Modeling of the signaling networks in patterning and growth control of the *Drosophila* wing disc

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Lin Lin

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Hans Othmer

January, 2016

Acknowledgements

First and foremost I would like to express my sincere gratitude to my advisor Hans Othmer for his tremendous mentorship, support and guidance during the process of my Ph.D study. My sincere thanks also goes to my co-advisor Michael O'Connor, who gave me the access to the lab facilities and taught me how to design and conduct biological experiments. I would like to thank the rest of my thesis committee, David Odde, Tay Netoff, and Yoichiro Mori, for their insightful comments, questions and encouragement. I would like to thank all the members of Hans Othmer's group for their advice, collaboration and friendship. I would like to thank a former group member, Zhan Chen, for his stimulating discussion on FRAP work.

Dedication

To my parents and my brother

Abstract

A major challenge in understanding patterning and growth control is how the signaling pathways are balanced to produce normal pattern and growth and how they interact to respond to aberrant signals. In the dissertation, we aim to develop a mathematical model which incorporates the Hippo pathway locating at the center of different regulatory pathways in the wing disc of *Drosophila* so as to be able to understand existing experimental results, to make experimentally-testable predictions, and to provide a platform for integrating and testing new results and incorporating other signaling pathways. The model we developed addresses the limitation of previous models due to lack of mechanistic details, and predicts all the primary characteristic phenotypes associated with the pathway. Moreover, the model supports two hypotheses, one of which have been confirmed by experiments.

As using a mathematical model to facilitate the development of biology is contingent on parameters, the other specific aim of our work is to propose a new way to improve parameter estimation from experimental data. We identify the source for poor estimation in Fluorescence recovery after photobleaching (FRAP), a widely-used technique for quantitative measurement of molecular dynamics, and propose three feasible ways to improve parameter estimation. In addition, we also introduce sensitivity analysis to improve model identification in FRAP.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Patterning and growth control

The control of patterning and growth in developmental biology is currently an active interdisciplinary area between biologists and physical scientists because the interaction of experimentation and modeling has produced significant new insights into a number of model systems. In biology, patterning means the creation of spatial, temporal, or spatio-temporal structure, while growth control refers to the regulation of growth rate, size, and shape of an organism or a tissue. For instance, the size of many organisms and their constituent tissues and organs vary substantially but differ little in morphology; they appear to be scaled versions of a common pattern. In addition, nutrient availability and competition within a growing tissue can affect the final size, but individual body parts scale properly to match overall size even in the face of harsh environmental conditions. The mechanisms that regulate and coordinate growth and patterning to produce scale invariance are poorly understood. Meanwhile, defects in the mechanisms that govern growth and patterning can cause major health issues such as cancer and birth defects. Of course the control of growth and patterning ultimately depends on the control of gene expression, but translating molecular-level information into an understanding of tissue- and organism-level functioning is one of the most challenging problems in biology. Advances in our understanding of this integrated process in a developing organism will certainly help to unravel the mechanisms as to how the various levels of biological complexity are integrated to control growth and patterning in normal development

and provide insight into our understanding of pathological process that contribute to diseases.

1.2 Mathematical modeling of signal transduction networks

The control of patterning and growth at the cellular, tissue and organism levels involves complex interactions within and between the diverse signal transduction and gene control networks. Despite molecular-level discoveries and insights from experiments, the *in vivo* behavior of signal transduction networks involved in development cannot be understood solely at the molecular level – integration of the signal transduction and transport processes involved in development at the tissue and organ level is necessary, and this in turn requires mathematical modeling and computational analysis. Moreover, because different pathways are tightly linked, interactions within the pathways are delicately balanced, and both points further the case for detailed mathematical models of these pathways and their interactions. A mathematical model, which incorporates the signaling pathways and their interactions, and explains the existing experimental results, can be used to explore how the developmental regulators influence growth and to evaluate postulated mechanisms that are difficult to explain by using verbal arguments and thought experiments alone. In addition, the plethora of experimental results need to be synthesized into a more comprehensive structure by modeling, and they provide data to test the models.

The availability of experimental data for signaling pathways has led to a shift from predominantly phenomenological models of patterning and growth control to mechanism-based models, the purposes of which are not only to explain the existing observations within a mechanistic framework, but also to serve as tools for discovery by experimentalists. Mathematical models for *Drosophila* oogenesis, Bcd patterning, BMP-mediated patterning, planar cell polarity, EGF patterning, and segment polarity have all led to experiments that may not have been carried out otherwise, and contributed greatly to our understanding of those systems [16, 17, 18, 19, 20, 21, 22].

The Hippo pathway functions as the hub of a variety of regulatory mechanisms that control growth and patterning of the wing disc, therefore, a mechanistic model that describes the Hippo pathway would provide the framework for integrating other

pathways. The Hippo pathway involves highly connected networks of interacting components. Most, if not all, existing models for this pathway attempt to stay simple to capture some features of the signaling pathway, and have provided considerable insights [23, 24, 12, 25, 26, 27]. However, they omit most of the details of the signaling network and provide little mechanistic understanding as how molecular interactions are translated to cellular function and tissue organization. Therefore, we sought to develop a more detailed model that incorporates the major molecular components in this pathway and the interaction of a cell with its neighbors via cadherins, and thereby to formulate the first detailed model of this pathway that accounts for signal transduction at cell level and cell-cell interactions at tissue level. Our detailed model provides a mechanistic understanding of existing experimental results, makes experimentally-testable predictions, and provides a platform for integrating new results and testing new hypotheses. We anticipate that our model together with the existing models will better predict the spectrum of the signaling network behavior, and might elucidate the mechanisms that control and coordinate growth and patterning at different levels.

An important and usually difficult step in testing models against experimental observations hinges on the determination of appropriate parameters for the models, since this involves extracting parameters from limited data in situations where many of the underlying steps may not be known. This interplay between models and experimental data raises the question of how useful the former are for extracting parameters from the available experimental data. When the underlying system involves the interaction of several transport and reaction steps, the question arises as to what level the system should be modeled at and how one is to estimate the parameters in the model from the available data. Here the issue of a phenomenological model versus a mechanism-based model arises again, and as we discuss later, the parameters estimated using a phenomenological model may be of marginal value in understanding mechanistic details in a system. An example described later that arose from studies of the *Drosophila* wing disc [15, 28] illustrates this in detail.

1.3 Parameter estimation using FRAP

FRAP is a widely-used technique for quantitative measurement of molecular dynamics. In a FRAP experiment, the fluorescence-tagged molecules in a region of interest (ROI) are first photobleached, and then the recovery of fluorescence within the ROI due to transport from the surrounding region is recorded [29]. By fitting the recovery data to a mathematical model, parameters that measure transport due to diffusion, binding and chemical reactions can be estimated. Traditionally FRAP experiments were used for cellular or sub-cellular level processes that occur on short time scales, and by fitting parameters such as diffusion coefficients and binding rates to the data, properties of cell- or sub-cellular-level processes could be inferred. More recently FRAP has been used for tissue-level studies that occur on a long time scale, where the results may be influenced by the interactions of production, transport, decay and other processes.

However, when the underlying system involves the complex interaction of multiple processes, the question arises as to what level the system should be modeled at and how one is to estimate the parameters in the model. This latitude can lead to wide discrepancies in the estimated parameters as one model may omit a process included in another. In recent studies, the complexity involved in the interactions of transport process with binding, internalization and kinetic transformation is hidden by analyzing the spread of morphogens using a simple reaction-diffusion system such as (4.16), and how one relates the 'real' parameters in individual steps to parameters in such a high-level description is usually difficult to determine analytically. This has lead to dramatic differences in the estimates of parameters for a simplified description such as (4.16), and raises the question as to what those parameters represent. In particular, estimates of diffusion coefficients in the wing disc reported from FRAP analysis in [15] differ by a factor of 200 from those measured in [28]. Even if the recovery models are identical, the parameter estimates may vary widely due to differences in the assumptions about the parameters, as will be described in an example later. Therefore, to the extent possible, a careful assessment of whether and how the transport and reaction processes couple should be made before a FRAP model is formulated, because otherwise the results may bear little relationship to the actual processes that determine the recovery curve. Our objective is to propose an approach to investigate this discrepancy, and improve

parameter estimation and model identification for FRAP experiments.

1.4 Aims and outline of the thesis

A major challenge in understanding patterning and growth control is how the signaling pathways are balanced to produce normal pattern and growth and how they interact to respond to aberrant signals. In this thesis, we aim to develop a mathematical model which incorporates the Hippo pathway locating at the center of different regulatory pathways in the wing disc of *Drosophila* so as to be able to understand existing experimental results, to make experimentally-testable predictions, and to provide a platform for integrating and testing new results and incorporating other signaling pathways. As using a mathematical model to facilitate the development of biology is contingent on parameters, the other specific aim of our work is to propose a new way to investigate the discrepancies in parameter estimation, identify the source for poor estimation, and improve parameter estimation and model identification.

In chapter 2, we first introduce the background on *Drosophila* melanogaster and its wing disc as our model system. It is followed by the review of the primary signaling pathways that control growth and patterning during the development of the wing disc, with a focus on the Hippo pathway. Then the current state of research on modeling the Hippo pathway is discussed. In addition to signal transduction pathways within cells, the signal transport mechanism across cells are also discussed. After that, we discuss the FRAP background and the application of FRAP in parameter estimation. Finally, the mathematical basis used in this paper is briefly introduced.

In chapter 3, we first introduce how the model is formulated based on our review and analysis of literature. It is a significant effort to evaluate all the existing regulatory mechanisms embedded in numerous studies of this signaling network, and to identify and incorporate the primary contributors to growth control and patterning, driven by their known impact and the questions we intend to answer. The model is divided into several modules based on the function and localization of the components within the network. For the mechanisms that have not yet been identified, plausible hypotheses have been made. The model is then validated by matching disc-wide and clone experimental results, and used to make experimentally-testable predictions. In addition, given the

large uncertainty in many parameters, we introduce parameter sensitivity analysis and perturbation analysis to identify key parameters and mechanisms responsible for control of growth and patterning. The model aims to complement the existing models without oversimplification. The model can serve as a foundation for future work which can extend the findings, much as the models for morphogen and hormone pathways have done.

In chapter 4, we first give the mathematical framework of FRAP analysis in general and specifically as applied to the wing disc. Then in order to demonstrate the effect of model assumptions and data utilization on parameter estimation, we generate data computationally for a known model with known parameters and then test our recovery of parameters from the data, which avoids the difficulties to evaluate estimates due to unknown mechanisms and other factors in FRAP experiments. By using a recovery model identical to the theoretical model, we identified the factors that lead to poor estimates and proposed feasible ways to improve parameter estimation. By varying the recovery model from the theoretical model, we investigate the relationship between parameters in the theoretical model and those in the recovery model, and under what conditions some processes can be neglected in the recovery model. Lastly, we introduce sensitivity analysis as a technique to better understand FRAP data and to improve FRAP model identification and parameter estimation.

Finally, in chapter 5 conclusions and future directions are discussed

Chapter 2

Background

2.1 *Drosophila* melanogaster

The common fruit fly, *Drosophila* melanogaster, is one of the most commonly used and genetically tractable model organisms in biology and physiology, due to their short life cycle, large brood size,and the availability of well-developed molecular techniques. Studies of various vertebrates have revealed that the signal transduction and gene regulatory networks that control growth and patterning in *Drosophila* are highly conserved across species. As a result, elucidating the mechanisms that control developmental processes in Drosophila further our understanding of such processes in other organisms, including humans, and expands our knowledge and treatment of conditions such as birth defects and cancer.

The life cycle of Drosophila consists of a number of stages: embryogenesis, three larval stages, the pupal stage, and the adult stage. Its growth is restricted to the embryonic and larval stages, and the final size of an adult is determined by the size of the larva when it stops feeding. Therefore, studies on patterning and growth regulation usually focus on the larval stages and, in particular, on imaginal discs. Imaginal discs are pouches of epithelium composed of undifferentiated cells invaginated from embryonic ectoderm and formed and matured at larval stage, and they are the precursors of the appendages in the adult. Wing imaginal discs are specified at the embryonic stage in response to two secreted factors, Wingless (Wg) and Decapentaplegic (Dpp), which are called morphogens, since they are produced from local sources and spread through the

wing tissue, forming a spatial distribution that specifies cell fate in a concentration-dependent manner. Dpp and Wg continue to function throughout the larval stages to affect both patterning and growth of the disc.

At the beginning of the 1st instar (24h After Egg Laying (AEL)) the wing disc contains 40 cells. During the larval development, there is 1000 fold increase in cell number. By late third instar a disc has $\sim$ 50K cells and the disc pouch is 150 μm (Dorsal-Ventral, DV) x 300 μm (Anterior-Posterior, AP). Wing discs are two-sided sacs comprising two layers separated by a lumen ($\sim$ 6μm), a layer of epithelial columnar cells with their apical side at the lumen and an overlying peripodial epithelium. The plasma membrane of each columnar cell is divided into apical and basolateral domains. These domains are separated by adherens junctions (AJs - aka zonula adherens (ZA)), an adhesive belt around the cell, comprised of cadherins and catenins, and linked to the cytoskeleton, and by septate junctions (SJs) that lie basal to the AJs and constitute a second permeability barrier for extracellular apical-basal transmission. These junctions maintain the cell shape in the epithelium, link cells to their neighbors, and divide the extracellular fluid into apical and basal layers, separated by a distinct layer between the AJs and SJs.

2.2 Signal transduction and gene control networks

The *Drosophila* wing disc is an elegant and self-organizing system that is ideal for the study of the signal transduction and gene control networks for patterning and growth control. The pathways that regulate growth and patterning of the wing disc can be divided into two categories: the intrinsic pathways, which control growth and patterning in response to the local signals, and the extrinsic pathways, which coordinate growth of the different tissues in the organism. The local signal pathways include morphogen-controlled pathways such as Dpp and Wg pathways and cell-cell interaction pathways such as the Hippo pathway. The global signal pathways include nutrient-response pathways such as target of rapamycin (TOR) signaling pathways, and hormone pathways such as insulin and ecdysone pathways. Patterning is coupled to growth control via interactions between the Dpp-Wg patterning pathways, and the Hippo pathway for

growth control. Local and global growth control are coordinated via interactions between nutrient and hormone pathways and the Hippo pathways.

2.2.1 Morphogen pathways

Early patterning of the disc defines compartments, which behave as independent units of size control, and the boundaries of which are not crossed by cells [31, 2]. The anterior and posterior compartments are divided early in the embryo and are maintained throughout the larval stage, and are defined by whether or not cells express the selector gene *engrailed* (*en*). En is a transcription factor that identifies cells in posterior compartment and changes cell surface properties to set up the compartment boundary. En autonomously induces *hedgehog* (*hh*) expression and represses activation of the transcription factor cubitus interruptus in the posterior compartment. Hh functions as a short-range signal to induce expression of Dpp, which is both a morphogen and a growth factor, in a stripe of cells anterior to the AP compartment boundary [32, 2, 33]. This leads to a spatially graded distribution of Dpp in the AP direction [15, 34].

