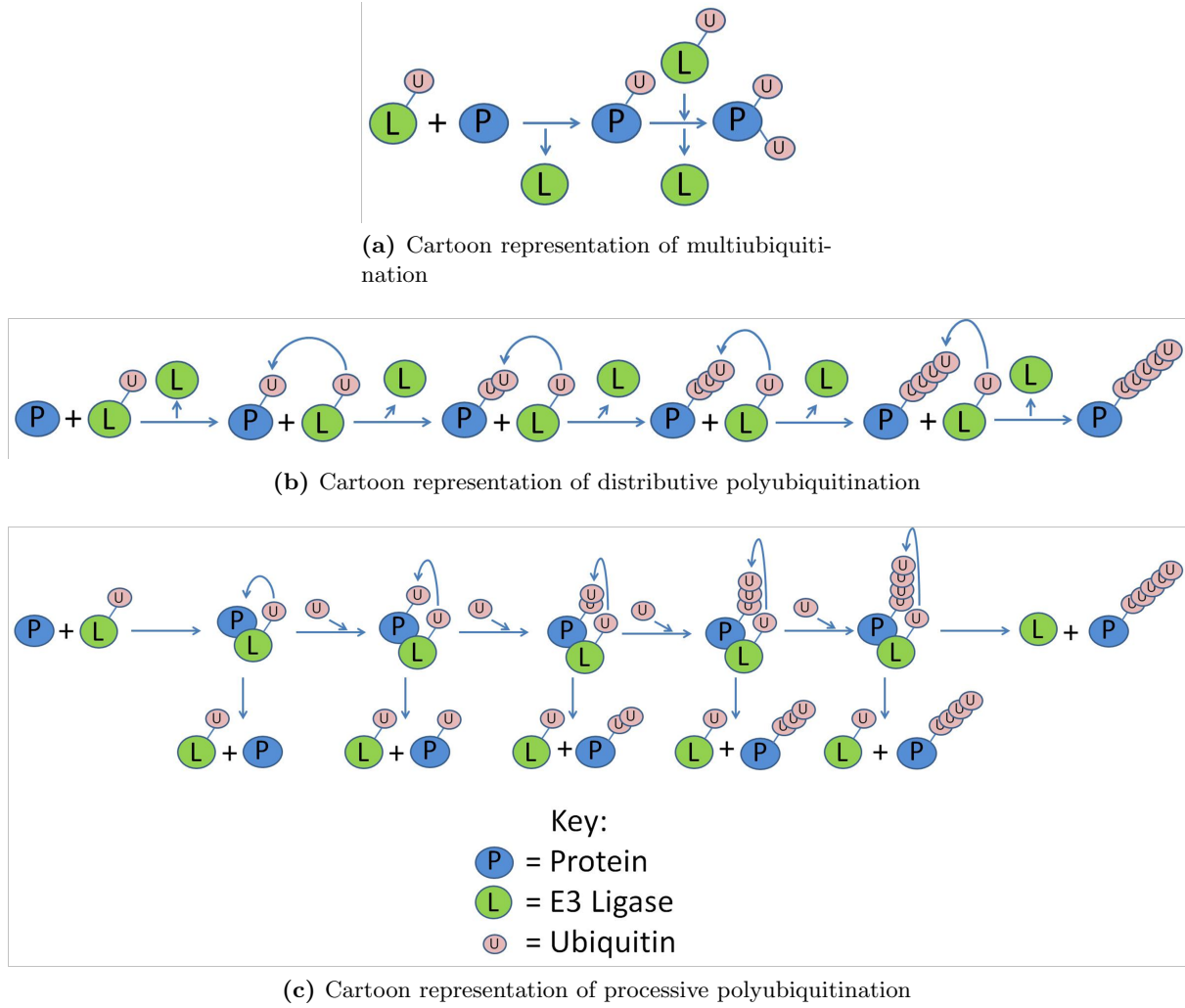


Figure 1: Cartoon representations of ubiquitination. A) is multiubiquitination, in which the protein is ubiquitinated at two different sites. B) is distributive polyubiquitination, in which the protein P is polyubiquitinated in a process involving multiple association and dissociation steps between the protein and the ligase L. C) is processive polyubiquitination, in which the protein P is polyubiquitinated in a process involving a single association event involving the protein and the ligase L.

chains is vital for proteolysis as the ubiquitinated protein is targeted for degradation by the proteasome.^{3,14,15} Polyubiquitination can occur in a processive or distributive manner.¹⁶ Processive polyubiquitination involves the serial addition of ubiquitin to a growing chain while the ligase remains bound to the substrate protein. Conversely during distributive polyubiquitination the ligase releases the substrate after the addition of each ubiquitin monomer. Multiubiquitination is the addition of single ubiquitin monomers to multiple sites of the same protein. This often occurs on receptor proteins and leads to their internalisation and lysosomal degradation.^{17–20} These processes are illustrated in figure 1.

Recent evidence has suggested that chains of ubiquitin are built up in a stepwise fashion and

are not passed to the substrate protein fully formed as was previously believed.²¹ This was discovered by Pierce et al²² following the development of a new experimental technique which allows the ubiquitination process to be followed on a millisecond timescale which was not previously possible. This technique allowed intermediates during the buildup of polyubiquitin chains to be captured and analysed, showing that the chains are built up in a stepwise fashion and not transferred whole to the substrate protein. This development will allow ubiquitination to be studied in more depth experimentally, which was previously challenging as ubiquitination occurs over a short timescale.

Polyubiquitination determines the substrate ordering of anaphase-promoting complex (APC) substrates. Substrates are polyubiquitinated in a well defined order by the APC. The complex sequence of protein ubiquitinations is termed substrate ordering and is required for proper cellular function. In 2006, Rape et al¹⁶ showed that the substrate ordering is determined by the processivity of the substrate: proteins are ubiquitinated according to where they sit on a continuum between processive and distributive polyubiquitination. Their processivity has implications for how the proteins compete for the ligase. Processively polyubiquitinated proteins bind the ligase once to achieve polyubiquitination, whereas distributively polyubiquitinated proteins must rebind the ligase multiple times to become polyubiquitinated. This means that processively polyubiquitinated proteins will outcompete distributively polyubiquitinated proteins and so are ubiquitinated more rapidly. Examples of processive proteins are securin and geminin. They are degraded in the presence of low concentrations of the ligase UbcH10 because they only need to bind the ligase once. Cyclin A is not degraded in the presence of low concentrations of UbcH10²³ as it is outcompeted for binding by the more processive proteins securin and geminin. Rapid degradation is due to more processive ubiquitination. The position of the protein on the processivity scale determines the speed at which substrate is degraded and defines the substrate ordering.

The pace of the ubiquitination process, along with its complexity at the molecular level, makes it challenging to study *in vitro* or *in vivo*. Coupled with its emerging clinical significance, this makes the process a suitable candidate for investigation by mathematical modelling. The development of a simplified model of the process will improve our understanding of this complex process and may allow the mode of ubiquitination to be determined from experimental results. Modelling the process will also improve our understanding of which rate constants

are important. Rate constants control the rates of reaction and one step may control the whole process. This is named the rate limiting step, and is often of therapeutic importance: by targeting the rate limiting step a process can be sped up or inhibited, with therapeutic implications.

This study will aim to create a novel series of models of ubiquitination. Multiubiquitination and polyubiquitination will be modelled by creating a series of ordinary differential equations (ODEs) describing the processes. These ODEs will be solved computationally, and the results will be plotted. The plots will be analysed and compared to find the implications for our understanding of ubiquitination.

2 Methods and Materials

The method used to generate the models is best described using the example process of an enzyme (E) binding to its substrate (S) to form the enzyme-substrate complex (ES) which reacts to form the product (P) and E. This is shown below:



This process can be described as a series of ODEs, which describe the rate of change of the concentration of each species as a function of the concentrations of all species. The ODEs for this process are shown below:

$$\frac{\delta E}{\delta t} = k_{\text{off}}[ES] + k_{\text{cat}}[ES] - k_{\text{on}}[E][S] \quad (2)$$

$$\frac{\delta S}{\delta t} = k_{\text{off}}[ES] - k_{\text{on}}[E][S] \quad (3)$$

$$\frac{\delta ES}{\delta t} = k_{\text{on}}[E][S] - k_{\text{off}}[ES] - k_{\text{cat}}[ES] \quad (4)$$

$$\frac{\delta P}{\delta t} = k_{\text{cat}}[ES] \quad (5)$$

These ODEs can be numerically solved to give information on the process. This approach was followed to create models of the more complex ubiquitination pathways investigated.

The Python language²⁴ and the SciPy toolkit²⁵ was used to numerically solve the models using

the odeint module. 2D plots were created using the matplotlib module of Python (see figure 2). MATLAB²⁶ was used to create 3D plots of data generated in Python (see figure 7).

3 Results

3.1 Modelling Multiubiquitination

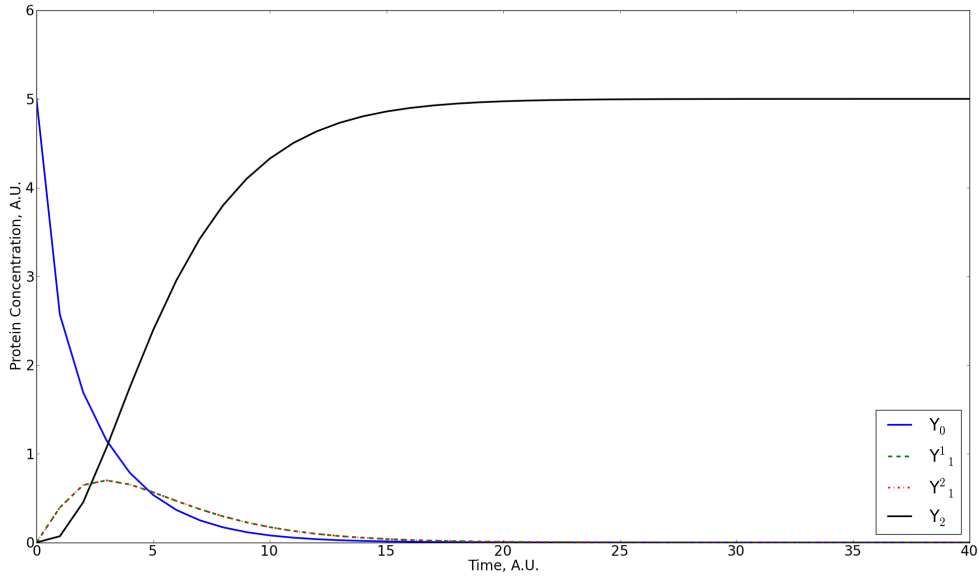


Figure 2: Numerical solutions to our model of multiubiquitination for species Y_0 , Y_1^1 , Y_1^2 and Y_2 , defined in appendix 1. The model (see appendix 1) consist of a series of ordinary differential equations which were numerically solved. Here, all rate constants were set at 1 except k_2 and k_6 which were both 0.5.

Multiubiquitination involves the covalent attachment of single ubiquitin monomers to two different sites on a target protein by an E3 ligase. The cartoon in figure 1a demonstrates an example of this. The corresponding mathematical model is given in appendix 1. This model was numerically solved. The resulting series of concentrations were plotted against time. This is shown in figure 2.

Over time, the concentration of the initial substrate decreases rapidly to zero. The intermediate monoubiquitinated proteins initially increase however their concentration slowly decreases to zero as these species are used up in the formation of the multiubiquitinated final product. Both monoubiquitinated proteins have an identical profile here, as the rate constants k_2 and

k_6 which control their formation are equal. If the rate constants were altered, the profiles of the intermediates would change however they would both decrease to zero by the end of the simulation. The final product increases in concentration over time until its concentration is equal to the concentration of protein added to the system. No other species are present at the end of the model.

3.2 Polyubiquitination

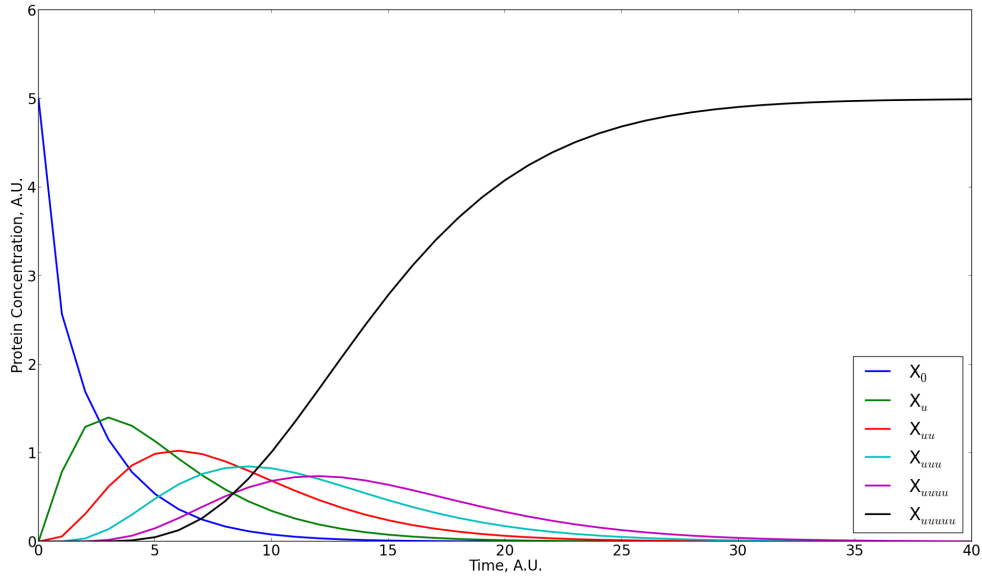
Polyubiquitination is the formation of a chain of covalently bonded ubiquitin monomers on one lysine residue of the substrate protein. We considered two mechanisms for polyubiquitination: distributive polyubiquitination (see figure 1b) and processive polyubiquitination (see figure 1c). In distributive polyubiquitination, the addition of a ubiquitin monomer to the growing chain occurs following the association of the protein with a ligase. This complex dissociates after the monomer is added to the chain. This contrasts with processive polyubiquitination in which the protein stays associated with the same E3 ligase enzyme throughout the formation of the ubiquitin chain. The ligase and substrate protein may dissociate during the process however they do not reassociate. We determined to create abstract models of these processes.

3.2.1 Distributive Polyubiquitination

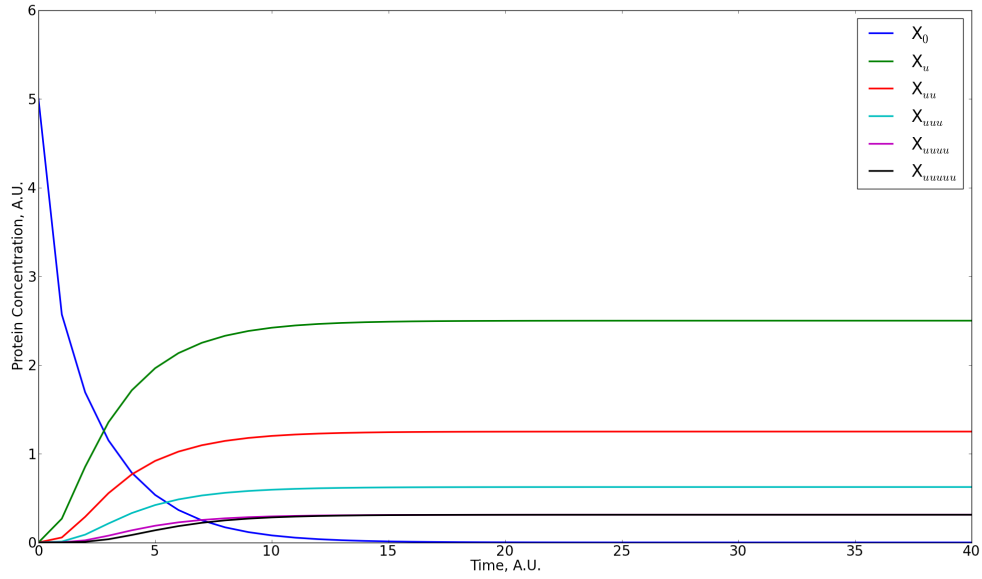
The model of distributive polyubiquitination is given in appendix 2. The model was numerically solved. The results were plotted against time and can be seen in figure 3a. Over time, the concentration of the initial substrate decreased rapidly to zero. Concentration of the intermediate species sequentially increased and decreased. The concentration of the polyubiquitinated protein X_{uuuuu} increased in a sigmoidal fashion until it reached 5 A.U., the starting concentration of unubiquitinated protein X_0 . No other species were present at the end of the simulation. In order to compare the behaviour of distributive and processive polyubiquitination, processive polyubiquitination was modelled.

3.2.2 Processive Polyubiquitination

The model of processive polyubiquitination is given in appendix 3. It has a constant k_{off} rate. The model was numerically solved. The results were plotted against time and are shown



(a) Numerical evaluation of a model of distributive polyubiquitination



(b) Numerical evaluation of a model of processive polyubiquitination

Figure 3: Numerical solutions to models of processive and distributive polyubiquitination for species X_0 , X_u , X_{uu} , X_{uuu} , X_{uuuu} and X_{uuuuu} defined in appendix 2. Models shown in appendices 2 and 3, respectively, were numerically solved. The results of the models were plotted against time. Here, all the rate constants were set at 1.

in figure 3b. The concentration of the initial substrate X_0 decreased rapidly to zero. The concentration of the other species increased and plateaued at sequentially lower concentrations. The end levels of each species can be altered by changing the rates of reaction (data not shown).

The processive model reached thermodynamic equilibrium after between 10 and 15 A.U., which was much faster than the distributive model which reached equilibrium after between 35 and 40 A.U.