Dpp binding to the receptor thickvein (Tkv) induces expression of *spalt (sal)* at high levels and *optomotorblind (omb)* at lower levels, and indirectly represses *tkv* [35, 36]. pMad and Medea, the effectors of Dpp signaling, bind with the repressor protein Schnurri and thence to the *brinker* (*brk*) enhancer, thereby relieving *brk*-mediated repression of *omb* and *sal*. *sal* activity is required to promote cell proliferation in the center of the disc [37], and *omb* exerts a negative feedback on Hh signaling [38], and both inhibit apoptosis. How Dpp is transported from the production site to establish the gradient is not well understood [39]. Free or hindered diffusion in the extracellular space has been suggested [28, 39], but there is also evidence for transcytosis [15]. As we show later, the elusive transport mechanisms have led to dramatic discrepancies in estimates of parameters, which, on the other hand, play an important role in studies of transport and downstream mechanisms.

The second compartment boundary between the dorsal and ventral regions develops during the second larval instar of development. The DV boundary separates cells that express the homeobox gene *apterous* (*ap*) (D) from those that don't (V). Wg is initially expressed throughout the ventral compartment, but during the third instar this is restricted to a strip at the DV boundary [40]. The Wg gradient along the DV axis induces

expression of the downstream genes *vestigal* (*vg*) and *distalless*. Wg affects disc growth through the vestigal protein (Vg) [41, 24]. Thus the initial fate of cells in the DP is determined by whether or not they express *en* and/or *ap*, and the level of Wg, Hh and Dpp to which they are exposed.

Hh, Dpp and Wg are the primary factors in patterning and much is known about their spatio-temporal patterns of expression. However, the details of their long-range transport and the downstream mechanisms that integrate signals from these networks are not fully understood. Furthermore, recent experimental studies on the wing disc have elucidated the role of morphogen signaling pathways in growth control, and results on the interaction of these pathways with the Hippo pathway, which controls cell proliferation and apoptosis, have emerged. These and other genetic studies illustrate the tight linkage between the patterning and growth processes, but despite such discoveries, how the morphogen-controlled pathways are integrated with the Hippo pathway to regulate patterning and growth is not understood. In addition, studies suggest that cell-cell interactions mediated by the Hippo pathway also play an important role in achieving normal disc size.

2.2.2 The Hippo pathway

The Hippo pathway, also known as Salvador/Warts/Hippo(SWH) pathway, consists of a highly conserved core kinase cascade the serine/threonine kinases Hippo (Hpo) and Warts (Wts) and the adaptor proteins Salvador (Sav) and Mob as tumor suppressor, and takes its name from the fact that mutation of the *hippo* gene gives rise to a large fly with folds of overgrown tissues that resemble the body folds of a hippopotamus. The key effector in the Hippo pathway is Yorkie (Yki), a co-transcription factor whose nuclear localization is controlled by the kinase Warts (Wts). Wts phosphorylates Yki to prevent it from entering the nucleus. Yki acts by binding to transcription factors such as Scalloped (Sd) to activate the expression of *cyclin E, myc, DIAP1*, and *bantam* that control cell proliferation, and controls expression of genes upstream in the Hippo pathway, such as *expanded, merlin, kibra*, and *four-jointed* (*fj*).

The Hippo pathway, which regulates many downstream transcription factors involved in cell growth, proliferation and apoptosis, is regulated by many upstream factors. Two atypical cadherins, Fat and Dachsous (Ds) are key upstream components in

the Hippo pathway. Both are enormous molecules with intracellular, transmembrane and extracellular domains, and can be localized in the apical cell membrane. The intracellular domains (ICDs) of each can independently mediate signaling through the Hippo pathway within the cell, while Fat and Ds on adjacent cell membranes can also bind with each other to strengthen the signaling and hence mediate cell-cell interaction and regulate growth.

Signaling from FatICD has been shown to suppress growth via the downstream factor Dachs, an unconventional myosin. Although the mechanism is still unclear, *dachs* is epistatic to *fat* in terms of both growth and gene expression phenotype, i.e. loss of *dachs* completely suppresses the ability of the *fat* loss-of-function mutant to induce overgrowth or target gene overexpression. In normal development, Dachs is low in the cytoplasm and high near the adherens junctions, and preferentially accumulates on the distal sides of cells. In *ds* mutants, strong but unpolarized membrane localization of Dachs is detected, just as it is in *fat* mutants. In *fat* mutants, the Dachs protein level does not have visible change, which indicates that the influence of Fat on Dachs most probably reflects protein re-localization. Experiments suggest that while the polarization of Dachs controlled by Fat and Ds is essential for cell polarization, the amount of Dachs localized on the membrane controls cell growth. Localized Dachs can associate with the kinase Warts (Wts) and promote the degradation of Wts protein [42]. In the wing, overexpression of *dachs* increases wing size, while *dachs* loss-of-function mutant decreases wing size [43].

Signaling from Ds ICD positively regulates growth via direct interaction with Riquiqui (Riq), a 343-amino-acid protein with five WD40 repeats predicted to mediate protein-protein interaction, and Minibrain (Mnb), a DYRK family kinase. Ds is required for the Riq protein to localize to the apical junction of cells in the wing disc. Localized Riq alone does not affect Wts ability, but rather in turn potentiates the ability of Mnb to phosphorylate Wts to reduce its activity and hence increase Yki activity. Depletion of Riq or Mnb decreases the wing size while overexpression of Riq or Mnb increases the wing size. While Ds binding to Fat enhances the inhibitory effect of Fat on Dachs localization, Fat binding to Ds also increases the interaction between Ds and Riq and thereby enhances Riq localization. Recent studies suggest a potential interaction between Ds ICD and Dachs [25]. It is likely that the interaction is also reinforced upon Fat binding to Ds. However, since modulating the expression of either Riq or Mnb does

not influence Dachs levels or localization in wing discs [44], it is possible that either Ds ICD has two independent binding sites for Riq and Dachs respectively, or Ds only interacts with localized Dachs.

The interaction between Ds and Fat mediates cell-cell interaction, and is regulated by another protein Four-jointed (Fj), which itself is one of the target genes of the Hippo pathway. Fj functions in the Golgi as a kinase, and it phosphorylates cadherin domains of Fat and Ds that are destined to be extracellular [45]. By directly phosphorylating cadherin domains of Fat and Ds, Fj promotes the ability of Fat to bind to Ds, and inhibits the ability of Ds to bind to Fat [46]. However, the weaker phenotype of *fj* mutants in comparison to *ds* mutants, and the ability of cells expressing high levels of Fat and Ds to aggregate without exgenous Fj, imply that Fat has some Ds-binding ability even without Fj [46].

A plethora of experimental observations resulting from mutant clones and various other interventions raises numerous questions concerning how Fat and Ds regulate the Hippo pathway to mediate cell-cell interaction and to control cell and tissue growth. Recent studies have shown the non-monotonic effect of Fat on the growth [47], i.e. overexpression of *fat* decreases the wing size and the complete knockout of *fat* increases the wing size, but the partial knockout of *fat* decreases, rather than increases, the size. Likewise, it appears that the effect of Ds is also non-monotonic given the existing experimental results. While loss of Ds results in enlarged wings and wing discs [48], overexpression of Ds using different Gal4 drivers may lead to the opposite effect on target gene expression and growth, i.e. it may either reduce [48, 49] or enhance growth [44]. In addition, double mutants of *fat* and ds overgrow more than either of the single mutants [10]. Furthermore, the expression level of Fj, the modulator of Fat and Ds as well as the target gene of this signaling pathway, also causes non-monotonic response in growth, i.e. either loss of fj or overexpression of fj across the disc is associated with modest reductions in wing size. Finally, when Fj and Ds are co-overexpressed, the reduction in wing size is greater than overexpression of either of them [48, 49].

These observations resulting from interference across the disc are not intuitive, given that Fat and Ds are recognized as the growth suppressors, and so also are the observations with clone cells. Overexpression of *ds* in clone cells upregulates Hippo target genes in cells on both sides of the border [9, 49], while *ds* lost-of-function mutant clones

upregulates Hippo target genes outside of clone borders but not in *ds* mutant cells [9], which requires both Fat and Dachs since loss of either of them suppresses the effects. Moreover, although the overexpression of target gene expression is observed both autonomously and non-autonomously when Ds is overexpressed in posterior compartment, the autonomous effect in the posterior is greatly reduced when either Riq or Mnb is depleted and the non-autonomous Yki hyperactivation in anterior cells near the AP boundary is unaffected [44]. Such a non-autonomous effect is also reported when Fat is overexpressed [50, 51] although no such effect has been observed when Fat is lost partially or completely. In addition, both *fj* loss-of-function clones and *fj* overexpression clones upregulate Hipp target genes on both sides around the clone border. Furthermore, the non-autonomous effects of Fj and Ds are more pronounced on the side that clone cells and wild-type cells have greater difference. In addition to the observations in space, investigations have been conducted in time. It has been found that the inhibitory effect on growth induced by overexpression of either Fj or Ds is transient, and the suppressed proliferation begins to recover by 50hrs and is detected throughout the disc later. The phenotypes and experimental observations related to the Hippo pathway are summarized in Table 2.1.

The growth in the disc is quite uniform in normal development. However, besides the fact that morphogens such as Dpp and Wg form gradients along their respective axes, the Hippo pathway regulators, Ds and Fj, are also expressed unevenly in the disc. The expression of Ds is lower in the distal region and higher in the proximal domain, and vice-versa for Fj. The non-uniform expression of growth regulators further show that growth is controlled and accommodated by multiple pathways. In fact, how the gradients are converted to uniform growth and how the size of tissue is determined remain unanswered.

We are interested in the Hippo pathway as it provides an integration point for all the local effects including Dpp, Wg morphogen signaling pathways, Frizzled-Dishevelled (Fz-Dsh) planar cell polarity (PCP) pathways, and cell mechanics pathways, and interacts with global regulatory signals such as TOR and insulin pathways. For instance, it was shown that Dpp signaling through the Fat/Hippo pathway is necessary for the effect of Dpp on proliferation [54, 49], and the Hippo pathway also modulates morphogen signaling through transcriptional control of *dally* and *dlp* [55] that affect Dpp transport and

gradient in the disc. Driving uniform expression of *vg*, the target gene of Wg signaling pathway, downregulates *ds* expression and upregulates *fj* expression. *fat* mutant causes significant developmental delay, which might be through its impact on the insulin signaling pathway. The interaction between the Hippo pathway with other signal pathways happens almost on every level and only the part which is closely related to our discussion is touched upon here. A few recent reviews are given in [7, 56, 57, 58, 59, 60, 61]. Given its central role in growth control, it is no surprise that disrupted Hippo pathways can cause major health issues either directly through pathway mutations or indirectly through crosstalk from other signaling network. Studies on the Hippo pathway will certainly shed light on other signaling pathways and the mechanism of overall patterning and growth control, and might suggest opportunities for therapeutic intervention for diseases.

Most of the current mathematical models on the Hippo pathway concentrate on the cell polarity, and address various aspects of the signaling pathway. Some of them focus on the role of Dachs on cell mechanics and suggest polarized Dachs induced by the gradient of Ds within the disc promotes anisotropy of cell-cell junction tension, which contributes to cell orientation arrangement [25]. Others incorporate the Fz-Dsh pathway and point out that the interaction between the Fat-Ds pathway and Fz-Dsh pathways is the key to generating the domineering nonautonomy phenomenon, in which mutations in PCP signaling components cause polarity disruptions of neighboring wild-type cells [17]. In addition, several groups favor that the opposing gradients of Fj and Ds in the disc induce subcellular asymmetries of DsFt heterodimers and provide subcellular asymmetric molecular cues that are available to orient PCP to the tissue axes [27, 12, 23, 47]. It is also worth noticing that there are some preliminary efforts on integrating the Hippo pathway and morphogen pathways [24, 26]. Nevertheless, none of the existing models attempt to model the Hippo pathway mechanistically, and despite the insights, how change in Fat and Ds is converted to cell response in growth is hard to be understood from these high-level models. Our research represents the first attempt to develop a detailed model of this pathway with the goal to understand this type of questions.

2.2.3 Transport mechanisms

While the signal transduction pathways within a cell has been extensively studied, how signals are transported across cells is still unclear. Several mechanisms for signal transport have been proposed and further discussion of different modes can be found in [62]. One possible way to transfer the signal produced by source cells is via secretion into the extracellular space followed by re-uptake for other cells by receptor internalization or other mechanisms – a process known as transcytosis . For instance, Hh transport may involve transcytosis, in that Hh is secreted apically, re-internalized in the secreting cells, and then re-secreted baso-laterally [63, 64].

The second way is called 'Quorum communication', in which all cells release the signaling species and all detect and perhaps internalize the signal [65]. While not useful for establishing a spatial pattern, this mode does serve as a quorum-sensing mechanism [66, 67, 68] that can control gene expression and cell motility, and may be involved in size-regulation in some systems [69]. In the context of the *Drosophila* wing disc, this form arises if signals are released into the luminal space. For example, inhibition of Dpp signaling only in the peripodial cells nevertheless disrupts growth and patterning of the wing [70], suggesting that mechanisms that govern the growth and patterning of peripodial cells coordinate with those of columnar cells.

Juxtacrine signaling is another way, which invloves direct cell-cell communication, either via gap junctions or direct receptor-ligand interactions. In the *Drosophila* wing disc, direct communication between cells occurs via Fat in the membrane of one cell binding to Ds in the membrane of a neighboring cell, and this is used for establishing compartment boundaries or directing neighboring cells to adopt different fates [42].

In addition, active transport, for instance along microtubules or cytonemes [71], may play a role in polarizing cells. While some evidence supports the idea that Hh is transported by transcytosis, other evidence suggests that cytonemes are involved in Hh transport [72, 73, 74].

2.3 Parameter estimation and FRAP

The FRAP technique was originally developed to study protein mobility in living cells and successfully measured the diffusion on cell membranes. Recent development of both

fluorescent protein technology and confocal microscopy have led to a marked increase in the use of FRAP not only for studying molecular mobility in the cell, but also for assessing molecular dynamics and interactions with others. Nowadays, FRAP has been adopted as a common technique for studying all aspects of biology, including chromatin structure, transcription, mRNA mobility, protein recycling, signal transduction, versicle transport and etc. The literature on FRAP analysis is large and can only be touched upon here, but a recent review is given in [75]. Some background information relevant to our analysis is given as follows.

Confocal laser scanning microscopes (CLSMs) are currently the widely used platforms for FRAP experiments. CLSMs are equipped with lasers for photobleaching and detection. The setup and the sequence of steps in a typical FRAP experiment that uses a CLSM are as follows. Assuming that diffusion is the only transport process, the steps are modeled as shown below, which also explains Figure 2.9. We describe them in generality here, and detailed descriptions of the models used for parameter estimation are given later. Both bleaching and scanning steps are done pixel by pixel and line by line.

Step 1: Prebleaching Use a low intensity laser to scan the fluorescence density in the entire domain Ω which includes the ROI. The evolution of the concentrations of fluorescent molecules during the pre-bleaching time $[0, T_0]$ (*cf.* Fig. 2.9) is governed by

$$\frac{\partial c}{\partial t} = D\bigtriangledown^2 c + R(c) - I_p c$$

where R(c) represents the reaction processes, and I_p is the pre-bleaching function. The pre-bleaching process is usually modeled as a first-order process as above, and in a 2D domain the intensity is given by

$$I_p(t, x, y) = I_{p_0} exp\left(-\frac{2[(x - X(t))^2 + (y - Y(t))^2]}{r_0^2}\right),$$

wherein $x = X(t), y = Y(t)$ and $x, y \in \Omega$ describes the pre-bleaching path of the laser.

Step 2: Bleaching Use a high intensity beam to bleach the ROI for an interval of length T_1. In this phase

$$\frac{\partial c}{\partial t} = D\bigtriangledown^2 c + R(c) - Ic$$

with initial conditions obtained from the end of the pre-bleaching period. The differences between pre-bleaching and bleaching are the laser intensity and pixel dwell time for scanning, as well as the scanning domain.