3.2.3 Comparison to Experimental Results

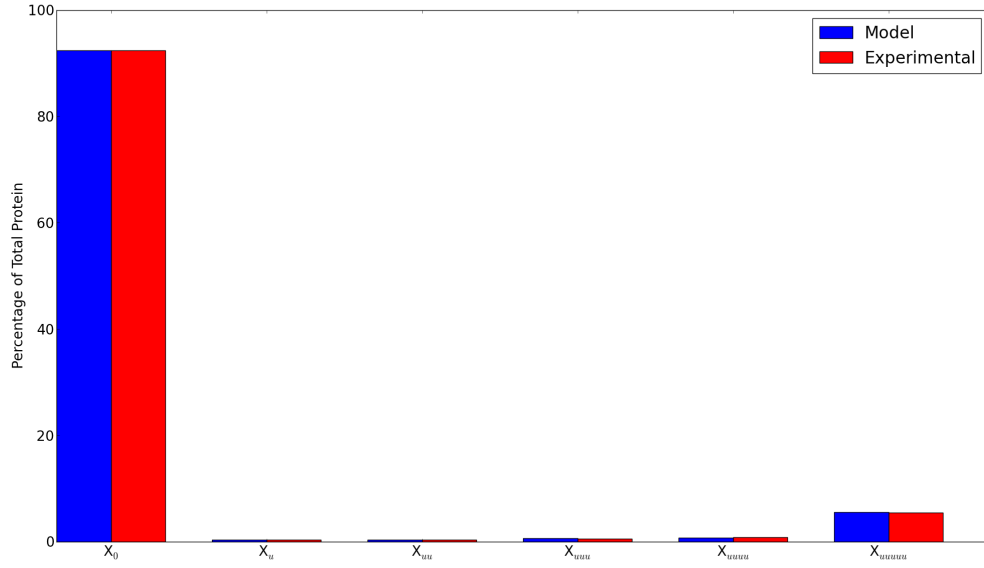
Recently, Pierce et al²² developed a new technique to allow the determination of the rate constants and final levels of the species involved in an *in vitro* polyubiquitination model. In this paper, the polyubiquitination of β -Cat and CYCE were investigated and were determined to display processive polyubiquitination.

In order to validate the model we developed, the rate constants from the paper were used in our processive polyubiquitination model. In these simulations, the $k_{off\ 0}$ rate was set at 0 and a $X_{0.L}$ concentration of 5 A.U. was used. These changes were made to reflect the *in vitro* model used in the paper.

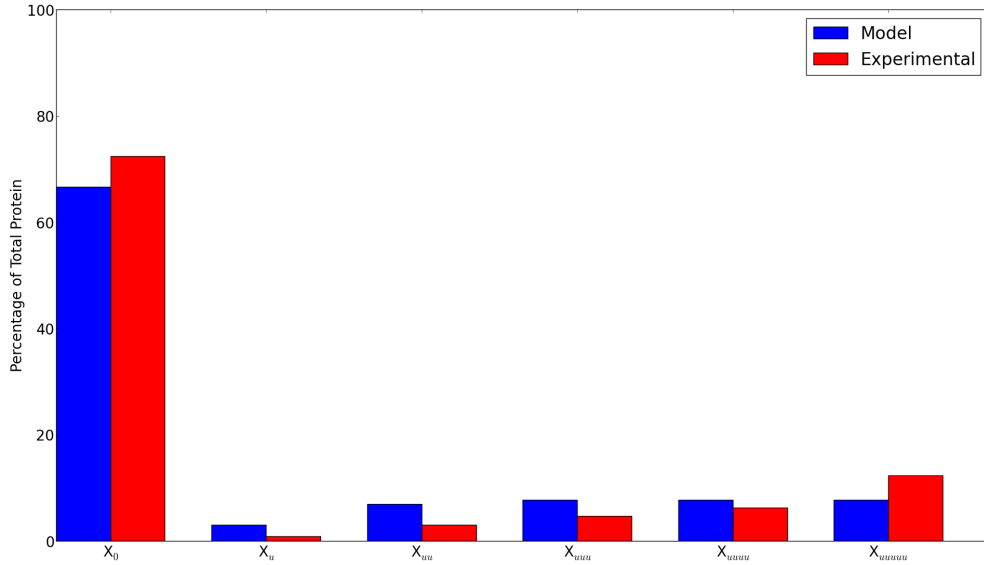
The modified models of processive polyubiquitination were numerically solved as before. The concentration of each species was plotted as a percentage of all species. Each species was plotted next to the experimentally determined value from the paper (figure 4). As the experimental system involved species with ubiquitin chains up to nine ubiquitins long, whereas our model only ran to five ubiquitins long, the experimental data at the point X_{uuuuu} is the sum of the protein from X_{uuuuu} to $X_{uuuuuuuu}$.

Pearson correlation coefficients²⁷ were calculated for the data. For both proteins there was a strong correlation between the model and experimental data. The β -cat data had a P-value of 4.4195e-12 and the CYCE data had a P-value of 4.4860e-5. This shows that for both β -Cat and CYCE, there was no significant difference between our *in silico* model predictions and the experimental *in vitro* data from Pierce et al.²² This showed that this model of processive polyubiquitination can return the distribution of protein species at the end of a process, given the rate constants for that process.

This experiment showed that our model can produce a varied output. The general trend of our predictions matched the experimental data. The results were further fine tuned by the parameters which were experimentally determined. The disparity in CYCE results suggests that these parameters may not be correct.



(a) Comparison of a numerically solved model of processive polyubiquitination with experimentally derived values of β -Cat polyubiquitination



(b) Comparison of a numerically solved model of processive polyubiquitination with experimentally derived values of CYCE polyubiquitination

Figure 4: Comparison of experimentally determined results with predicted results from numerical evaluation of a processive polyubiquitination model. Plots show experimentally determined concentrations of proteins in different ubiquitination states at the end of the given reaction period. The model of processive polyubiquitination was adapted to reflect the *in vitro* experimental work carried out by Pierce et al.²² The experimentally determined rate constants were used in the models, which were evaluated. The final concentrations of each species were plotted and compared to the experimentally determined final concentration. Note that the model's final species is X_{uuuuu} whereas the experimental system ran to $X_{uuuuuuuuuu}$ and so the X_{uuuuu} term for the experimental data is the sum of the concentrations of X_{uuuuu} onwards.

3.2.4 Unified Model of Polyubiquitination

Within polyubiquitination, it is known that there is a sliding scale of processivity between entirely processive and entirely distributive. This creates a mechanism for substrate ordering in late mitosis and G1, which is vital for regulating these processes.¹⁶

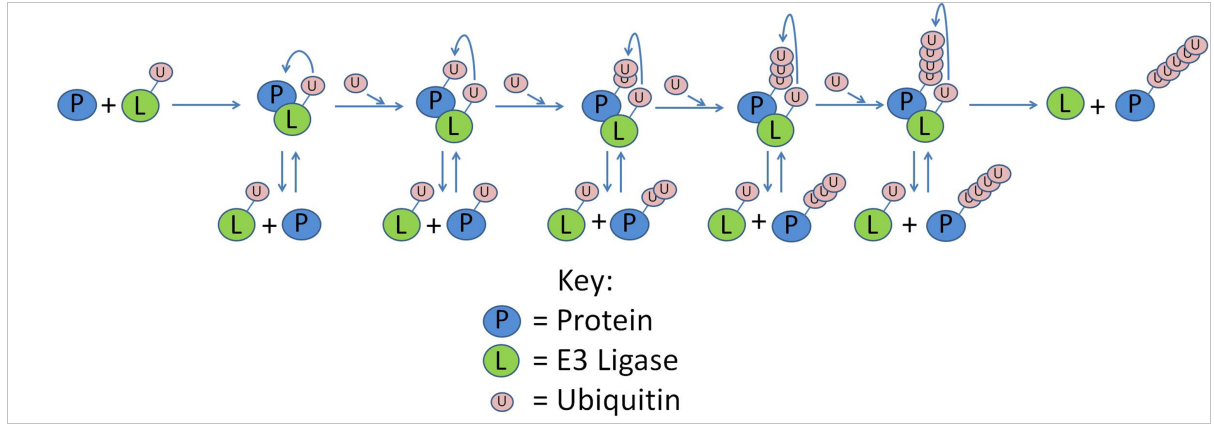
We modelled this sliding scale by altering the processive model of polyubiquitination. Following the dissociation of the ligase from the substrate, the species here can reassociate with the ligase and re-enter the pathway. In this scenario, the processive model has distributive characteristics and so is termed the unified model of polyubiquitination. This model is illustrated in figure 5a. This was modelled (see appendix 4) and numerically solved. The results were plotted and can be seen in figure 5b.

The results of this simulation closely resemble those of the distributive polyubiquitination model (see figure 3a). The concentration of the initial substrate X_0 decreased rapidly to zero. The intermediate species sequentially increased before decreasing slowly to zero. Unlike the processive polyubiquitination model (see figure 3b) intermediate species do not remain at the end of the simulation. The concentration of X_{uuuuu} increased in a sigmoidal fashion. At the end of the simulation, only X_{uuuuu} was present in the system.