Step 3: Postbleaching This usually comprises a waiting time T_2 between the end of bleaching and the beginning of observation. During the observation time use a low intensity beam to image the fluorescence recovery process. During the waiting time

$$\frac{\partial c}{\partial t} = D\bigtriangledown^2 c + R(c)$$

with initial conditions obtained from the end of step 2.

During the observation period T_3

$$\frac{\partial c}{\partial t} = D\bigtriangledown^2 c + R(c) - I_p c$$

with initial conditions obtained from the end of the waiting period.

The notation used in the figure is as follows. Let $b = \Omega_{ROI}/\Omega$ denote the ratio of the area of the ROI to the area of the entire domain. Let a be the ratio of the remaining fluorescent intensity after bleaching to the fluorescent intensity before bleaching, assuming that bleaching is homogeneous and instantaneous, and let c be the fraction of fluorescent molecules which are immobile on the time scale of the experiment. The loss of fluorescence due to bleaching reflected in a recovery curve is $b(1-a)$, and the immobile fraction outside the ROI is $c(1-b)$.

The FRAP recovery curve is usually normalized as

$$\mathrm{F}_{norm}(\mathrm{t}) = \frac{\mathrm{F(t)\ -\ F(T_0)}}{\mathrm{(1\text{-}b)(1\text{-}c)+ab}},$$

wherein F is the fluorescence intensity and $t \geq T_0 + T_1$. Later we assume that bleaching is complete, i.e. $a = 0$, and that no immobile fraction exists, i.e. $c = 0$

The data obtained is usually averaged over the ROI and presented as a recovery curve. The parameter estimation step consists in fitting this data with a 'suitable' model, but since the recovery portion typically can be fit with a sum of time-dependent exponential terms, this leaves wide latitude as to what underlying processes are to be included, and once that is fixed, what meaning can be ascribed to those parameters.

The inconsistencies in FRAP also come from different assumptions for models. One of the disparate assumptions lie in that of initial conditions for recovery phase, the

obtainment of which can be divided into two categories among existing FRAP models. The first one is modeling all the previous processes before the recovery phase to get the initial conditions [76, 77, 29, 78]. By modeling the pre-bleaching, bleaching and recovery phase, the analytical solution representing FRAP data can be derived in some special cases and with certain assumptions [29]. It provides us insights about how the bleaching process and others affect the FRAP recovery, however, it requires prior knowledge of some parameters to model activities of the laser beam. Another way which is more popular to estimate parameters is to only model the recovery phase with assumptions about the initial conditions which can be obtained by the first post-bleach image of fluorescence. In this category, there are three different kinds of assumptions which all neglect the bleaching effect with imaging process in recovery phase. One natural assumption is the piece-wise constant initial conditions from direct measurement of the size of the photobleaching spot [79]. This is validated when the bleaching is homogeneous and the bleaching time is negligible. Another one is also the piece-wise constant initial conditions but are deduced from the final recovery concentration and the conservation of fluorescence [80]. It is an improvement from the first one, but still is limited to the cases where the boundaries of the bleached region are relatively sharp and the observational photobleaching is negligible. When the boundaries of the bleached region are smoothed by diffusion before observation, the piece-wise constant is not able to capture the characteristics of the initial condition, which may cause large errors in parameter estimation [81]. Thus for the third kind, some Gaussian or Gaussian-edge function fitted by the initial data, which results from the Gaussian assumption of the laser profile, is used for the initial condition [81]. It has been shown that using the Gaussian expression of the initial post-bleach profile alters the estimates of parameters in comparison to the first or second utilization of initial condition, however, the utilization of Gaussian function needs to be carefully justified. It is not convincing that the utilization of a Gaussian function as the initial condition produces better estimations, because in a real biological system the true values of binding/unbinding rates as well as the diffusion coefficient are unknown. In our simulations, we only simulate the data with homogeneous and instantaneous bleaching process, and thus we estimate parameters by only modeling the recovery with piece-wise constant initial conditions.

Traditionally FRAP experiments were used for cellular or sub-cellular level processes

that occur on short time scales, and by fitting parameters such as diffusion coefficients and binding rates to the data, properties of cell- or sub-cellular-level processes could be inferred. More recently FRAP has been used for tissue-level studies that occur on a long time scale, where the results may be influenced by the interactions of production, transport, decay and other processes, and a major issue in the use of FRAP in this context is what model should be used as the basis for parameter estimation. This latitude can lead to wide discrepancies in the estimated parameters, since one recovery model may omit a process included in another. Even if the recovery models are identical, the parameter estimates may vary widely due to differences in the assumptions about the parameters, as will be described in an example later. Therefore, to the extent possible, a careful assessment of whether and how the transport and reaction processes couple should be made before a FRAP model is formulated, because otherwise the results may bear little relationship to the actual processes that determine the recovery curve.

2.4 Mathematical Background

2.4.1 Perturbation Analysis

Suppose the system at steady state is $G(u, p) = 0$ where u represents the state variables and p represents the parameters with solution u^s

(1) Perturbation of the steady state

Linearize the system around the steady state:

$$G_u(u^s, p)\xi = 0$$

where $u^s = u - \xi$

(2) Structural perturbation

$$G(u, p) + R(u, p) = 0$$

Linearize around the steady state:

$$G(u^s + \xi, p) + R(u^s + \xi, p) = 0$$

$$G(u^s, p) + G_u(u^s, p)\xi + R(u^s, p) + R_u(u^s, p)\xi = 0$$

$$\xi = -[G_u(u^s, p) + R_u(u^s, p)]^{-1} R(u^s, p)$$

2.4.2 Parameter Sensitivity Analysis

Given the uncertainty of parameters in the model, sensitivity analysis is necessary for the investigation of how the change or error of the parameter values affect the model output and their impacts on conclusions to be drawn from the model. There is a very large and diverse literature on sensitivity analysis, and we here only focus on the techniques used in our analysis.

Local sensitivity analysis

The dependence of steady state on parameters is calculated as :

$$G_u(u^s, p)u_p + G_p(u^s, p) = 0$$

$$u_p = -G_u^{-1} G_p$$

It can be improved in such a way as to obtain a sensitivity measure that would rank parameters consistently, i.e. $\frac{\partial u}{\partial p}\frac{p}{u}$

The derivative-based local sensitivity analysis has the attraction of being very efficient in computational time. However, it is limited when the parameters have large uncertainty and when the model has non-linearity.

Input/output scatterplots

Input/output scatterplots are in general a very simple and informative way of running a sensitivity analysis, since they can provide an immediate visual depiction of the relative importance of the factors. For example, the existence of shape or pattern in the points of a scatterplot identifies an important factor as shown on the left of Figure 2.10, while little shape, which presents a rather uniform cloud of points over the range of the input factor, is a sign of a non-influential factor as shown on the right of Figure 2.10.

Variance-based sensitivity analysis

The non-local analysis described here is more informative than simply computing the derivatives of the output at a local state because parameters can be varied over large intervals. Here we use a form of variance-based sensitivity for models given in the form $Y = f(X_1, X_2, ...X_k)$, where Y is a model output and $X_1, X_2, ...X_k$ are factors with respect to which the sensitivity of the output is to be determined [82, 83]. In applying this to parameter sensitivity analysis, the factors X are the parameters, and the model is written as $Y = f(p_1, p_2, ...p_k) = f(P)$. In general, and unless otherwise specified, we assume that the input factors are independent of each other, so that each one can be independently sampled from its marginal distribution.

One could conceive of using $V_{P\sim i}(Y \mid p_i)$ as a measure of the relative importance of P_i, reasoning that the smaller $V_{P\sim i}(Y \mid p_i)$ is, the greater the influence of P_i is. To remove the dependency on P_i, the average of this measure is taken over all possible p_i, and is written as $E_{p_i}(V_{P\sim i}(Y \mid p_i))$. The law of total variance states that

$$V(Y) = V_{p_i}(E_{P\sim i}(Y \mid p_i)) + E_{p_i}(V_{P\sim i}(Y \mid p_i)) \tag{2.1}$$

Hence a small $E_{p_i}(V_{P\sim i}(Y \mid p_i))$, or a large $V_{p_i}(E_{P\sim i}(Y \mid p_i))$, will imply that P_i is an important parameter. $V_{p_i}(E_{P\sim i}(Y \mid p_i))$ is called the first-order effect of P_i on Y and the first-order sensitivity index of p_i on the output Y is defined as

$$S_i = \frac{V_{p_i}(E_{P\sim i}(Y \mid p_i))}{V(Y)} = 1 - \frac{E_{p_i}(V_{P\sim i}(Y \mid p_i))}{V(Y)} \tag{2.2}$$

where $V(\cdot)$ is the variance and $P_{\sim i}$ indicates that the expectation of the variance is taken with respect to all but the i^{th} parameter. Thus $E_{P\sim i}(Y \mid p_i))$ is the expected value of Y that results from averaging over all but p_i. To remove the dependency on the fixed value of p_i, the variance is taken with respect to p_i.

The first order index measures the contribution of parameter p_i to the total variance, and since it is normalized, it lies in $[0, 1]$. A large S_i indicates that the parameter p_i contributes a large fraction of the total variance, and thus can be regarded as an important parameter in setting the error. For an additive model, $\sum_{i=1}^{k} S_i = 1$, while for a non-additive model the first order terms do not add up to one, and higher-order interactions amongst the parameters account for some of the variance. For example, if we describe the parabolic cylinder as $p_1^2 + p_2^2$, then there is no interaction between

the parameters and $\sum S_i = 1$. This is not the case in our model, since the processes described by the parameters are highly coupled and interacted.

A second measure of sensitivity is obtained as follows [82]. The total variance can also be written

$$V(Y) = V_{p_i}(E_{P\sim i}(Y \mid p_i)) + V_{P\sim i}(E_{p_i}(Y \mid P_{\sim i})) + V_{p_i,P\sim i}$$

wherein the last term accounts for the variance due to the interactions. If the parameter p_i contributes little to the total variance then the sum of the first and last terms is approximately zero, which means that

$$V(Y) \sim V_{P\sim i}(E_{p_i}(Y \mid P_{\sim i})).$$

Thus an alternate measure of a parameter's effect is the total order sensitivity index of p_i on Y, which is defined as

$$S_{Ti} = 1 - \frac{V_{P\sim i}(E_{p_i}(Y \mid P_{\sim i}))}{V(Y)} = \frac{E_{P\sim i}(V_{p_i}(Y \mid P_{\sim i}))}{V(Y)}$$

From the first equality one sees that S_{Ti} is the expected variance due to the first and higher order effects of p_i on Y. For the following simulations the first and total order indices are calculated by using the method of Sobal' which costs $(k+2)N$ model runs, where k is the number of parameters and N is the number of sample points in parameter space. $N = 1000$ is used for all our simulations.

Sensitivity analysis is driven by the setting. When the purpose of the analysis is to prioritize factors, the first-order sensitivity index is a a natural choice. If the objective is to fix non-influential factors, the total sensitivity index comes to use. All of these settings, the computation of derivatives is advisable for a general understanding of the model.

Implementation of sensitivity analysis

The indices of variance-based sensitivity measures are usually done by the Monte Carlo Method. In practice, instead of generating pseudo-randomly distributed points in the parameter space in traditional Monte Carlo Method, the low-discrepancy quasi-random number generator is used to improve the efficiency of the estimators. This is known as the Quasi-Monte Carlo method. It is implemented as follows (taken from [83]).

1. "Generate an $N \times 2k$ sample matrix with respect to the probability distributions of the input variables- the parameters. N is the number of sample points. k is the dimension of parameter space, i.e. the number of parameters.

2. Use the first k columns of the matrix as matrix A, and the remaining k columns as matrix B, which generates two independent samples of N points in the k-dimensional parameter space

3. Build k further $N \times k$ matrices A_B^i for $i = 1, 2, ..., k$, such that the ith column of A_B^i is equal to the ith column of B, and the remaining columns are from A

4. The A, B, and the k A_B^i matrices specify $N \times (k+2)$ points in the parameter space (one for each row). Run the model at each point, giving a total of $N \times (k+2)$ model evaluations, i.e. the corresponding$f(A)$, $f(B)$ and $f(A_B^i)$ values

5. Calculate the sensitivity indices using the estimators discussed below."

There are a number of Monte Carlo estimators for both indices. Two that are currently widely used are according to the rule proposed by Saltelli et al. (2010).

$$V_{p_i}(E_{P\sim i}(Y \mid p_i) \approx \frac{1}{N}\sum_{j=1}^{N} f(B)_j(f(A_b^i)_j - f(A)_j)$$

$$E_{X\sim i}(V_{p_i}(Y \mid p_{\sim i}) \approx \frac{1}{2N}\sum_{j=1}^{N}(f(A_b^i)_j - f(A)_j)^2$$

They are used for the estimation of S_i and S_{Ti} respectively.

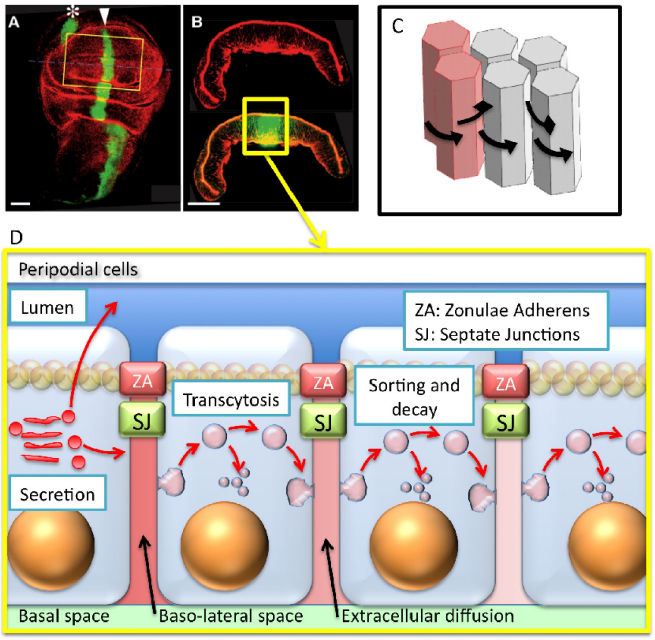


Figure 2.1: Patterning of epithelial cells in the Drosophila wing imaginal disc. The morphogen Dpp patterns the Anterior/Posterior compartments of the Drosophila wing imaginal disc in (A) top view showing the pouch and (B) slice (along dotted line in (A)) showing the geometry of the columnar cells. (C-D) Dpp establishes a non-uniform distribution to pattern the anterior/posterior axis by transport and reaction. Numerous processes may contribute to formation of the Dpp distribution including diffusion around columnar cells (C) or transcytosis through columnar cells (D). Dpp secreted in the basolateral space cannot enter the lumen and vice versa due to the presence of septate junctions (SJ) in (D). From [30] with permission.

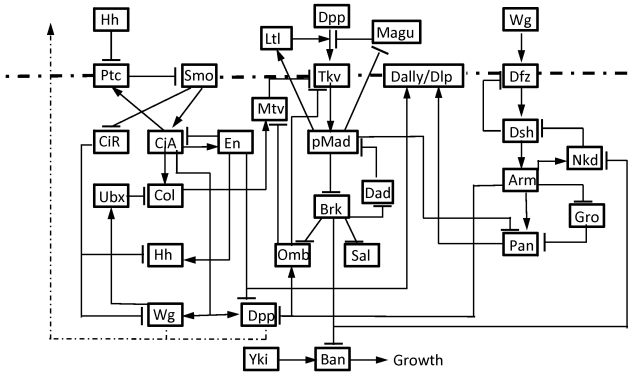


Figure 2.2: The Hh, Dpp and Wg pathways. Ptc: patched, Smo: smoothened, CiA(R): cubitus interruptus activator (repressor), Ubx: ultrabithorax, Col: collier, Dfz: *Drosophila* frizzeled, Dsh: disheveled, Nkd: naked, Gro: groucho, Pan: pangolin. Constructed from [1, 2, 3, 4, 5, 6, 7].

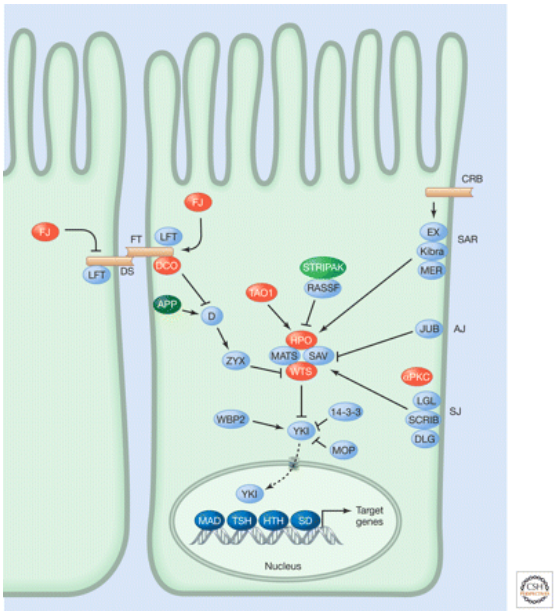


Figure 2.3: The Hippo pathway. App, Approximated; Crb, Crumbs; Dco, Discs overgrown; Dlg, Discs large; Ds, Dachsous; Ex, Expanded; Fj, Four-jointed; Hth, Homothorax; Jub, Drosophila Ajuba; Lats, Large tumor suppressor; Lft, Lowfat; Lgl, Lethal giant larvae; Mer, Merlin; Mats, Mob as a tumor suppressor; Rassf, Ras-associated factor; Sav, Salvador; Scrib, Scribble; Sd, Scalloped; Tsh, Teashirt; Yki, Yorkie [8]

.

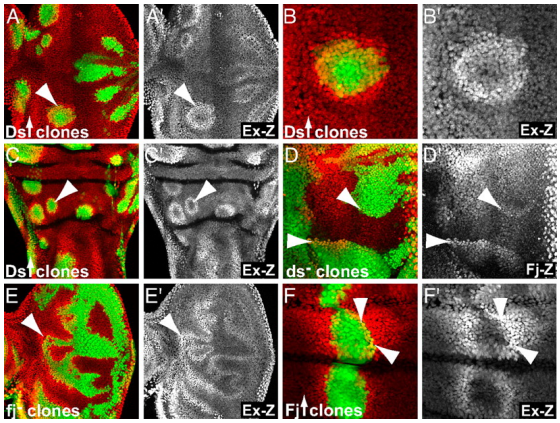


Figure 2.4: Boundaries of Ds and Fj activity induce Hippo target genes. Third instar imaginal discs, which contain clones of cells either overexpressing or mutant for *ds* or *fj* [9].

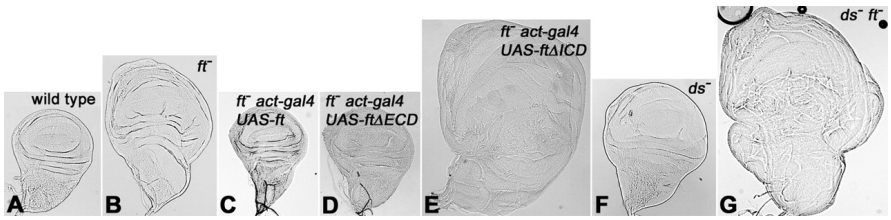


Figure 2.5: Comparison of overgrowth of wing imaginal discs. Typical wing discs for each genotype are shown at the same magnification. (A) Wild type. (B) Overgrowth in *fat* mutant ($ft^{-/-}$). (C) Rescue of $ft^{-/-}$ overgrowth with UAS-ft and act-gal4. (D) Rescue of $ft^{-/-}$ overgrowth with UAS-ftECD and act-gal4. (E) Potentiation of $ft^{-/-}$ overgrowth with UAS-ftICD and act-gal4. (F) Mild overgrowth in *ds* mutant ($ds^{-/-}$). (G) Potentiation of $ft^{-/-}$ overgrowth by $ds^{-/-}$ [10].

Table 2.1: Experiments and phenotypes

Experiments	Phenotypes	Reference
Disc-wide manipulation		
fat$^{-/-}$ mutant	Discs overgrow, cells grow faster, cell size smaller, cell density higher, delay development	[52, 53, 43, 47]
Partial knockout of *fat*	Discs undergrow	[47]
Overexpression of *fat*Δ*ECD* in *fat*$^{-/-}$ background	Reduce the overgrowth of *fat*$^{-/-}$ mutant	[10]
Overexpression of *fat*Δ*ICD* in *fat*$^{-/-}$ background	Enhance the overgrowth of *fat*$^{-/-}$ mutant	[10]
ds$^{-/-}$mutant	Enlarge wing and wing disc	[48]
Overexpression of *ds* using *tub-Gal4* driver	Reduce wing size	[48, 49]
Overexpression of *ds* using *en-Gal4* driver in posterior compartment	Reduce compartment size	[49]
Overexpression of *ds* using *hh-Gal4* driver in posterior compartment	Overexpression of target genes	[44]
Overexpression of *ds*Δ*ECD*	Very weak effect on growth	[51]
Overexpression of *ds*Δ*ECD* in *ds*$^{-/-}$ background	Overexpression of target genes	[51]
fj$^{-/-}$ mutant	Reduce wing size	[48, 49]
Overexpression of *fj*	Reduce wing size	[48, 49]
Co-overexpression of *fj*and *ds*	Greater reduction in wing size than overexpression of either of them	[48]
fat$^{-/-}$, *ds*$^{-/-}$ double mutant	Discs overgrow more than either of single mutants	[10]
Cell clone experiments		
Overexpression of *ds* or *ds*Δ*ICD* in clones	Overexpression of target genes on both sides of the border	[49, 9]
Overexpression of *ds* in clones in either *fat*$^{-/-}$ or *dachs*$^{-/-}$ background	No boundary effects	[9]
ds$^{-/-}$mutant clones	Overexpression of target genes only in neighbor wild-type cells but not in clone cells	[9]
Overexpression of *ds* in clones in *ds*$^{-/-}$ background	Overexpression of target genes only in clone cells but not in neighbor cells lacking of Ds	[9]
Overexpression of *ds*Δ*ICD* in clones in *ds*$^{-/-}$ background	No boundary effects	[9]
Overexpression of *ds* using *hh-Gal4* driver in posterior compartment	Overexpression of target genes both autonomously and non-autonomously	[44]
Overexpression of *ds* using *hh-Gal4* driver together with the depletion of Riq or Mnb in posterior compartment	The autonomous effect is weakened, and the non-autonomous effect is unaffected.	[44]
Overexpression of *fat* or *fat*Δ*ICD* in posterior compartment	Overexpression of target genes in neighbor wild-type cells	[51]
Overexpression of *fat* or *fat*Δ*ICD* in posterior compartment in either *fat*$^{-/-}$ or *ds*$^{-/-}$ background	No boundary effects	[51]
Overexpression of *fat*Δ*ICD* in posterior compartment in posterior compartment	Overexpression of target genes both autonomously and non-autonomously	[51]
Overexpression of *fat*Δ*ICD* in posterior compartment in posterior compartment in *ds*$^{-/-}$ background	The autonomous effect is weakened, and the non-autonomous effect is gone	[51]
Overexpression of *fat*Δ*ICD* in posterior compartment in posterior compartment in *ds*$^{-/-}$ *fat*$^{-/-}$double mutant background	Both of the autonomous and non-autonomous effects are gone	[51]
Overexpression of *fj* in cell clones	Overexpression of target genes on both sides of the border	[9]
fj $^{-/-}$ clones	Overexpression of target genes on both sides of the border	[9]

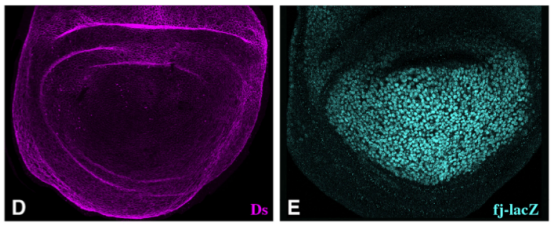


Figure 2.6: Ds and Fj expression in wing disc. (Left) Ds protein staining in a wild-type wing disc; (Right) Fj expression in wild-type revealed by a *fj-lacZ* transgene [11].

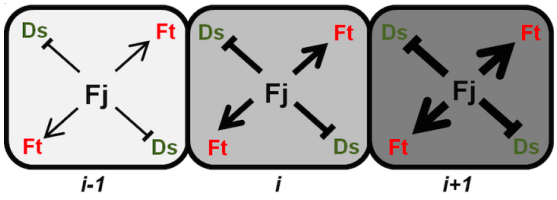


Figure 2.7: Arrangement of cells seen in the contexts of proximally-distally increasing gradient of Fj as well as its activity [12].

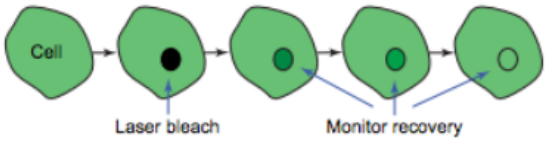


Figure 2.8: Schematic illustrating a FRAP experiment [13].

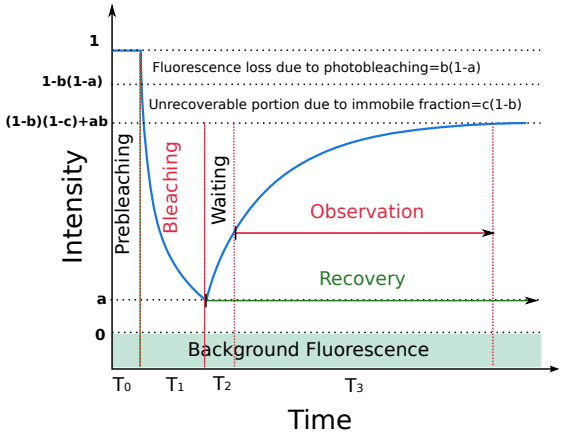


Figure 2.9: A typical recovery curve for a FRAP experiment. After subtraction of the background fluorescence and correction of the observed photobleaching, a FRAP recovery curve is normalized by the fluorescent intensity before bleaching. See text for explanation of the symbols.

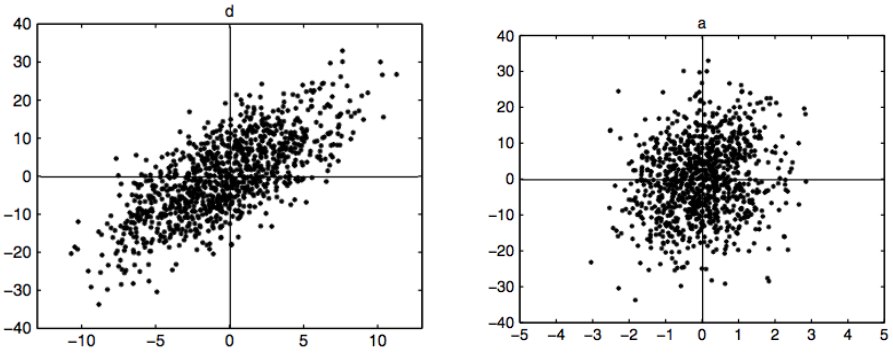


Figure 2.10: Scatterplots of output versus input 1 and input 2

Chapter 3

The Hippo pathway

The objective of this chapter is to develop a mechanistic model of the Hippo pathway to systematically explore the properties and behaviors of this network. We structure the chapter in the following way.

First, we introduce how the model is formulated by choosing the components that primarily contribute to phenotypes observed in experiments, and the kinetic processes that properly capture the interactions between components. Although variations in network structure are explored in the course of model assembly, only the final configuration is described here. It represents the most compact description of the network that captures the network behaviors as realistically as possible. From there, we justify model assumptions and generate hypotheses based on phenotypes and biochemical evidence in the literature. We then simulate the model, analyze the steady state behaviors and demonstrate the validity of the model by comparing model predictions and experimental data. Finally, we analyze the sensitivity of the model with respect to the parameter uncertainties from local and global points of views. We also present the computational analysis of two general models embedded in the network, which could be generalized to help analyze similar modules in other systems.

3.1 The Model

3.1.1 Model philosophy and introduction

To approach the model, we thoroughly evaluated the existing regulatory steps within the network and experimental data available, and analyzed the balances by incorporating the primary genes and proteins, and their interactions, with the goal of understanding whether the structure of the known network could predict experimentally-observed outcomes. In fact, this is how an earlier version of the model was first built with two hypotheses. The qualitative analysis of our model predicted the necessity of two hypotheses: a pathway from Ds that activates Yki, and a pathway from Ds, which is independent of Fat, that represses Yki. Both are needed to explain the results. While searching for a biochemical basis for such pathways, a paper describing one of the two hypotheses, the positive pathway, appeared.

Given the complexity of the network, we do not attempt to incorporate all species and their interactions in the model, but rather retain the central components. Our objective is to develop a detailed, but manageable model which involves selecting a limited set of the protein components that participate in signaling and represent the aggregate effect of upstream components. The major species that are incorporated in the model are Fat, Ds, Dachs, Riq, Wts, and Yki, which are subject to constitutive production rates, and the complexes formed between them. Although there are only 6 distinct species, the number of emerging species complexes due to binding reactions and phosphorylation is rather large. The notations for all the species in the model are summarized in Table 3.1. The variables in the model are listed, of which differential equations can be derived directly.

For each cell in the model, the schematic diagram of the signaling pathway in our model is shown in Figure 3.1. All protein-protein interactions are described by a reversible reaction step for the binding and dissociation of complex partners with the kinetic rate constants k^+ and k^- with appropriate subscripts. An irreversible catalytic step describes subsequent protein modification including decay and phosphorylation of Wts. For the phosphorylation of Fat and Ds by Fj, dephosphorylation steps are included.