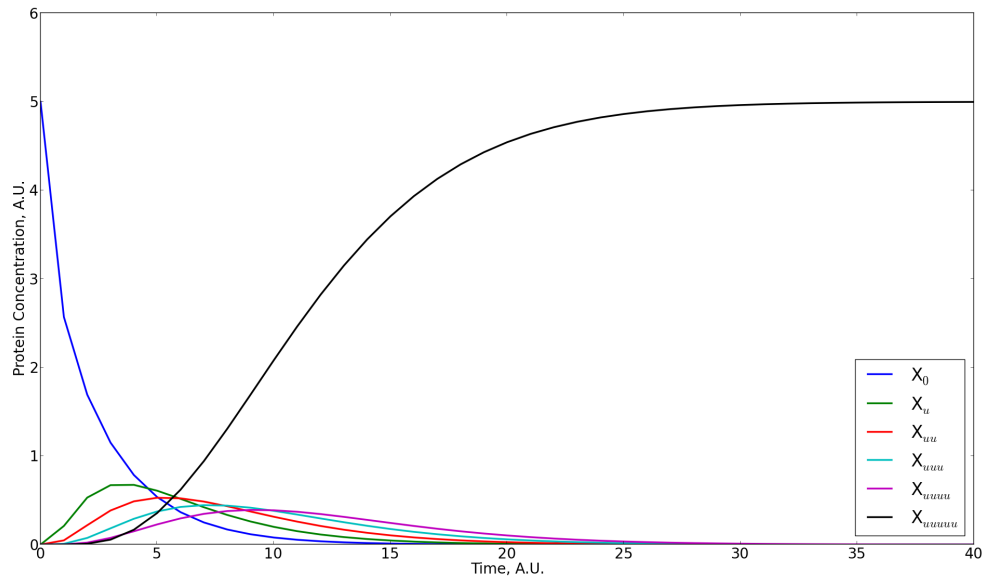
3.3 Predicted Western Blot

One use of the models we have created is to determine from the final result of an experiment which form of ubiquitination has occurred. To determine the mode of ubiquitination, the protein levels in a sample must be determined. One way to determine protein levels in a sample is to perform a Western blot, where the proteins in the mixture are separated by size on a gel before being transferred to a membrane. Antibodies are incubated with the membrane and bind to the protein they were raised against. The visualisation of the binding of these antibodies allows the species present in a sample to be determined, along with their concentration relative to one another. A Western blot could indicate which form of ubiquitination has occurred.

Figure 6 shows a predicted Western blot we have made for an experiment in which multi-ubiquitination, processive polyubiquitination, distributive polyubiquitination and unified polyubiquitination have all taken place, along with a ubiquitination negative control sample. We have predicted the appearance of this Western blot based on numerical solutions presented in



(a) Cartoon of the unified model of polyubiquitination



(b) Numerical evaluation of a unified model of polyubiquitination

Figure 5: Cartoon and results of numerical evaluation of a unified model of polyubiquitination. Part a) shows a cartoon representation of the process. It demonstrates the option for intermediates to reassociate with the ligase to continue the process, which differentiates it from processive polyubiquitination. Part b) shows the results of the numerical evaluation of the model plotted against time. All rate constants were set at 1.

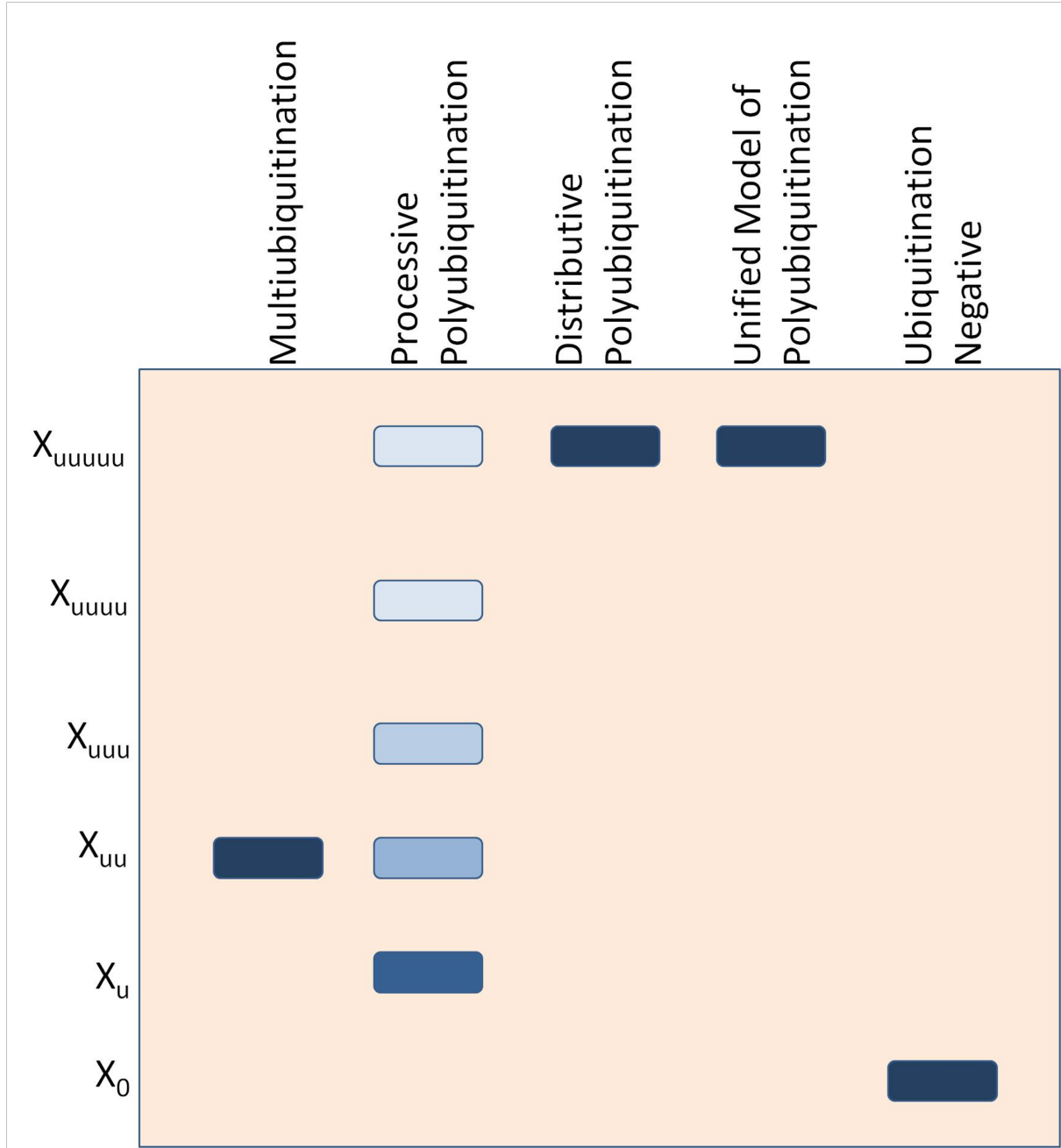


Figure 6: Simulated Western blot of the model outputs, where all rate constants are equal to 1. The outputs of the models, shown in figures 2, 3, and 5, were used to predict the appearance of a Western blot containing samples from systems in which each form of ubiquitination investigated had taken place. A ubiquitination-negative sample was included as a negative control.

sections 3.1, 3.2.1, 3.2.2 and 3.2.4.

Multiubiquitination, distributive polyubiquitination and unified polyubiquitination all lead to only one dark band being predicted as only one species is present. Processive polyubiquitination leads to a ladder of bands being predicted, with the higher molecular weight species giving progressively fainter bands. This is due to the sequentially lower concentrations of species as

they progress through the simulation. This shows that while distributive polyubiquitination and the unified model of polyubiquitination cannot be distinguished from this Western blot, processive polyubiquitination, distributive polyubiquitination and multiubiquitination can be distinguished. However, we would be interested to know if this is true only for the parameters used herein.

3.4 Sampling the Polyubiquitination Models

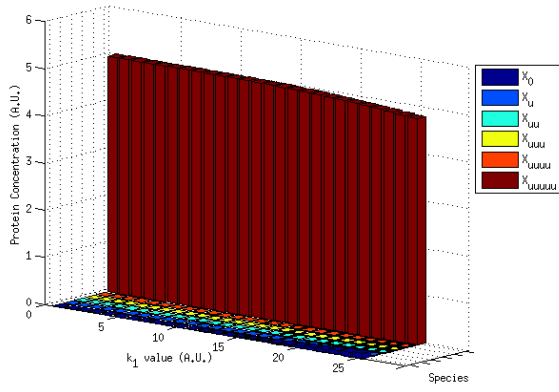
To investigate how the end state of our models of protein polyubiquitination change with rate constant, and to further investigate if the mode of polyubiquitination can be determined from observing the end state of the system, we sampled the dissociation constants of the three polyubiquitination models. The rates chosen for sampling were k_1 from the distributive model (see appendix 2), $k_{off\ 0}$ for the processive model (see appendix 3) and $k_{off\ 0}$ for the unified model (see appendix 4). These rates were chosen as we were attempting to replicate the *in vivo* conditions. Data from Pierce et al²² shows that a high proportion of the substrate is unubiquitinated suggesting a high dissociation rate which regulates the pathway. Therefore the dissociation rates were of interest.