All species decay via first-order kinetics in the cell with rates denoted as β_X and on

Table 3.1: Species notations in the model

Notation	Protein
Cytosol species	
F_c, Fp_c	Fat and phosphorylated Fat
S_c, Sp_c	Ds and phosphorylated Ds
J_c	Fj
$\overline{JF}_c$	Fj-Fat complex
$\overline{JS}_c$	Fj-Ds complex
R_c	Riq
A_c	Dachs
W_c	Wts
Y_c, Yp_c	Yki and phosphorylated Yki
Membrane species	
F_m, Fp_m	Fat and phosphorylated Fat
S_m, Sp_m	Ds and phosphorylated Ds
$\overline{FS}_m, \overline{FpS}_m, \overline{FSp}_m, \overline{FpSp}_m$	Fat-Ds complex
A_m	Dachs
$\overline{AS}_m, \overline{ASp}_m$	Dachs-Ds complex
$\overline{AFS}_m, \overline{AFpS}_m, \overline{AFSp}_m, \overline{AFpSp}_m$	Dachs-Fat-Ds complex
$\overline{RS}_m, \overline{RSp}_m$	Riq-Ds complex
$\overline{RFS}_m, \overline{RFpS}_m, \overline{RFSp}_m, \overline{RFpSp}_m$	Riq-Fat-Ds complex
$\overline{WA}_m$	Wts-Dachs complex
$\overline{WAS}_m, \overline{WASp}_m$	Wts-Dachs-Ds complex
$\overline{WAFS}_m, \overline{WAFpS}_m, \overline{WAFSp}_m, \overline{WAFpSp}_m$	Wts-Dachs-Fat-Ds complex
$\overline{WRS}_m, \overline{WRSp}_m$	Wts-Riq-Ds complex
$\overline{WRFS}_m, \overline{WRFpS}_m, \overline{WRFSp}_m, \overline{WRFpSp}_m$	Wts-Riq-Fat-Ds complex

the membrane with rates denoted as γ_X. The decay rate of species on the membrane might result from endocytosis or other degradation mechanisms. In addition, as all the Wts-complex species on the cell membrane are either degraded or phosphorylated due to Dachs or Riq, there are no additional decay rates for them. All the species in the cytosol diffuse. The detailed justifications for each reaction are presented in the following.

In the model we consider an array of discrete cells, as indicated schematically in

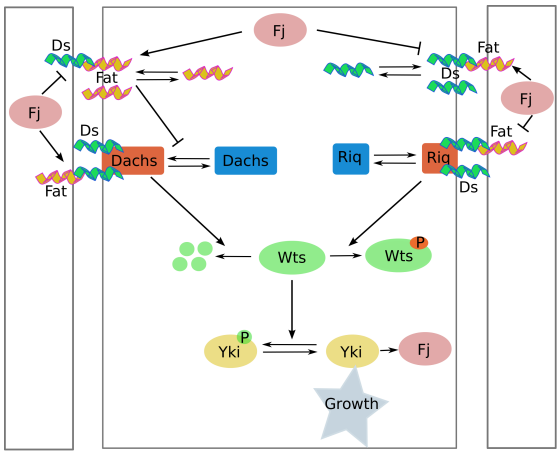


Figure 3.1: Schematic diagram of our model of the cell

Figure 3.2, and we assume that adjacent cells are in close proximity. The model allows reactions and transport within cells, reactions between species at the membrane and either species in the associated cell or on the membrane of adjacent cells, and reactions between species on each of the opposing membranes. The movement of proteins within each cell has been modeled as diffusion, although more complex mechanisms may be involved. We will discuss how the assumptions affect the results later.

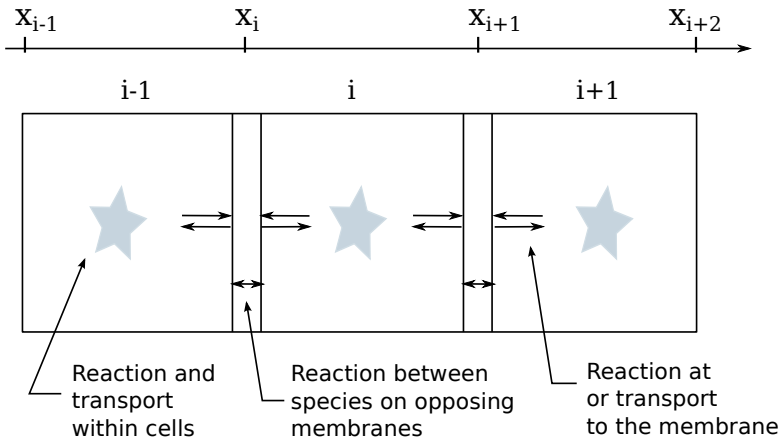


Figure 3.2: A schematic of a 1D network of coupled cells in which cells can interact at their opposing membranes

The reactions involving Fat in the i^{th} cell are binding to the membrane and binding to Ds on either of the adjacent cells. (Hereafter c denotes cytosol, m membrane and i the cell index). The left and right cell membranes of the i^{th} cell are labeled as x_i and x_{i+1}, respectively. As the space between cells is ignored, x_i represents the membrane common to the i^{th} cell and the $(i-1)^{th}$ cell, while x_{i+1} represents the membrane common to the i^{th} cell and the $(i+1)^{th}$ cell. X_c^i denotes the concentration of species X in the cytosol of the i^{th} cell. $X_m^i(x_i)$ and $X_m^i(x_{i+1})$ denote the concentration of species X at the two cell membranes of the i^{th} cell. It should be noted that $\overline{FS}_m^i$ represents the concentration of Fat-Ds complex at the membrane between i^{th} cell and $(i-1)^{th}$ cell, where the Ds part is located on the membrane of the i^{th} cell and the Fat part is located on the membrane of the $(i-1)^{th}$ cell. Therefore, for the i^{th} cell this complex functions as enhanced Ds, while for the $(i-1)^{th}$ cell it functions as enhanced Fat.

3.1.2 Module I: Upstream Intercellular Signaling

The upstream intercellular signaling processes are represented in Figure 3.3, and describe the dynamics of the cadherins Fat and Ds.

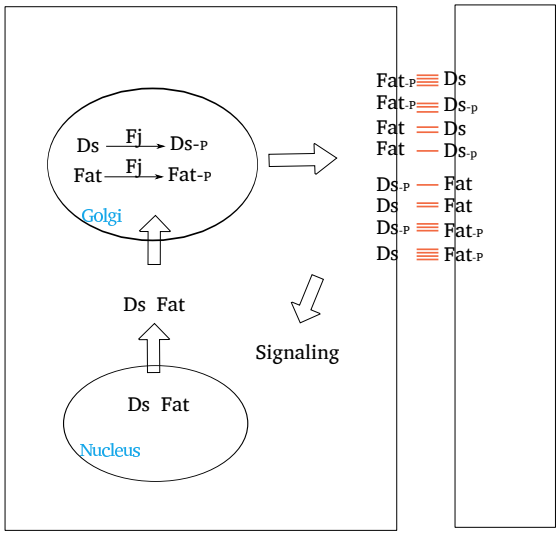


Figure 3.3: Upstream Events: cadherin phosphorylation, membrane localization, and heterodimer formation

Cadherin phosphorylation

The phosphorylation of Fat and Ds catalyzed by the kinase Fj are modeled as enzymatic processes. Dephosphorylation, usually catalyzed by protein phosphatases, is acting in concert with phosphorylation to balance the protein activity in the cell. Here we assume that phosphatases are in abundance, i.e., their concentration does not change significantly over the course of the reaction. Then dephosphorylation processes can be approximated as first-order reactions which depend solely on the concentration of the substrates. Although phosphorylation and dephosphorylation processes happen in the Golgi where Fj is located, we assume that the distribution of Golgi bodies is quite uniform, and hence these processes occur uniformly in cytosol.

$$F_c^i + J_c^i \underset{k_{JF}^-}{\overset{k_{JF}^+}{\rightleftharpoons}} \overline{JF}_c^i \xrightarrow{k_{c1}} Fp_c^i + J_c^i$$

$$Fp_c^i \xrightarrow{k_{c2}} F_c^i$$

$$S_c^i + J_c^i \underset{k_{JS}^-}{\overset{k_{JS}^+}{\rightleftharpoons}} \overline{JS}_c^i \xrightarrow{k_{c3}} Sp_c^i + J_c^i$$

$$Sp_c^i \xrightarrow{k_{c4}} S_c^i$$

Cadherin membrane localization

Both Fat and Ds are transmembrane proteins, thus they are assumed to translocate to the cell membrane reversibly with constant rates. Furthermore, unphosphorylated and phosphorylated forms are assumed to have the same rates, i.e., Fat and Ds have intrinsic rates of binding to the cell membrane, and phosphorylation of Fat and Ds does not change this.

$$F_c^i \, (Fp_c^i) \underset{k_F^-}{\overset{k_F^+}{\rightleftharpoons}} F_m^i \, (Fp_m^i)$$

$$S_c^i \, (Sp_c^i) \underset{k_S^-}{\overset{k_S^+}{\rightleftharpoons}} S_m^i \, (Sp_m^i)$$

Heterodimer (Fat-Ds) formation

Fat and Ds bind with each other heterophilically, and the binding activity is modulated by Fj. Specifically, Fj phosphorylates the extracellular domain of Fat and thus enhances its ability to bind with Ds, and on the other hand, Fj also phosphorylates the extracellular domain of Ds, but decreases its ability to bind with Fat [46]. The opposing effects of Fj on Fat and Ds are represented by four binding affinities of the reactions between unphosphorylated and phosphorylated forms, specifically, the binding affinity between Fat-P and Ds is highest among the possible reactions, and it is much higher than all the other three [46]. In addition, the binding between Fat and Ds promotes the recruitment of each other to cell membrane of adjacent cells. This must be taken into account in the parameter selection process.

$$S_m^i\,(F_m^i) + F_m^{i-1,i+1}\,(S_m^{i-1,i+1}) \underset{k_{FS}^-}{\overset{k_{FS}^+}{\rightleftharpoons}} \overline{FS}_m^i\,(\overline{FS}_m^{i-1,i+1})$$

$$S_m^i\,(Fp_m^i) + Fp_m^{i-1,i+1}\,(S_m^{i-1,i+1}) \underset{k_{FS1}^-}{\overset{k_{FS1}^+}{\rightleftharpoons}} \overline{FpS}_m^i\,(\overline{FpS}_m^{i-1,i+1})$$

$$Sp_m^i\,(F_m^i) + F_m^{i-1,i+1}\,(Sp_m^{i-1,i+1}) \underset{k_{FS2}^-}{\overset{k_{FS2}^+}{\rightleftharpoons}} \overline{FSp}_m^i\,(\overline{FSp}_m^{i-1,i+1})$$

$$Sp_m^i\,(Fp_m^i) + Fp_m^{i-1,i+1}\,(Sp_m^{i-1,i+1}) \underset{k_{FS3}^-}{\overset{k_{FS3}^+}{\rightleftharpoons}} \overline{FpSp}_m^i\,(\overline{FpSp}_m^{i-1,i+1})$$

Studies have shown that the binding process between Fat and Ds (heterodimer formation) modulates the phosphorylation of the cytoplasmic/intracellular domain of each other and hence potentiates their intracellular signaling abilities [48, 11]. In order to represent the effect of phosphorylation of the intracellular domains caused by the heterodimer formation process without further complicating the model, we assume that all the Fat-Ds complex forms have a stronger ability to affect downstream factors than their single forms without introducing another phosphorylated form for each of them. This assumption is valid when the phosphorylation of the intracellular domain is fast enough and thereby the concentration of phosphorylated form is proportional to the total concentration for each species.

It should be emphasized once more that the phosphorylated forms of Fat and Ds in our model refer to the phosphorylation of the extracellular domains caused by Fj, which is different from phosphorylation induced due to heterodimer formation; the former happens to the extracellular domains of Fat and Ds which changes the strength of binding between each other, whereas the latter affects the intracellular regions of Fat and Ds which change their own potency to signal downstream.

3.1.3 Module II: Intermediate Signal Transduction

Dachs localization

Dachs is a major determinant in Fat signal transduction, as *dachs* mutants show the opposite phenotype of *fat* mutants. Although experiments suggest that Fat inhibits Dachs apical localization and the palmitoyltransferase App can relieve Fat's inhibition of Dachs, it is fair to assume a constant background of App when exploring the effect of Fat on Dachs. It should also be noted that Dachs protein is strongly polarized in normal disc cells. However, it has been suggested that the direction in which Dachs localization is polarized influences planar cell polarity (PCP), whereas the amount of Dachs on the membrane influences Hippo signaling [84].

As Fat, Ds and Dachs tend to accumulate at apical junctions of cell membrane, the pools at different junctions are decoupled from each other, and are connected via molecules transported between junctions. At each junction, Dachs localization is assumed to be sensitive to Fat and Ds localized at that junction. .

$$A_c^i \underset{k_A^-}{\overset{k_A^+}{\rightleftharpoons}} A_m^i$$

Since Fat on the cell membrane hinders Dachs localization with an undefined mechanism, this inhibitory effect is modeled as a reduction in the forward binding rates, which is represented as a decreasing Hill function of the concentration of all forms of Fat and Fat complex given by

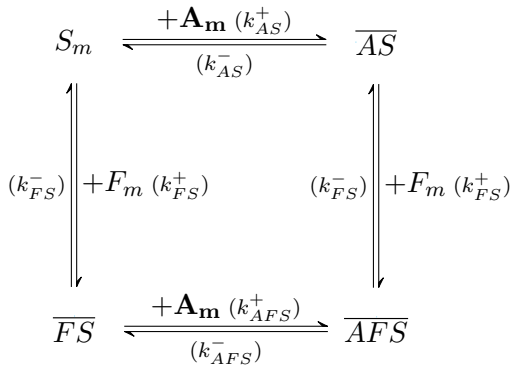
$$k_A^+ = \frac{V_m}{[1 + (\frac{F_m + Fp_m}{K_1})^{n_1}][1 + (\frac{\overline{FS}_{tot}}{K_2})^{n_2}]}$$

where $\overline{FS}_{tot}$ is the sum concentration of all the forms of Fat complex, specifically

$$\begin{aligned}\overline{FS}_{tot} = &\overline{FS} + \overline{FpS} + \overline{FSp} + \overline{FpSp} \\ &+ \overline{AFS} + \overline{AFpS} + \overline{AFSp} + \overline{AFpSp} \\ &+ \overline{RFS} + \overline{RFpS} + \overline{RFSp} + \overline{RFpSp} \\ &+ \overline{WAFS} + \overline{WAFpS} + \overline{WAFSp} + \overline{WAFpSp} \\ &+ \overline{WRFS} + \overline{WRFpS} + \overline{WRFSp} + \overline{WRFpSp}.\end{aligned}$$

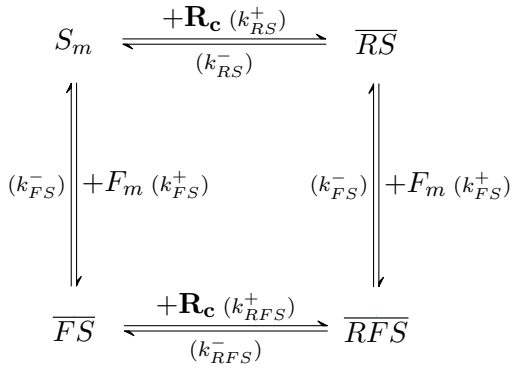
In the function, n_1 and n_2 are used to approximate the non-linearity in the inhibitory effect, and are set to be equal in our simulations. K_1 and K_2 are used to separate the inhibitory effect by the Fat complex from that by single forms of Fat.

As we discussed before, the fact that *fat* and *ds* double mutants overgrow more than either of single mutants indicates a negative regulatory step from Ds that is independent of Fat. Meanwhile, the evidence that Dachs binds to the intracellular domain of Ds observed in *Drosophila* scutellum cells [25] may account for this type of regulation, which leads to our first hypothesis: localized Dachs can bind to Ds to form a complex which decays faster than Dachs alone on the cell membrane, probably via speeding up some internalization processes. With this hypothesis, Ds has been assigned a role of negatively regulating growth via promoting the decay of Dachs on the membrane. Furthermore, the interaction between Fat and Ds usually leads to incremental change rather than radical change in each other's signaling potency. Therefore, we assume that Ds complex also can bind to Dachs with higher affinity.