The sampling was implemented by solving the models with a series of values for the chosen rate. The ranges chosen were between 0 and 25 in steps of 1, and between 0 and 1000 in steps of 20. The models were solved, and the data were plotted. The results can be seen in figure 7.

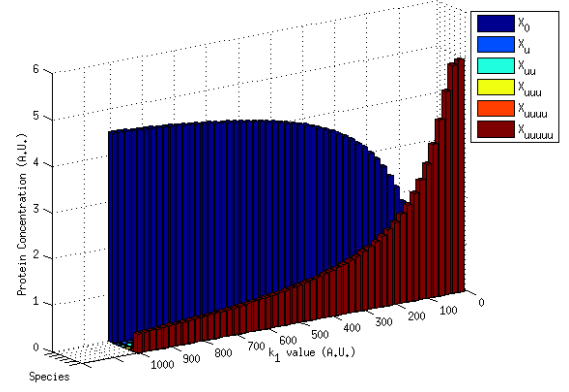
From these graphs, it is clear that the models behave in different ways upon sampling of the dissociation constants. The distributive model at lower k_1 values has high X_{uuuuu} concentration. As k_1 increases, the level of X_{uuuuu} decreases and X_0 increases. No other species is present at the end of the simulation. This shows that in this model, for k_1 values greater than 20, two species will be detectable if distributive polyubiquitination occurs.

In processive polyubiquitination, at low $k_{off\ 0}$ rates all substrate is ubiquitinated. Most substrate is monoubiquitinated and subsequent species have sequentially lower concentration. As $k_{off\ 0}$ increases, the level of X_0 increases gradually at the expense of the other species. Initially, the level of X_u decreases. Subsequent species decrease at higher $k_{off\ 0}$ values.

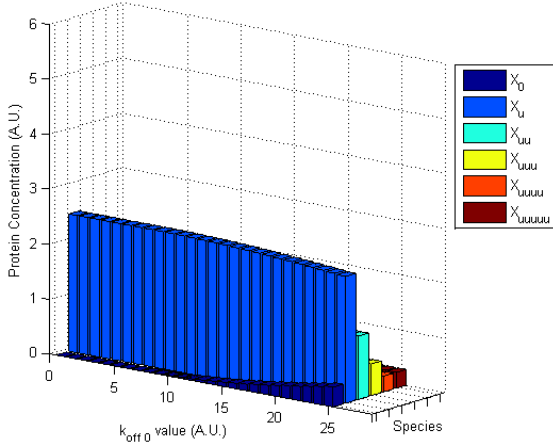
The unified model of polyubiquitination resembles distributive polyubiquitination. At low $k_{off\ 0}$ values, all protein is in the X_{uuuuu} state. As the $k_{off\ 0}$ value increases, the level of X_0 increases



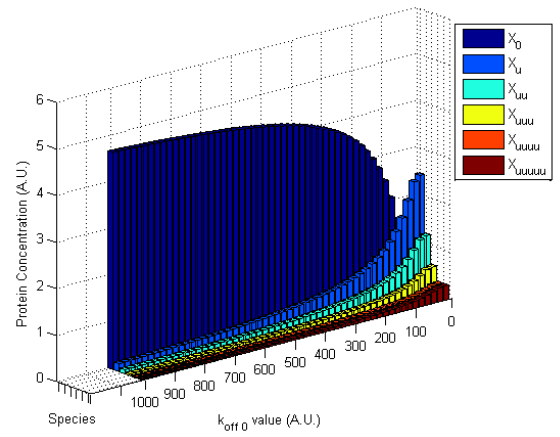
(a) Sampling k_1 between 0 and 25 in distributive polyubiquitination



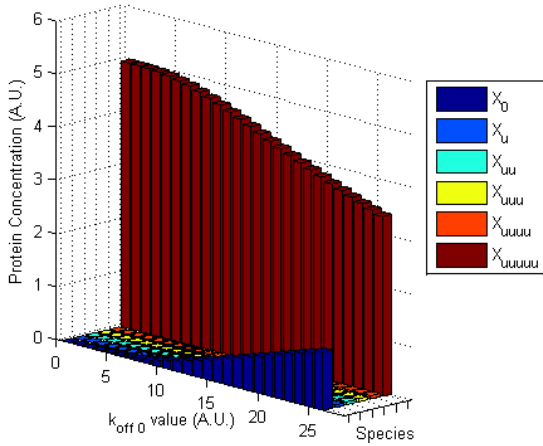
(b) Sampling k_1 between 0 and 1000 in distributive polyubiquitination



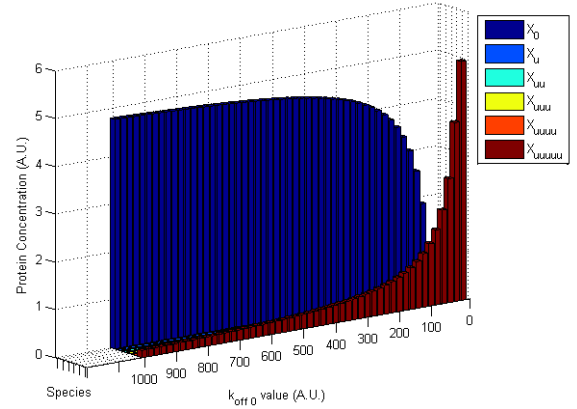
(c) Sampling k_{off0} between 0 and 25 in processive polyubiquitination



(d) Sampling k_{off0} between 0 and 1000 in processive polyubiquitination

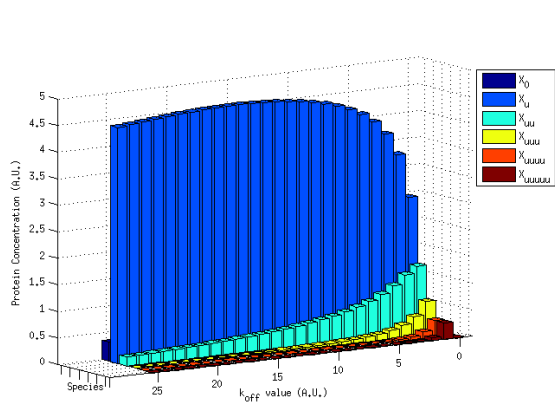


(e) Sampling k_{off0} between 0 and 25 in a unified model of polyubiquitination

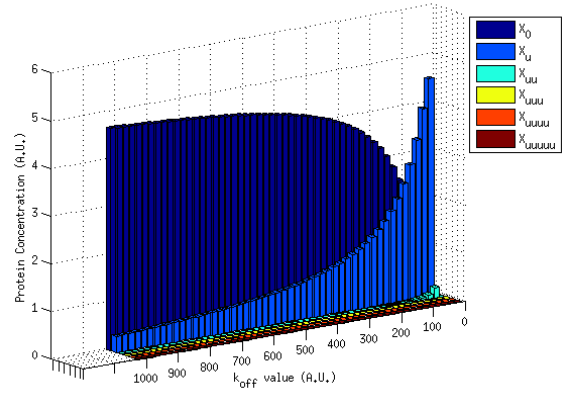


(f) Sampling k_{off0} between 0 and 1000 in a unified model of polyubiquitination

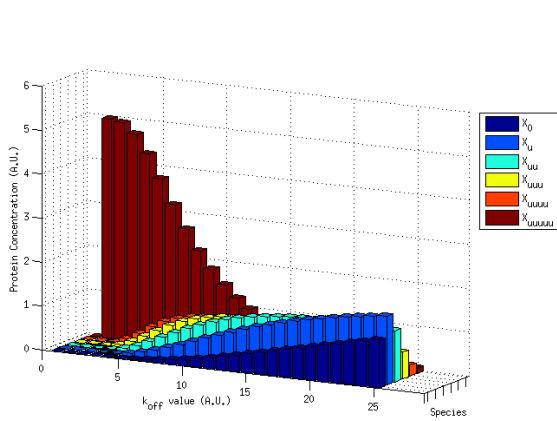
Figure 7: Sampling the dissociation rate constants for the initial encounter between the ligase and unubiquitinated protein in models of polyubiquitination. The three models of polyubiquitination, a)-b) distributive (see appendix 2), c)-d) processive (see appendix 3) and e)-f) unified (see appendix 4), were all evaluated with a range of dissociation rate constants. Each plot shows the distribution of ubiquitin at steady state given a different dissociation rate. Each model was evaluated twice with rate constants between 0 and 25 in steps of 1 and between 0 and 1000 in steps of 20. The models were evaluated, and the final concentrations of each species was plotted.



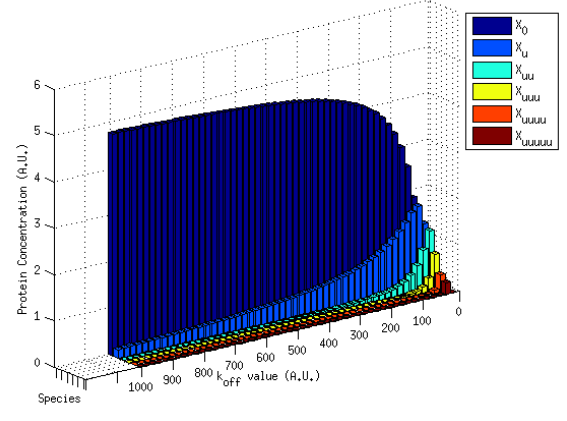
(a) Sampling k_1 between 0 and 25 in processive polyubiquitination



(b) Sampling k_1 between 0 and 1000 in processive polyubiquitination



(c) Sampling k_{off} between 0 and 25 in a unified model of polyubiquitination



(d) Sampling k_{off} between 0 and 1000 in a unified model of polyubiquitination

Figure 8: Sampling the dissociation rate constants in models of polyubiquitination. The processive model of polyubiquitination (see appendix 3) and the unified model of polyubiquitination (see appendix 4) were modified such that all k_{off} values were equal. The rate constants were then sampled. Each model was evaluated twice with rate constants between 0 and 25 in steps of 1 and between 0 and 1000 in steps of 20. The models were evaluated, and the final concentrations of each species was plotted.

and the level of X_{uuuuu} decreases. For most k_{off} values, two protein species will be detectable. Therefore while processive polyubiquitination should be easily distinguishable from distributive and unified mechanisms of polyubiquitination, the distributive and unified mechanisms may be indistinguishable.

We investigated how the processive and unified models varied if all the k_{off} rates were sampled simultaneously with the same value. This sampling was carried out as before, and the results can be seen in figure 8. In processive polyubiquitination, at low k_{off} rates all substrate is ubiquitinated. Most substrate is monoubiquitinated and subsequent species have sequentially

lower concentration. As k_{off} increases, the level of X_u initially increases at the expense of other ubiquitinated protein species. X_{uu} is the last ubiquitinated species other than X_u to remain. This species eventually reaches zero before k_{off} reaches 100. When k_{off} is greater than 15, the concentration of X_0 starts to increase at the expense of X_u . As k_{off} continues to increase, X_0 increases and X_u decreases symmetrically. This shows that the number of species present at the end of a process if processive polyubiquitination is occurring varies with k_{off} .

In the unified model of polyubiquitination, at low k_{off} values the only species present is X_{uuuuu} . As k_{off} increases, the ubiquitinated species sequentially increase in concentration and X_{uuuuu} decreases in concentration. All species other than X_0 eventually decrease; all except X_u decrease to zero. X_0 increases with k_{off} . Similarly to the processive model, the species present at the end of the simulation depend on the k_{off} value.

The unified model, with different combinations of parameters, results in a distribution of proteins similar to either the processive model or the distributive model. Figure 8d at high k_{off} values resembles figures 7d and 8b. The unified model results in figures 7e and 7f at low and high k_{off} values resemble the distributive models in figures 7a and 7b. This implies that our unified model can represent both the processive and distributive models, given appropriate parameters.

4 Discussion

4.1 Multiubiquitination

Multiubiquitination is the addition of single ubiquitin monomers to multiple sites on the substrate protein by an E3 ligase. Figure 2 shows a numerical solution for our model of this process, given in appendix 1. The simulation led to one species, multiubiquitinated protein. Figure 6 shows a simulated Western blot where we show the predicted appearance of reaction where multiubiquitination has occurred in the first lane.

The implications of the model are that if the system runs to thermodynamic equilibrium, only one species will be present. Therefore it can be determined if the process has reached its conclusion from analysing its makeup. This means that once a substrate protein enters this pathway, it is likely to be multiubiquitinated. Multiubiquitination is vital for the regulation of

transport of proteins during internalisation,^{17–20} and therefore its strict regulation is important: incorrectly internalised receptor proteins prevent proper cellular function.

The regulation of this pathway may occur via kinetic proofreading:²⁸ a high k_1 rate in the initial encounter will lead to very few proteins being multiubiquitinated as it will be unlikely for a stable protein-ligase complex to form. However, if a protein is targeted for ubiquitination it could be altered by another modification such as phosphorylation. Phosphorylation is a post-translational modification which can mark proteins for ubiquitination. Phosphorylation leading to ubiquitination can depend on the presence of PEST domains^{29–31} or can occur on single sites.^{32,33} This event could change the protein conformation, leading to altered binding to the ligase. This in turn could significantly lower the k_1 rate, allowing multiubiquitination to occur. The possibility of kinetic proofreading regulating this pathway should be investigated further by sampling the k_1 rate and investigating its impact on the entry of protein into the system.

A further avenue for future work is the interplay between k_2 and k_6 . If one site is preferentially ubiquitinated over the other, its rate constant will be greater, leading to greater flux through that side of the pathway. Similarly interesting the concept of feedback: flux through the pathway may alter rate constants in the pathway, altering the fate of subsequent substrate. For example, the buildup of Y_1^1 may increase k_3 . Another example is substrate level control of the rate of reaction. Mitogen-activated protein kinase phosphorylation has been shown to be regulated by substrate concentration.³⁴ This could be implemented into our model.

4.2 Polyubiquitination

We investigated polyubiquitination, the construction of chains of ubiquitin on a single site of the substrate protein. We modelled both distributive and processive polyubiquitination (figure 3). The distributive polyubiquitination model led to a single protein species forming, X_{uuuuu} . This contrasted with the processive polyubiquitination model which resulted in a series of proteins with sequentially lower concentration as their point in the process increased. When the dissociation rate constants were sampled, see figure 7, the results of the two models were distinct. The distributive and processive polyubiquitination models are distinct, both in terms of the composition of their results and the speed that the systems reached equilibrium. This is in keeping with the findings of Rape et al¹⁶ who showed that the processivity of a substrate determines

how rapidly it is polyubiquitinated. When our model of processive polyubiquitination was compared to experimental results gained by Pierce et al,²² (see figure 4) the polyubiquitination of β -Cat and CYCE matched the model. This finding adds confidence to our model of processive polyubiquitination as it can return experimental results given the rate parameters.

Polyubiquitination can act as a form of kinetic proofreading.²⁸ In kinetic proofreading, multiple steps of a biological process have stringent kinetic parameters, meaning that it is highly unlikely for an unsuitable substrate to reach the end of a process. The kinetics of each reaction add a layer of stringency to a pathway beyond the signaling which leads to its activation. Only proteins with ubiquitin chains of length four or longer are recognised by the proteosome³⁵ and degraded. If the polyubiquitination process does not reach this key chain length, the proteosome does not read the protein as polyubiquitinated. As polyubiquitination involves multiple reversible reactions, each step represents an opportunity for the cell to proofread the process to ensure that only appropriate substrate is polyubiquitinated.

Polyubiquitination *in vivo* occurs on a sliding scale of processivity.¹⁶ Some proteins are polyubiquitinated more processively and some more distributively. Each reaction in polyubiquitination can have processive or distributive elements. The processivity of individual reactions determines the processivity of the overall process. In order to investigate the sliding scale of processivity, a unified model of polyubiquitination is presented (see figure 5).

Sampling the models (see figures 7 and 8) demonstrated that the unified model can resemble the distributive or processive models, depending on the rate constants chosen. This shows that the unified model encompasses both extremes of the processivity scale, and therefore can represent proteins which are at intermediate points on the scale. In the sampling experiment, very high k_{off} values were chosen up to 1000 times the other rate constants. These high constants invariably resulted in most substrate protein being unubiquitinated. The biological relevance of these high constants is that if a protein is not marked for polyubiquitination, by phosphorylation or other mechanisms, the high k_{off} rate for that protein prevents its polyubiquitination and subsequent degradation.

The discrete models of polyubiquitination we have presented impose an artificial structure on the process, whereas the unified model of polyubiquitination is closer to the continuum of processivity observed *in vivo*. Therefore, it is more interesting for further study. One avenue for further investigation is which rates of reaction determine the processivity of the model: there may be

important rate limiting steps which cause the overall process to be more or less processive. Similarly, the rate constants which are typical of processive or distributive polyubiquitination could be determined. Certain combinations of rate constants may impose a distributive or processive structure on the process, and these would be interesting to investigate further. Subsequently, there may be concentrations of protein in the final state of the system which imply that the process was more or less processive. Characterising the makeup of the system which is a hallmark of a certain degree of processivity would be instrumental in allowing the determination of the mechanism of polyubiquitination of real experimental results. It would allow the comparison of the processivity of different substrate proteins. The processivity determines how they compete for the ligase. This determines how rapidly they are polyubiquitinated and is important for substrate ordering.

The processivity of a substrate determines how it competes for the ligase. In our model we have not investigated the importance of the enzyme concentration. Modelling the ligase enzyme explicitly will be important for investigating competition between substrates for ligase, and the impact of this on ubiquitination.