Riq localization

Experiments show that Ds is required to recruit Riq to apical junctions, and Fat on adjacent cells binding to Ds promotes the recruitment [47]. Hence we assume that Dachs on the membrane binds to Riq in the cytosol directly, in contrast to the assumption that Ds binds to Dachs after Dachs is localized to membrane. The interactions between species are represented as binding to form complexes. The effect of Fat on Riq localization via Ds is represented by the higher binding affinity between Fat-Ds complex and Riq than that between Ds and Riq.



3.1.4 Module III: Downstream Effectors

Wts decay

Localized Dachs can bind to Wts and promote the turnover of Wts. We assume Dachs on the membrane can recruit cytosolic Wts to membrane directly

$$W^i_c + \mathbf{A^i_m} \underset{k^-_{WA}}{\overset{k^+_{WA}}{\rightleftharpoons}} \overline{WA}^i_m \xrightarrow{k_d} \phi + \mathbf{A^i_m}$$

and similarly for $\overline{AS}_m$, $\overline{ASp}_m$,$\overline{AFS}_m$,$\overline{AFpS}_m$,$\overline{AFSp}_m$,$\overline{AFpSp}_m$.

Wts phosphorylation

Localized Riq can bind to Wts and promote its phosphorylation which inhibits Wts activity. This occurs through the interaction of Riq with Mnb to potentiate the ability of

Mnb to phosphorylate Wts. However, since our aim is to analyze Fat/Ds induced Hippo pathway behaviors, it is appropriate to lump two successive reactions and represent the ability of Riq-Mnb complex to phosphorylate Wts by the concentration of Riq. Whether these proteins interact directly, or if they interact through a series of intermediate proteins is not expected to alter the nature of this model.

It is worth noting that there are multiple phosphorylation events on Warts kinases. Expanded(Ex), a cytoplasmic protein that suppresses growth via the Hippo pathway, has been suggested to regulate Yki activity [85, 86, 87]. Ex phosphorylates Wts on the activation loop and hydrophobic motif, which is an activating phosphorylation. Meanwhile, Riq/Mnb phosphorylates Wts on independent sites, which is a repressive phosphorylation. However, experiments show that overexpression of Ex alone has no effect on Wts mobility (coexpression of Ex and Merline(Mer) enhances the effect on Wts mobility induced by overexpression of Mer alone [88]). Since there is no direct interaction between Fat/Ds signaling and Mer, the model does not include Ex and Mer. Thus we represent the effect as

$$W_c^i + \overline{\mathbf{RS}}_{\mathbf{m}}^{\mathbf{i}} \underset{k_{WR}^-}{\overset{k_{WR}^+}{\rightleftharpoons}} \overline{WRS}_m^i \xrightarrow{k_p} \varphi + \overline{\mathbf{RS}}_{\mathbf{m}}^{\mathbf{i}},$$

and similarly for $\overline{RSp}_m$, $\overline{RFS}_m$, $\overline{RFpS}_m$, $\overline{RFSp}_m$,$\overline{RFpSp}_m$.

Yki phosphorylation

Active Wts binds to Yki in the cytosol to promote Yki phosphorylation, which prevents Yki from entering the nucleus and hence inhibits its activity.

$$W_c^i + Y_c^i \underset{k_{WY}^-}{\overset{k_{WY}^+}{\rightleftharpoons}} \overline{WY}_c^i \xrightarrow{k_{py}} W_c^i + Yp_c^i$$

In normal growth, Yki nuclear localization is not detectable, which suggests that Yki is a potent growth regulator, and many players inhibit its localization in normal development. Therefore, in the model, Y_c should be low at the steady state under normal conditions.

3.1.5 The governing equations

All the species in the cytosol are assumed to diffuse freely and isotropically within the cell interior. For each of them, the boundary conditions for two edges are symmetric, i.e, fluxes with opposite signs. For simplicity, only one of the two boundary conditions is listed.

We assume that all the species on cell membrane are immobile. For each of them, there are two governing equations describing its dynamics on two edges of a cell. For simplicity, only one of the two equations for each species is listed.

Species in the cytosol

The equations for unphosphorylated Fat and Ds share the same forms with different kinetic parameters. Consider unphosphorylated Fat as an example:

$$\frac{\partial F_c}{\partial t} = D_F \frac{\partial^2 F_c}{\partial x^2} + \alpha_F - \beta_F F_c - k^+_{JF} F_c J_c + k^-_{JF} \overline{JF}_c + k_{c2} F p_c \qquad \text{in } \Omega \tag{3.1}$$

$$-\frac{D_F}{\varepsilon} \frac{\partial F_c}{\partial x}|_{\partial\Omega} = -k^+_F F_c + k^-_F F_m$$

[1]

The equations for phosphorylated Fat and Ds share the same form with different kinetic parameters. Consider phosphorylated Fat as an example:

$$\frac{\partial F p_c}{\partial t} = D_{Fp} \frac{\partial^2 F p_c}{\partial x^2} - \beta_F F p_c + k_{c1} \overline{JF}_c - k_{c2} F p_c \qquad \text{in } \Omega \tag{3.2}$$

$$-\frac{D_{Fp}}{\varepsilon} \frac{\partial F p_c}{\partial x}|_{\partial\Omega} = -k^+_F F p_c + k^-_F F p_m$$

Fj concentration:

$$\frac{\partial J_c}{\partial t} = D_J \frac{\partial^2 J_c}{\partial x^2} + \alpha_J - \beta_J J_c - k^+_{JF} F_c J_c + (k^-_{JF} + k_{c1}) \overline{JF}_c - k^+_{JS} S_c J_c + (k^-_{JS} + k_{c3}) \overline{JS}_c \qquad \text{in } \Omega \tag{3.3}$$

$$-D_J \frac{\partial J_c}{\partial x}|_{\partial\Omega} = 0$$

[1] The factor ε reflects the conversion to volume and is assumed to be the thickness of the cell membrane.

The equations for Fj-Fat complex and Fj-Ds complex share the same form with different kinetic parameters. Consider Fj-Fat complex as an example:

$$\frac{\partial \overline{JF}_c}{\partial t} = D_J \frac{\partial^2 \overline{JF}_c}{\partial x^2} + k^+_{JF} F_c J_c - (k^-_{JF} + k_{c1})\overline{JF}_c \tag{3.4}$$

$$-D_{JF}\frac{\partial \overline{JF}_c}{\partial x}|_{\partial\Omega} = 0$$

Dachs concentration:

$$\frac{\partial A_c}{\partial t} = D_A \frac{\partial^2 A_c}{\partial x^2} + \alpha_A - \beta_A A_c \qquad \text{in } \Omega \tag{3.5}$$

$$-\frac{D_A}{\varepsilon}\frac{\partial A_c}{\partial x}|_{\partial\Omega} = -k^+_A A_c + k^-_A A_m$$

Note that k^+_A depends on Fat concentration on cell membrane.

Riq concentration:

$$\frac{\partial R_c}{\partial t} = D_R \frac{\partial^2 R_c}{\partial x^2} + \alpha_R - \beta_R R_c \qquad \text{in } \Omega \tag{3.6}$$

$$\begin{aligned}-\frac{D_R}{\varepsilon}\frac{\partial R_c}{\partial x}|_{\partial\Omega} = &-k^+_{RS}(S_m + Sp_m)R_c + k^-_{RS}(\overline{RS}_m + \overline{RSp}_m)\\ &- k^+_{RFS}R_c(\overline{FS}_m + \overline{FpS}_m + \overline{FSp}_m + \overline{FpSp}_m)\\ &+ k^-_{RFS}(\overline{RFS}_m + \overline{RFpS}_m + \overline{RFSp}_m + \overline{RFpSp}_m)\end{aligned}$$

Wts concentration:

$$\frac{\partial W_c}{\partial t} = D_W \frac{\partial^2 W_c}{\partial x^2} + \alpha_W - \beta_W W_c - k^+_{WY} W_c Y_c + (k^-_{WY} + k_{py})\overline{WY}_c \qquad \text{in } \Omega \tag{3.7}$$

$$\begin{aligned}-\frac{D_W}{\varepsilon}\frac{\partial W_c}{\partial x}|_{\partial\Omega} = &-k^+_{WA}A_m W_c + k^-_{WA}\overline{WA}_m\\ &- k^+_{WR}W_c(\overline{RS}_m + \overline{RSp}_m) + k^-_{WR}(\overline{WRS}_m + \overline{WRSp}_m)\\ &- k^+_{WR}W_c(\overline{RFS}_m + \overline{RFpS}_m + \overline{RFSp}_m + \overline{RFpSp}_m)\\ &+ k^-_{WR}(\overline{WRFS}_m + \overline{WRFpS}_m + \overline{WRFSp}_m + \overline{WRFpSp}_m)\\ &- k^+_{WA}W_c(\overline{AS}_m + \overline{ASp}_m) + k^-_{WA}(\overline{WAS}_m + \overline{WASp}_m)\\ &- k^+_{WA}W_c(\overline{AFS}_m + \overline{AFpS}_m + \overline{AFSp}_m + \overline{AFpSp}_m)\\ &+ k^-_{WA}(\overline{WAFS}_m + \overline{WAFpS}_m + \overline{WAFSp}_m + \overline{WAFpSp}_m)\end{aligned}$$

Yki concentration:

$$\frac{\partial Y_c}{\partial t} = D_Y \frac{\partial^2 Y_c}{\partial x^2} + \alpha_Y - \beta_Y Y_c - k^+_{WY} W_c Y_c + k^-_{WY} \overline{WY}_c \qquad \text{in } \Omega \tag{3.8}$$

$$-D_Y \frac{\partial Y_c}{\partial x}|_{\partial\Omega} = 0$$

Phosphorylated Yki concentration:

$$\frac{\partial Yp_c}{\partial t} = D_{YP} \frac{\partial^2 Yp_c}{\partial x^2} + k_{py} \overline{WY}_c - \beta_{Yp} Yp_c \qquad \text{in } \Omega \tag{3.9}$$

$$-D_{YP} \frac{\partial Yp_c}{\partial x}|_{\partial\Omega} = 0$$

Wts-Yki concentration:

$$\frac{\partial \overline{WY}_c}{\partial t} = D_{WY} \frac{\partial^2 \overline{WY}_c}{\partial x^2} + k^+_{WY} W_c Y_c - (k^-_{WY} + k_{py}) \overline{WY}_c - \beta_{WY} \overline{WY}_c \qquad \text{in } \Omega \tag{3.10}$$

$$-D_{WY} \frac{\partial \overline{WY}_c}{\partial x}|_{\partial\Omega} = 0$$

Species on the cell membrane

The equations for unphosphorylated Fat and phosphorylated Fat on cell membrane share the same form with different kinetic parameters. Consider unphosphorylated Fat as an example:

$$\begin{aligned}
\frac{dF^i_m}{dt} &= k^+_F F^i_c - k^-_F F^i_m - \gamma_F F^i_m \\
&\quad - k^+_{FS} F^i_m (S^{i-1}_m + \overline{RS}^{i-1}_m + \overline{WRS}^{i-1}_m + \overline{AS}^{i-1}_m + \overline{WAS}^{i-1}_m) \\
&\quad + k^-_{FS} (\overline{FS}^{i-1}_m + \overline{RFS}^{i-1}_m + \overline{WRFS}^{i-1}_m + \overline{AFS}^{i-1}_m + \overline{WAFS}^{i-1}_m) \qquad \text{on } \partial\Omega \\
&\quad - k^+_{FS2} F^i_m (Sp^{i-1}_m + \overline{RSp}^{i-1}_m + \overline{WRSp}^{i-1}_m + \overline{ASp}^{i-1}_m + \overline{WASp}^{i-1}_m) \\
&\quad + k^-_{FS2} (\overline{FSp}^{i-1}_m + \overline{RFSp}^{i-1}_m + \overline{WRFSp}^{i-1}_m + \overline{AFSp}^{i-1}_m + \overline{WAFSp}^{i-1}_m)
\end{aligned} \tag{3.11}$$

The equations for unphosphorylated Ds and phosphorylated Ds on the membrane share the same form with different kinetic parameters. Consider unphosphorylated Ds as an example:

$$\begin{aligned}
\frac{dS^i_m}{dt} &= k^+_S S^i_c - k^-_S S^i_m - \gamma_S S^i_m \\
&\quad - k^+_{FS} S^i_m F^{i-1}_m + k^-_{FS} \overline{FS}^i_m - k^+_{FS1} S^i_m Fp^{i-1}_m + k^-_{FS1} \overline{FpS}^i_m \\
&\quad - k^+_{RS} S^i_m R^i_c + k^-_{RS} \overline{RS}^i_m - k^+_{AS} S^i_m A^i_m + k^-_{AS} \overline{AS}^i_m
\end{aligned} \tag{3.12}$$

The equations for the 4 types of Fat-Ds complex formed on cell membrane, $\overline{FS}, \overline{FpS}, \overline{FSp}, \overline{FpSp}$, share the same form with different kinetic parameters. Consider $\overline{FS}$ as an example:

$$\begin{aligned}\frac{d\overline{FS}^i_m}{dt} &= k^+_{FS}F^{i-1}_m S^i_m - k^-_{FS}\overline{FS}^i_m - \gamma_{FS}\overline{FS}^i_m \\ &\quad - k^+_{RFS}\overline{FS}^i_m R^i_c + k^-_{RFS}\overline{RFS}^i_m - k^+_{AFS}\overline{FS}^i_m A^i_m + k^-_{AFS}\overline{AFS}^i_m\end{aligned} \tag{3.13}$$

Dachs concentration:

$$\begin{aligned}\frac{dA_m}{dt} &= k^+_A A_c - k^-_A A_m - \gamma_A A_m \\ &\quad - k^+_{AS}A_m(S_m + Sp_m) + k^-_{AS}(\overline{AS}_m + \overline{ASp}_m) \\ &\quad - k^+_{AFS}A_m(\overline{FS}_m + \overline{FpS}_m + \overline{FSp}_m + \overline{FpSp}_m) \\ &\quad + k^-_{AFS}(\overline{AFS}_m + \overline{AFpS}_m + \overline{AFSp}_m + \overline{AFpSp}_m) \\ &\quad - k^+_{WA}A_m W_c + (k^-_{WA} + k_d)\overline{WA}_m\end{aligned} \tag{3.14}$$

The equations for the 4 types of Dachs-Ds complex formed on cell membrane, $\overline{AS}, \overline{ASp}$, share the same form with different kinetic parameters. Consider $\overline{AS}$ as an example:

$$\begin{aligned}\frac{d\overline{AS}^i_m}{dt} &= k^+_{AS}S^i_m A^i_m - k^-_{AS}\overline{AS}^i_m - k^+_{WA}W^i_c\overline{AS}^i_m + (k^-_{WA} + k_d)\overline{WAS}^i_m \\ &\quad - k^+_{FS}\overline{AS}^i_m F^{i-1}_m + k^-_{FS}\overline{AFS}^i_m - k^+_{FS1}\overline{AS}^i_m Fp^{i-1}_m + k^-_{FS1}\overline{AFpS}^i_m - \gamma_{AS}\overline{AS}^i_m\end{aligned} \tag{3.15}$$

The equations for the 4 types of Dachs-Fat-Ds complex formed on cell membrane, $\overline{AFS}, \overline{AFpS}, \overline{AFSp}, \overline{AFpSp}$, share the same form with different kinetic parameters. Consider $\overline{AFS}$ as an example:

$$\begin{aligned}\frac{d\overline{AFS}^i_m}{dt} &= k^+_{AFS}A^i_m\overline{FS}^i_m - k^-_{AFS}\overline{AFS}^i_m - k^+_{WA}W^i_c\overline{AFS}^i_m + (k^-_{WA} + k_d)\overline{WAFS}^i_m \\ &\quad + k^+_{FS}\overline{AS}^i_m F^{i-1}_m - k^-_{FS}\overline{AFS}^i_m - \gamma_{AFS}\overline{AFS}^i_m\end{aligned} \tag{3.16}$$

The equations for the 2 types of Riq-Ds complex formed on cell membrane, $\overline{RS}, \overline{RSp}$, share the same form with different kinetic parameters. Consider $\overline{RS}$ as an example:

$$\begin{aligned}\frac{d\overline{RS}^i_m}{dt} &= k^+_{RS}S^i_m R^i_c - k^-_{RS}\overline{RS}^i_m - k^+_{WR}W^i_c\overline{RS}^i_m + (k^-_{WR} + k_p)\overline{WRS}^i_m \\ &\quad - k^+_{FS}\overline{RS}^i_m F^{i-1}_m + k^-_{FS}\overline{RFS}^i_m - k^+_{FS1}\overline{RS}^i_m Fp^{i-1}_m + k^-_{FS1}\overline{RFpS}^i_m - \gamma_{RS}\overline{RS}^i_m\end{aligned} \tag{3.17}$$

The equations for the 4 types of Riq-Fat-Ds complex formed on cell membrane, $\overline{RFS}$, $\overline{RFpS}$, $\overline{RFSp}$,$\overline{RFpSp}$, share the same form with different kinetic parameters. Consider $\overline{RFS}$ as an example:

$$\begin{aligned}\frac{d\overline{RFS}_m^i}{dt} &= k_{RFS}^+ R_c^i \overline{FS}_m^i - k_{RFS}^- \overline{RFS}_m^i - k_{WR}^+ W_c^i \overline{RFS}_m^i + (k_{WR}^- + k_p)\overline{WRFS}_m^i \\ &\quad + k_{FS}^+ \overline{RS}_m^i F_m^{i-1} - k_{FS}^- \overline{RFS}_m^i - \gamma_{RFS}\overline{RFS}_m^i\end{aligned} \tag{3.18}$$

Wts-Dachs complex concentration:

$$\frac{d\overline{WA}_m}{dt} = k_{WA}^+ A_m W_c - (k_{WA}^- + k_d)\overline{WA}_m \tag{3.19}$$

The equations for the 2 types of Wts-Dachs-Ds complex formed on cell membrane, $\overline{WAS}, \overline{WASp}$, share the same form with different kinetic parameters. Consider $\overline{WAS}$ as an example:

$$\begin{aligned}\frac{d\overline{WAS}_m^i}{dt} &= k_{WA}^+ W_c^i \overline{AS}_m^i - (k_{WA}^- + k_d)\overline{WAS}_m^i \\ &\quad - k_{FS}^+ \overline{WAS}_m^i F_m^{i-1} + k_{FS}^- \overline{WAFS}_m^i - k_{FS1}^+ \overline{WAS}_m^i F p_m^{i-1} + k_{FS1}^- \overline{WAFpS}_m^i\end{aligned} \tag{3.20}$$

The equations for the 4 types of Wts-Dachs-Fat-Ds complex formed on cell membrane, $\overline{WAFS}$, $\overline{WAFpS}$, $\overline{WAFSp}$, $\overline{WAFpSp}$, share the same form with different kinetic parameters. Consider $\overline{WAFS}$ as an example:

$$\frac{d\overline{WAFS}_m^i}{dt} = k_{WA}^+ W_c^i \overline{AFS}_m^i - (k_{WA}^- + k_d)\overline{WAFS}_m^i + k_{FS}^+ \overline{WAS}_m^i F_m^{i-1} - k_{FS}^- \overline{WAFS}_m^i \tag{3.21}$$

The equations for the 2 types of Wts-Riq-Ds complex formed on cell membrane, $\overline{WRS}$, $\overline{WRSp}$, share the same form with different kinetic parameters. Consider $\overline{WRS}$ as an example:

$$\begin{aligned}\frac{d\overline{WRS}_m^i}{dt} &= k_{WR}^+ W_c^i \overline{RS}_m^i - (k_{WR}^- + k_p)\overline{WRS}_m^i \\ &\quad - k_{FS}^+ \overline{WRS}_m^i F_m^{i-1} + k_{FS}^- \overline{WRFS}_m^i - k_{FS1}^+ \overline{WRS}_m^i F p_m^{i-1} + k_{FS1}^- \overline{WRFpS}_m^i\end{aligned} \tag{3.22}$$

The equations for the 4 types of Wts-Riq-Fat-Ds complex formed on cell membrane,

$\overline{WRFS}$, $\overline{WRFpS}$, $\overline{WRFSp}$, $\overline{WRFpSp}$, share the same form with different kinetic parameters. Consider $\overline{WRFS}$ as an example:

$$\frac{d\overline{WRFS}^i_m}{dt} = k^+_{WR} W^i_c \overline{RFS}^i_m - (k^-_{WR} + k_p)\overline{WRFS}^i_m + k^+_{FS}\overline{WRS}^i_m F^{i-1}_m - k^-_{FS}\overline{WRFS}^i_m \tag{3.23}$$

3.2 Parameter selection

Our model is based on partial differential equations (PDEs) and ordinary differential equations (ODEs) and consists of 46 state variables and 40 independent parameters. Compared with the number of state variables, the number of parameters is relatively small. Several general and feasible strategies have been implemented to ensure the minimal number of parameters used in the model. Firstly, the same parameters are used for similar processes unless experimental observations suggest a difference. For instance, the parameters for the binding process between Fat and Ds on adjacent cells are exactly the same with those between Fat and Ds complex, as no evidence shows that the interactions between Ds ICD and species within the cell have an impact on its binding with Fat on adjacent cells. Secondly, experimental data suggests the difference in the binding affinities for specific reactions in the system, and we select the corresponding association and dissociation rates to be in accordance with differential affinities. Lastly, we make use of common thermodynamic properties of molecules. For instance, molecules with large molecular weight move slower if they belong to the similar type of protein and their transport mechanisms are assumed to be the same. Slow movement usually in turn leads to slower binding process.

As most of the kinetic parameters in the Hippo pathway are unknown at present, we explored wide ranges of parameter space to understand the properties of the system. Detailed systematic parameter exploration compatible with experimentally observed behavior constitute an appropriate approach to tackle the lack of measured parameters. The model is insensitive to a large number of parameters when varied independently over 3 orders of magnitude. Most of the system behaviors can be observed in all these wide ranges of parameter choice. The observations are relatively sensitive to some parameters which will be discussed in detail later. All the parameters in the model parameters are listed in Table 3.2.3.

3.2.1 Parameters for binding processes

The binding rates in the model can be categorized into four types based on the types of interactions.

1. The binding rate constant between a cytosolic species and a membrane species with a conventional unit $M^{-1}min^{-1}$: k^+_{WA}, k^+_{WR}, k^+_{RS}, k^+_{RFS}
2. The binding rate constant between two cytosolic species with a conventional unit $M^{-1}min^{-1}$: k^+_{WY}, k^+_{JF}, k^+_{JS}
3. The binding rate constant of a species from cytosol to membrane with a conventional unit $m \cdot min^{-1}$: k^+_F, k^+_S, k^+_A
4. The binding rate constant between two membrane species with a conventional unit $m^2mol^{-1}min^{-1}$: k^+_{FS}, k^+_{FS1}, k^+_{FS2}, k^+_{FS3}, k^+_{AS}, k^+_{AFS}

Here M is the unit of concentration, and m is the unit of length.

For comparison of the magnitudes of the rate constants for different geometries, it is convenient to convert all the rate constants and concentrations to effective three-dimensional rate constants. To do that, we divide two dimensional concentrations by the membrane thickness ($\varepsilon \sim 10nm$) to give an estimated local volume for cell surface components, and convert them to three-dimensional concentrations. We also multiply the two dimensional rate constants (k^+_{FS}, k^+_{FS1}, k^+_{FS2}, k^+_{FS3}, k^+_{AS}, k^+_{AFS}) by the membrane thickness, and divide the binding rate constants from cytosol to membrane (k^+_F, k^+_S, k^+_A) by the membrane thickness. Therefore, all the binding rates and concentrations have the same units hereafter. This also explains why the boundary conditions have to be modified by the membrane thickness.

Knowledge of the rate constants of the forward and backward reactions are required to describe the behaviors of the system. In our model, these rates are restricted to be within the typical ranges for protein-protein interactions. For example, the association of protein molecules into dimers for larger complexes occurs in typical rate constants of the order of 10^{-3} to 10 $nM^{-1}min^{-1}$. The equilibrium dissociation constant K_D for ligand-receptor binding process has the typical range $0.1 \sim 10^3 nM$ with the range of association rate constant $10^{-3} \sim 10nM^{-1}min^{-1}$ and of the dissociation rate constant $10^{-3} \sim 1min^{-1}$ [89].

In addition, all the reaction rates are constrained to be no faster than predicted by the diffusion limit. The coupling reactions between two membrane species are usually diffusion-limited. The diffusion on cell membrane is much slower than that in cytosol, which means that the limit of binding rate constant between membrane species is much smaller, typically 2 orders of magnitude smaller. However, studies have shown that membrane localization can amplify the number of complexes that are formed between the signaling partners. When bound molecules are confined to a small volume, it results in 100-1000 fold increase in the apparent affinity [90]. Fat and Ds are localized to AJs, and therefore, it is possible that the binding between them can be much faster than normal binding between two membrane species.

In addition, the relative binding rates are assigned to be consistent with experimental observations. For example, as experiments show that Fat binding to Ds enhances the interaction between Ds and Riq, a higher forward rate is assigned for the binding process between Fat-Ds complex and Riq in comparison with that between Ds and Riq.

3.2.2 Diffusion coefficients

The simplest correlation based on molecular weights, temperature, and viscosity provides some basis for estimation of the diffusion coefficients. The Stokes –Einstein –Sutherland equation for diffusion of spherical particles through a liquid with low Reynolds number is given by

$$D = \frac{k_B T}{6\pi\,\eta\,r}$$

where D is the diffusion constant; k_B is Boltzmann's constant; T is the absolute temperature; η is the dynamic viscosity; r is the radius of the spherical particle.

Although we do not attempt to use this equation to estimate diffusion coefficients in the model, it shows that $D \propto \frac{1}{M}^{1/3}$ since $r \propto M^{1/3}$. Therefore, for similar types of molecules such as cadherins Fat and Ds, their diffusion coefficients maintain the ratio of the cubic root of the ratio between their molecular weights. For different types of molecules, the general assumption that molecules with larger weights move slower still holds.

3.2.3 Production and decay rates

A typical columnar cell in a wing disc at the larval stage can be approximated by a cylinder with a diameter $\sim 4\mu m$ and a height $\sim 20\mu m$, which gives the cell volume $\sim 250\mu m^3$. Receptor synthesis rate $10^2 \sim 10^4 min^{-1}$ then can be converted to $1 \sim 100nMmin^{-1}$. The typical range for receptor degradation rate is $10^{-3} \sim 0.1min^{-1}$ [89].

3.3 Numerical Methods

For non-linear highly coupled PDEs, it is almost impossible to analytically demonstrate that there is only one steady state for the system . However, we did use a variety of initial conditions for simulations to ascertain that all reach the same steady-state. The PDEs are discretized using standard finite difference methods. The numerical errors arising from discretization are well controlled so that they do not alter our simulations.

The periodic boundary conditions are used for the intercellular signaling terms, and the number of cells is large enough, such that the equations for every cell in the model are the same. When simulating patches of mutant cell clones, we choose the size of clones so that they are large enough for variations in patterns between cells to decay before reaching the edge of the simulated array in order to prevent these variations from interfering with each other.

3.4 Computational Results

The Hippo pathway is complicated yet interesting as all the components are highly connected and many of them play multiple roles, and the resulting nonlinear behaviors hamper an intuitive analysis. As opposed to some oversimplified models, our model is unique in the detail with which we model the Fat/Ds signaling via the Hippo pathway. By accounting for the specific interactions of proteins using mass action kinetics, we are able to explain many phenotypes associated with this pathway, including the non-monotonic effects of Fat, Ds and Fj on growth, the non-autonomous effect induced by cell clones, and etc. We find that the nature of this signaling pathway, whether it simulates or inhibits growth, depends on many factors, such as the strength of different

interactions, and the level of enzymatic activities of kinases. Our detailed model incorporates these quantitative properties and makes predictions with reasonable values of parameters.

Our model includes a large system of nonlinear reaction-diffusion equations, and it is important to first understand the structure of the system, which consists of three modules, upstream regulators Fat and Ds, intermediate transducers Dachs and Riq, and downstream effectors Wts and Yki. We analyze each of the modules as well as the communication between them, and predict the observed outcomes under various choices of parameters. The mechanisms of signal transduction within the cell and of signal transport between cells are discussed.

3.4.1 Non-monotonic Response

We begin by modeling single cell behaviors that allows us to explore the interactions of different components in the signaling pathway in the cell. If, for the present, we neglect the concentration gradients of Ds and Fj and assume uniform distributions, which appears to be a reasonable assumption especially in the wing pouch based on experimental observations, all the cells in the disc are identical, i.e., the state of this signaling pathway in each cell is the same and cells respond in the same way to the intervention. This will reflect the behaviors of this signaling pathway in the disc-wide experiments, where the expression of genes are changed in all cells across the entire disc. Under these circumstances, for each cell, the concentration of either Fat or Ds on cell membrane is exactly same with that on its neighbors. The reciprocal binding process between Fat and Ds on adjacent cells can be modeled as within a cell where Fat and Ds bind to each other. The full model is reduced to a single-cell model, which provides a simple and valid way to explore the behaviors of this pathway to disc-wide experiments.

With this reduction, we investigate Fat/Ds signal transduction via the Hippo pathway within the cell. We explore the steady-state concentration of Yki as a function of the production rates of Fat and Ds, and examine how upstream signals are transmitted downstream, resulting in the complex response of cell growth.

The color in Figure 3.4 corresponds to the concentration of the growth signal, Yki, as a function of Fat and Ds production rates, which are uniform inputs into the cell. Although details of this map depend on parameters, it has certain generic features

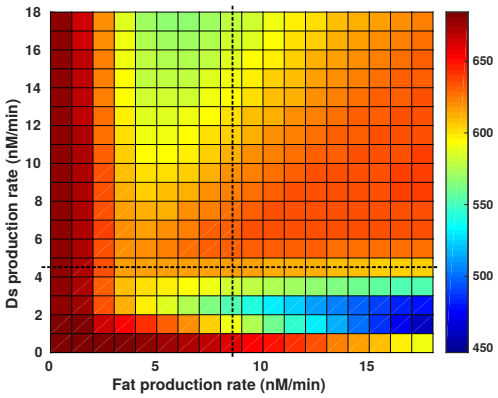


Figure 3.4: Predicted growth response measured through averaged Yki concentration as a function of Fat and Ds expression levels.

that we want to highlight. Slicing horizontally across this signal transduction map, as indicated by the dashed line in the figure generates a predicted growth curve as a function of Fat expression level at a normal Ds background. The curve shown in Figure 3.5 is non-monotonic and predicts that overexpression of Fat causes undergrowth, complete loss of Fat causes overgrowth, and more importantly, partial loss of Fat reduces, rather than enhances, growth. This is in good agreement with experimental observations [47].