If the model is further investigated, it may be possible to compare experimental results to the model and determine the mechanism of polyubiquitination as well as predict rate constants for the process under investigation. This would be a useful tool for molecular biologists investigating polyubiquitination. If the hallmark protein concentration trajectories for various levels of processivity were determined, experimental results could be compared to the model and their processivity could be predicted. If typical rate constants for different levels of processivity had been found, the rate constants for an experimental process could be predicted, giving laboratory scientists useful knowledge for further investigation. A further possibility would be the creation of a library of protein concentrations linked to the rate constants which resulted in this distribution, as well as where on the processivity scale the process sits. Experimental results could be compared to this library to predict rate constants and processivity levels.

A useful tool for life scientist working in polyubiquitination would be a user interface for the unified model of polyubiquitination which would allow each rate constant to be altered easily. The model would be evaluated and the output would be the resulting protein concentrations, which could be represented in a simulated Western blot. This would allow the investigator to predict how different rate constants affect the outcome, which would be useful for forming and

testing hypotheses.

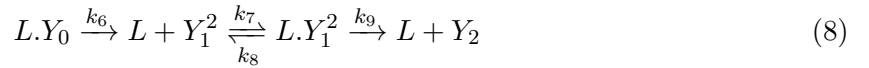
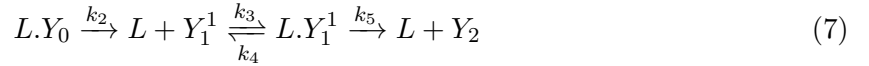
5 Conclusion

We have presented models for multiubiquitination, processive polyubiquitination, distributive polyubiquitination and a unified model for polyubiquitination. Polyubiquitination and multiubiquitination are distinct, as are processive and distributive polyubiquitination. These mechanisms of ubiquitination can be distinguished from one another. The unified model of polyubiquitination represents a more realistic continuum between the distributive and processive mechanisms of polyubiquitination, and should be investigated further.

6 Appendices

1 Multiubiquitination Model

Equations describing the model of multiubiquitination, where L is the E3 ligase and Y is the substrate protein. Y_0 is unubiquitinated protein, Y^1_1 is monoubiquitinated protein on site 1, Y^2_1 is monoubiquitinated protein on site 2, and Y_2 is multiubiquitinated protein. $L.Y$ is a complex between the E3 ligase and the substrate protein Y .



Ordinary differential equations:

$$\frac{\delta Y_0}{\delta t} = k_1 L.Y_0 - k_0 Y_0 L \quad (9)$$

$$\frac{\delta L.Y_0}{\delta t} = k_0 Y_0 L - k_1 L.Y_0 - k_2 L.Y_0 - k_6 L.Y_0 \quad (10)$$

$$\frac{\delta Y_1^1}{\delta t} = k_2 L.Y_0 + k_4 L.Y_1^1 - k_3 L.Y_1^1 \quad (11)$$

$$\frac{\delta L.Y_1^1}{\delta t} = k_3 L.Y_1^1 - k_4 L.Y_1^1 - k_5 L.Y_1^1 \quad (12)$$

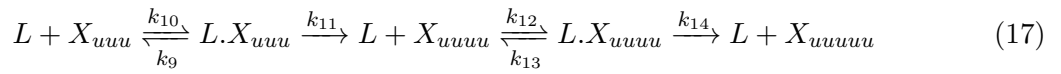
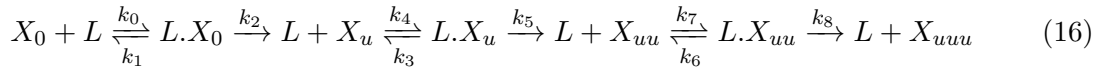
$$\frac{\delta Y_1^2}{\delta t} = k_6 L.Y_0 + k_8 L.Y_1^2 - k_7 L.Y_1^2 \quad (13)$$

$$\frac{\delta L.Y_1^2}{\delta t} = k_7 L.Y_1^2 - k_8 L.Y_1^2 - k_9 L.Y_1^2 \quad (14)$$

$$\frac{\delta Y_2}{\delta t} = k_5 L.Y_1^1 + k_9 L.Y_1^2 \quad (15)$$

2 Distributive Polyubiquitination Model

Equations describing the model of distributive polyubiquitination, where L is the E3 ligase and X is the substrate protein. X_0 is unubiquitinated protein. X_u is monoubiquitinated. The number of "u"s in the subscript denotes the number of ubiquitin monomers in the polyubiquitin chain. L.X is a complex between the E3 ligase L and the substrate protein X.



Ordinary differential equations:

$$\frac{\delta X_0}{\delta t} = k_1 L.X_0 - k_0 X_0 L \quad (18)$$

$$\frac{\delta L.X_0}{\delta t} = k_0 X_0 L - k_1 L.X_0 - k_2 L.X_0 \quad (19)$$

$$\frac{\delta X_u}{\delta t} = k_2 L.X_0 + k_3 L.X_u - k_4 X_u L \quad (20)$$

$$\frac{\delta L.X_u}{\delta t} = k_4 X_u L - k_3 L.X_u - k_5 L.X_u \quad (21)$$

$$\frac{\delta X_{uu}}{\delta t} = k_5 L.X_u + k_6 L.X_{uu} - k_7 X_{uu} L \quad (22)$$

$$\frac{\delta L.X_{uu}}{\delta t} = k_7 X_{uu} L - k_6 L.X_{uu} - k_8 L.X_{uu} \quad (23)$$

$$\frac{\delta X_{uuu}}{\delta t} = k_8 L.X_{uu} + k_9 L.X_{uuu} - k_{10} X_{uuu} L \quad (24)$$

$$\frac{\delta L.X_{uuu}}{\delta t} = k_{10} X_{uuu} L - k_9 L.X_{uuu} - k_{11} L.X_{uuu} \quad (25)$$

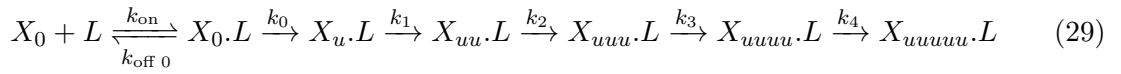
$$\frac{\delta X_{uuuu}}{\delta t} = k_{11} L.X_{uuu} + k_{13} L.X_{uuuu} - k_{12} X_{uuuu} L \quad (26)$$

$$\frac{\delta L.X_{uuuu}}{\delta t} = k_{12} X_{uuuu} L - k_{13} L.X_{uuuu} - k_{14} L.X_{uuuu} \quad (27)$$

$$\frac{\delta X_{uuuuu}}{\delta t} = k_{14} L.X_{uuuu} \quad (28)$$

3 Processive Polyubiquitination

Equations describing the model of processive polyubiquitination, where L is the E3 ligase and X is the substrate protein.



Ordinary differential equations:

$$\frac{\delta X_0}{\delta t} = k_{\text{off } 0} X_0.L - k_{\text{on}} X_0.L \quad (35)$$

$$\frac{\delta X_0.L}{\delta t} = k_{\text{on}} X_0.L - k_{\text{off } 0} X_0.L - k_0 X_0.L \quad (36)$$

$$\frac{\delta X_u.L}{\delta t} = k_0 X_0.L - k_{\text{off } 1} X_u.L - k_1 X_u.L \quad (37)$$

$$\frac{\delta X_u}{\delta t} = k_{\text{off } 1} X_u.L \quad (38)$$

$$\frac{\delta X_{uu}.L}{\delta t} = k_1 X_u.L - k_{\text{off } 2} X_{uu}.L - k_2 X_{uu}.L \quad (39)$$

$$\frac{\delta X_{uu}}{\delta t} = k_{\text{off } 2} X_{uu}.L \quad (40)$$

$$\frac{\delta X_{uuu}.L}{\delta t} = k_2 X_{uu}.L - k_{\text{off } 3} X_{uuu}.L - k_3 X_{uuu}.L \quad (41)$$

$$\frac{\delta X_{uuu}}{\delta t} = k_{\text{off } 3} X_{uuu}.L \quad (42)$$

$$\frac{\delta X_{uuuu}.L}{\delta t} = k_3 X_{uuu}.L - k_{\text{off } 4} X_{uuuu}.L - k_4 X_{uuuu}.L \quad (43)$$

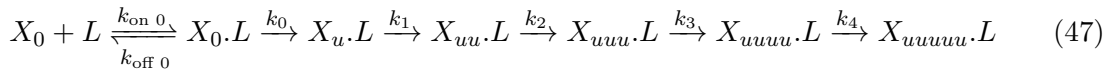
$$\frac{\delta X_{uuuu}}{\delta t} = k_{\text{off } 4} X_{uuuu}.L \quad (44)$$

$$\frac{\delta X_{uuuuu}.L}{\delta t} = k_4 X_{uuuu}.L - k_{\text{off } 5} X_{uuuuu}.L \quad (45)$$

$$\frac{\delta X_{uuuuu}}{\delta t} = k_{\text{off } 5} X_{uuuuu}.L \quad (46)$$

4 Unified Model of Polyubiquitination

Equations describing the unified model of polyubiquitination, where L is the E3 ligase and X is the substrate protein.



Ordinary differential equations:

$$\frac{\delta X_0}{\delta t} = k_{\text{off } 0} X_0.L - k_{\text{on } 0} X_0 L \quad (53)$$

$$\frac{\delta X_0.L}{\delta t} = k_{\text{on } 0} X_0 L - k_{\text{off } 0} X_0.L - k_0 X_0.L \quad (54)$$

$$\frac{\delta X_u.L}{\delta t} = k_0 X_0.L + k_{\text{on } 1} X_u L - k_1 X_u.L - k_{\text{off } 1} X_u.L \quad (55)$$

$$\frac{\delta X_{uu}.L}{\delta t} = k_1 X_u.L + k_{\text{on } 2} X_{uu} L - k_2 X_{uu}.L - k_{\text{off } 2} X_{uu}.L \quad (56)$$

$$\frac{\delta X_{uuu}.L}{\delta t} = k_2 X_{uu}.L + k_{\text{on } 3} X_{uuu} L - k_3 X_{uuu}.L - k_{\text{off } 3} X_{uuu}.L \quad (57)$$

$$\frac{\delta X_{uuuu}.L}{\delta t} = k_3 X_{uuu}.L + k_{\text{on } 4} X_{uuuu} L - k_4 X_{uuuu}.L - k_{\text{off } 4} X_{uuuu}.L \quad (58)$$

$$\frac{\delta X_{uuuuu}.L}{\delta t} = k_4 X_{uuuu}.L - k_{\text{off } 5} X_{uuuuu}.L \quad (59)$$

$$\frac{\delta X_u}{\delta t} = k_{\text{off } 1} X_u.L - k_{\text{on } 1} X_u L \quad (60)$$

$$\frac{\delta X_{uu}}{\delta t} = k_{\text{off } 2} X_{uu}.L - k_{\text{on } 2} X_{uu} L \quad (61)$$

$$\frac{\delta X_{uuu}}{\delta t} = k_{\text{off } 3} X_{uuu}.L - k_{\text{on } 3} X_{uuu} L \quad (62)$$

$$\frac{\delta X_{uuuu}}{\delta t} = k_{\text{off } 4} X_{uuuu}.L - k_{\text{on } 4} X_{uuuu} L \quad (63)$$

$$\frac{\delta X_{uuuuu}}{\delta t} = k_{\text{off } 5} X_{uuuuu}.L \quad (64)$$

References

- [1] Senadhi Vijay-Kumar, Charles E Bugg, and William J Cook. Structure of ubiquitin refined at 1.8 Å resolution. *Journal of molecular biology*, 194(3):531–544, 1987.
- [2] David H Schlesinger, Gideon Goldstein, and Hugh D Niall. Complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochemistry*, 14(10):2214–2218, 1975.
- [3] Daniel Finley, Seth Sadis, Brett P Monia, Paul Boucher, David J Ecker, Stanley T Crooke, and Vincent Chau. Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Molecular and cellular biology*, 14(8):5501–5509, 1994.
- [4] Avram Hershko and Aaron Ciechanover. The ubiquitin system. *Annual review of biochemistry*, 67(1):425–479, 1998.

- [5] Holger Bastians, Leana M Topper, Gary L Gorbsky, and Joan V Ruderman. Cell cycle-regulated proteolysis of mitotic target proteins. *Molecular biology of the cell*, 10(11):3927–3941, 1999.
- [6] Michael Glotzer, Andrew W Murray, Marc W Kirschner, et al. Cyclin is degraded by the ubiquitin pathway. *Nature*, 349(6305):132–138, 1991.
- [7] Vladimir Kirkin and Ivan Dikic. Ubiquitin networks in cancer. *Current opinion in genetics & development*, 21(1):21–28, 2011.
- [8] Hwan-Ching Tai and Erin M Schuman. Ubiquitin, the proteasome and protein degradation in neuronal function and dysfunction. *Nature Reviews Neuroscience*, 9(11):826–838, 2008.
- [9] Jodi L Stewart and Jennifer L Nemhauser. Do trees grow on money? Auxin as the currency of the cellular economy. *Cold Spring Harbor perspectives in biology*, 2(2), 2010.
- [10] Dior R Kelley and Mark Estelle. Ubiquitin-mediated control of plant hormone signaling. *Plant physiology*, 160(1):47–55, 2012.
- [11] Erich Kombrink. Chemical and genetic exploration of jasmonate biosynthesis and signaling paths. *Planta*, 236(5):1351–1366, 2012.
- [12] Laura B Sheard, Xu Tan, Haibin Mao, John Withers, Gili Ben-Nissan, Thomas R Hinds, Yuichi Kobayashi, Fong-Fu Hsu, Michal Sharon, Sheng Yang He, et al. Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature*, 468(7322):400–405, 2010.
- [13] Avram Hershko, Hannah Heller, Sarah Elias, and Aaron Ciechanover. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry*, 258(13):8206–8214, 1983.
- [14] Aaron Ciechanover, Daniel Finley, and Alexander Varshavsky. Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell*, 37(1):57–66, 1984.
- [15] Christoph H Emmerich, Anna C Schmukle, and Henning Walczak. The emerging role of linear ubiquitination in cell signaling. *Science signaling*, 4(204):re5, 2011.

- [16] Michael Rape, Sashank K Reddy, and Marc W Kirschner. The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell*, 124(1):89–103, 2006.
- [17] Giada Monami, Velia Emiliozzi, and Andrea Morrione. Grb10/Nedd4-mediated multiubiquitination of the insulin-like growth factor receptor regulates receptor internalization. *Journal of cellular physiology*, 216(2):426–437, 2008.
- [18] Olivier Staub and Daniela Rotin. Role of ubiquitylation in cellular membrane transport. *Physiological reviews*, 86(2):669–707, 2006.
- [19] Yaron Mosesson, Keren Shtiegman, Menachem Katz, Yaara Zwang, Gyorgi Vereb, Janos Szollosi, and Yosef Yarden. Endocytosis of receptor tyrosine kinases is driven by monoubiquitylation, not polyubiquitylation. *Journal of Biological Chemistry*, 278(24):21323–21326, 2003.
- [20] Kaisa Haglund, Sara Sigismund, Simona Polo, Iwona Szymkiewicz, Pier Paolo Di Fiore, and Ivan Dikic. Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nature cell biology*, 5(5):461–466, 2003.
- [21] Wei Li, Daqi Tu, Axel T Brunger, and Yihong Ye. A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. *Nature*, 446(7133):333–337, 2007.
- [22] Nathan W Pierce, Gary Kleiger, Shu-ou Shan, and Raymond J Deshaies. Detection of sequential polyubiquitylation on a millisecond timescale. *Nature*, 462(7273):615–619, 2009.
- [23] Michael Rape and Marc W Kirschner. Autonomous regulation of the anaphase-promoting complex couples mitosis to s-phase entry. *Nature*, 432(7017):588–595, 2004.
- [24] Guido Rossum. Python reference manual. Technical report, Amsterdam, The Netherlands, The Netherlands, 1995.
- [25] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001.
- [26] MATLAB. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts, 2010.
- [27] Philip Sedgwick. Pearson’s correlation coefficient. *BMJ*, 345, 7 2012.

- [28] John J Hopfield. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences*, 71(10):4135–4139, 1974.
- [29] Julia Yaglom, MH Linskens, Seth Sadis, David M Rubin, Bruce Futcher, and Daniel Finley. p34cdc28-mediated control of cln3 cyclin degradation. *Molecular and cellular biology*, 15(2):731–741, 1995.
- [30] Stefan Lanker, M Henar Valdivieso, and Curt Wittenberg. Rapid degradation of the g1 cyclin cln2 induced by cdk-dependent phosphorylation. *Science*, 271(5255):1597–1601, 1996.
- [31] Daniel Kornitzer, Bilha Raboy, Ro G Kulka, and GR Fink. Regulated degradation of the transcription factor gcn4. *The EMBO journal*, 13(24):6021, 1994.
- [32] Kwang-Ai Won and Steven I Reed. Activation of cyclin e/cdk2 is coupled to site-specific autophosphorylation and ubiquitin-dependent degradation of cyclin e. *The EMBO journal*, 15(16):4182, 1996.
- [33] J Alan Diehl, Frederique Zindy, and Charles J Sherr. Inhibition of cyclin d1 phosphorylation on threonine-286 prevents its rapid degradation via the ubiquitin-proteasome pathway. *Genes & Development*, 11(8):957–972, 1997.
- [34] Yoosik Kim, Ze’ev Paroush, Knud Nairz, Ernst Hafen, Gerardo Jiménez, and Stanislav Y Shvartsman. Substrate-dependent control of MAPK phosphorylation in vivo. *Molecular systems biology*, 7(1), 2011.
- [35] Julia S Thrower, Laura Hoffman, Martin Rechsteiner, and Cecile M Pickart. Recognition of the polyubiquitin proteolytic signal. *The EMBO journal*, 19(1):94–102, 2000.