Furthermore, the model predicts that although reducing Fat expression level monotonically decreases the level of localized Riq and increases the level of localized Dachs, the interactions between them and the corresponding nonlinearities lead to the opposite outcomes from partial and complete loss of *fat* as shown in Figure 3.6. The decline in the level of localized Riq dominates when a fraction of *fat* is lost, and hence leads to rise in Yki activity. In contrast, when *fat* is completely knocked out, it completely releases the inhibition on Dachs localization and thereby activates Yki activity. The different sensitivities of Dachs and Riq to the Fat expression level follow from the fact that Fat affects Dachs and Riq through distinct mechanisms. In addition, upregulation of Dachs localization in *fat* loss-of-function mutants is primarily due to the loss of Fat-Ds complex, while downregulation of Dachs localization in *fat* overexpression discs is mainly due to the increase of localized Fat on membrane. This reflects the relatively different

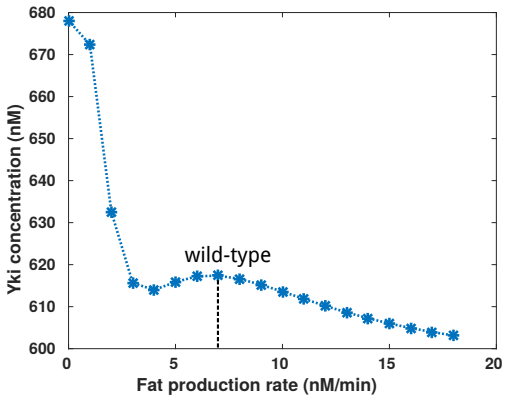


Figure 3.5: Horizontal slice of the growth response map shows a non-monotonic dependence of growth on Fat expression level

inhibitory strength of Fat and Fat-Ds complex on Dachs localization.

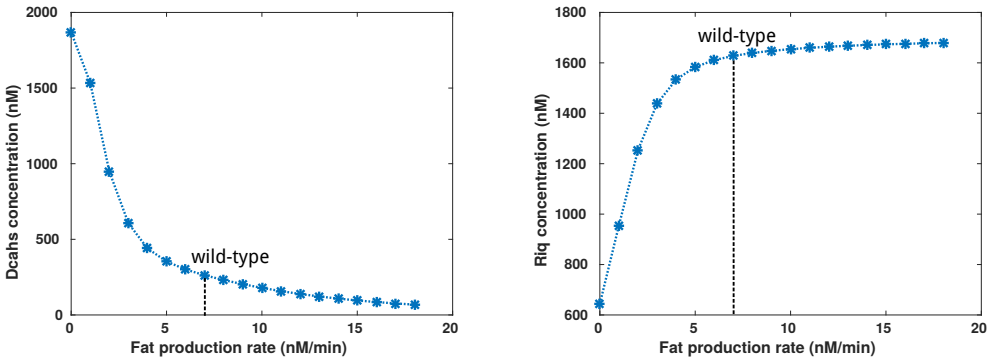


Figure 3.6: The concentrations of localized Dachs (left) and Riq (right) on cell membrane change with respect to Fat expression level.

The model also reveals a non-intuitive effect of the Ds expression level on growth. Previous work has shown that overexpression of Ds represses Yki activity [49, 48], and others have found that it simulates Yki activity [44]. Our results suggest that this paradox is due to difference in Ds overexpression level, which is likely to be caused by using different Gal4 drivers in experiments. Slicing vertically through the previous

response map generates a predicted growth curve as a function of Ds expression level at a normal Fat background shown in Figure 3.7. While strong overexpression of Ds inhibits Yki activity and growth, moderate or mild overexpression of Ds elevates Yki activity and stimulates cell growth. Moreover, the threshold that switches the role of Ds in growth depends on the amount of Fat. Furthermore, the model predicts the occurrence of a non-monotonic effect on growth when Ds function is lost, i.e. complete loss of Ds causes overgrowth, and partial loss of Ds reduces growth. This is remarkably similar to the observation when Fat function is lost, emphasizing the similarity of two atypical cadherins. It can be easily tested by experiments.

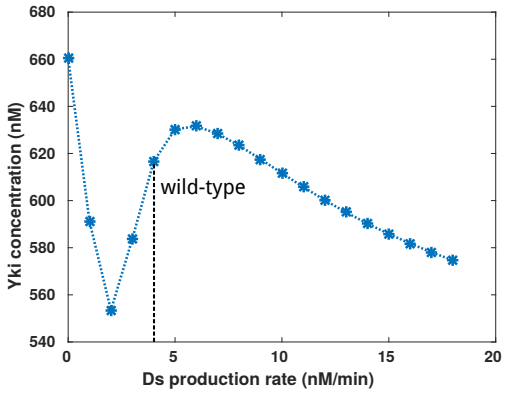


Figure 3.7: Vertical slice of the growth response map shows a non-monotonic dependence of growth on Ds expression level.

Another interesting prediction of the model is that double mutants of *fat* and *ds* overgrow more than either of single mutants (Figure 3.9), which is consistent with experimental observations. Previous work has shown that knockout of *ds* potentiates the overgrowth in *fat* mutants, but failed to elucidate the mechanistic basis. Our model provides a mechanistic explanation and suggests that the interaction between Ds and Dachs is responsible for the double mutant phenotype. In the model, we assume that the Ds-Dachs complex decays faster than Dachs alone on the cell membrane, which gives the interaction between Ds and Dachs a dual role; with Fat present, the localization of Dachs on the cell membrane is greatly suppressed, and thereby the interaction between Dachs and Ds favors the localization of Dachs. However, in the absence of Fat, the localized

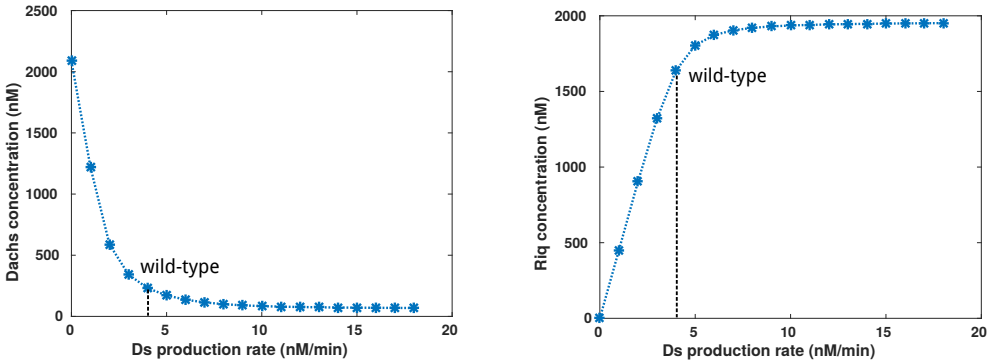


Figure 3.8: The concentrations of localized Dachs (left) and Riq (right) on cell membrane change with respect to Ds expression level

Dachs on the membrane is in abundance, and this interaction promotes the degradation of localized Dachs and hence reduces the amount of Dachs on cell membrane. In other words, the effect of Ds on Dachs is Fat-dependent, and is positive with Fat and negative without Fat.

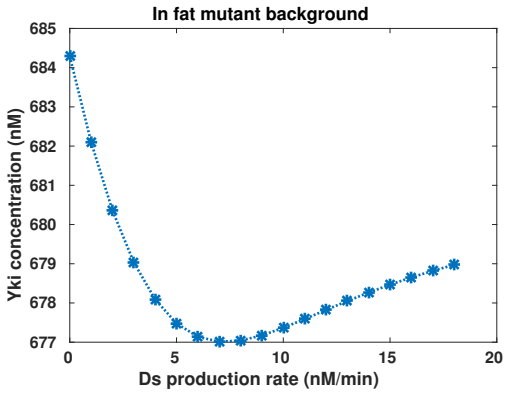


Figure 3.9: In *fat* mutant background, yki activity changes with respect to Ds expression level. It is generated from slicing the growth response map in Figure 3.4 vertically at zero Fat production rate

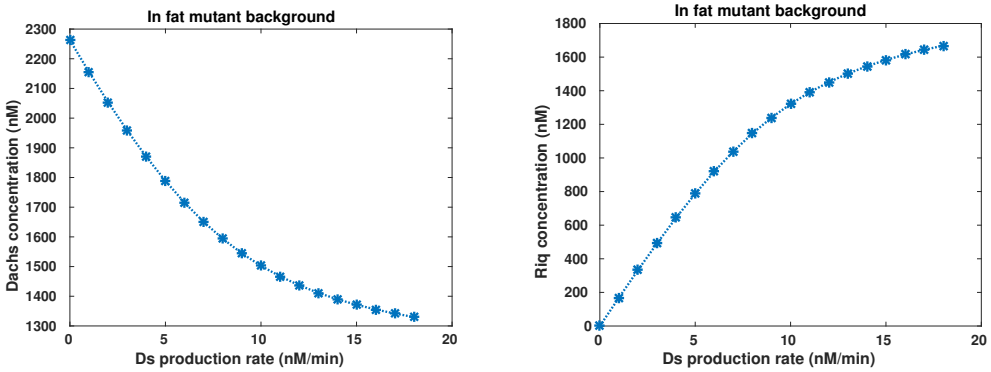


Figure 3.10: In *fat* mutant background, the concentrations of localized Dachs (left) and Riq (right) on the cell membrane change with respect to Ds expression level

3.4.2 Non-autonomous Response

Apart from the fact that Fat and Ds signal through the Hippo pathway to regulate cell growth, they also participate in cross-talk between cells, and are involved in signal transport across the disc. Here we focus on the way the cell detects and responds to the cues from its neighbors via the Fat/Ds signaling pathway, which in turn can modulate its growth. Non-autonomous responses, phenotypes induced in wild-type cells by mutant cells, have been observed in a variety of experiments when Fat/Ds signaling is disrupted in a patch of clone cells in an otherwise normal disc. For instance, experiments have shown that overexpression of Ds in cell clones stimulates Yki activity and growth both autonomously and non-autonomously. It induces hyperactivity of Yki and overexpression of target genes in cells around the boundary between wild-type and clone cells [9] that gradually declines away from the boundary. The predictions of our full model shown in Figure 3.11 incorporating cell-cell interactions match experimental observations. Moreover, the model predicts that the non-autonomous hyperactivity of Yki is due to elevation of localized Dachs concentration on the cell membrane, whereas the autonomous effect on Yki activity in clone cells is due to an increase in localized Riq concentration on the cell membrane (Figure 3.12).The non-autonomous effects suggest that cells have the potential to respond to abnormal activity and adjust growth. Interestingly, overexpression of Ds evenly in cell clones in a uniform expression of Fat and Ds

background induces asymmetric localization of Dachs within cells around the boundary. This is consistent with previous observations on cell polarity. Of course, it is caused by the asymmetric localization of Fat-Ds complex on the cell membrane. However, a considerable amplification is observed from Fat/Ds to Dachs.

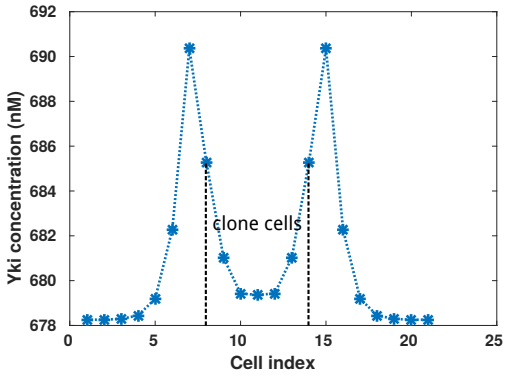


Figure 3.11: Predicted autonomous and non-autonomous growth responses around the boundary induced by overexpressing Ds in a cell clone. Growth response is measured by averaged Yki concentration. An array of 21 cells in total are simulated with a patch of 7 clone cells in the middle between the two vertical dashed lines

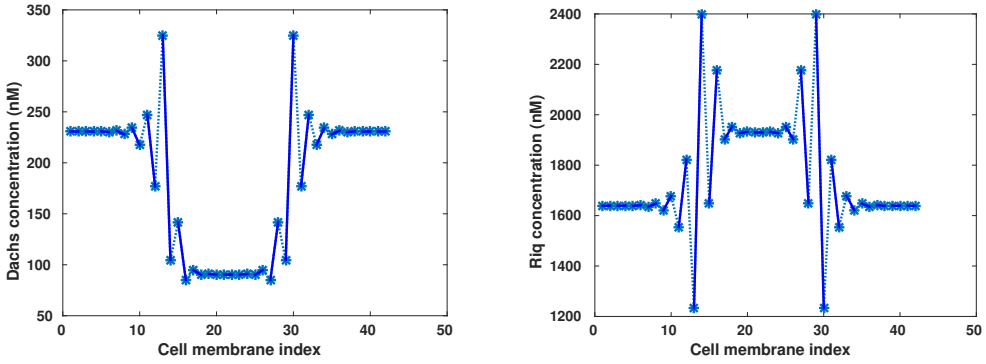


Figure 3.12: Level of Dachs and Riq localized on cell membranes with Ds overexpression clone cells. 42 locations of cell membranes from 21 cells are simulated with a patch of 14 cell membranes (7 clone cells) in the middle.

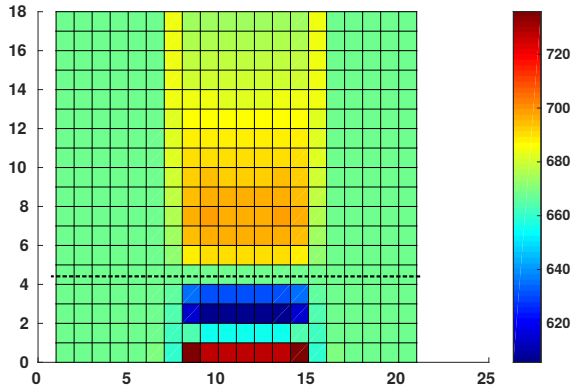


Figure 3.13: The effect of Ds expression level in clone cells on autonomous and non-autonomous growth responses measured by averaged Yki concentration reflected in color. 21 cells are simulated with a patch of 7 clone cells in the middle. The dashed line represents that Ds expression level in clone cells is the same with that in wild-type cells.

Figure 3.13 shows a colormap of the variation of Yki concentration in space, which represents variation in autonomous and nonautonomous responses, produced by changing the Ds expression level in clone cells. Each horizontal slice corresponds to an array of 21 cells where Ds is locally overexpressed in 7 clone cells at the center, and different horizontal slices correspond to different Ds expression level in clone cells. With Ds overexpression cell clones where Ds expression level is higher than in wild-type cells, although Yki hyperactivity in cells around the boundary between wild-type and clone cells becomes more pronounced as Ds expression level increases, non-autonomous response extends further from the boundary a bit but not significantly, and can be observed $2 \sim 4$ cells outside the cell clone. This agrees well with experimental observations, and suggests that cells maintain a robust response to abnormal input signal from its neighbors indicated by the difference in Ds expression level. In addition, the colormap shows that complete loss of Ds in clone cells induces non-autonomous response in wild-type cells, and simulates Yki activity autonomously. However, the Yki concentration in clone cells

is uniform, which is different from the autonomous response induced by Ds overexpression clone cells, where Yki concentration is gradually declining from the boundary to the inside of the clone. This indicates that Ds is essential for cells to respond to the boundary signal, which is consistent with experimental results [9]. Furthermore, the colormap also shows that partial loss of Ds in cell clones can elicit similar autonomous and non-autonomous responses as overexpression of Ds. Therefore, we go further and predict that the boundary effect can be activated as long as Ds is present and Ds expression level differs in two adjacent cells. This can be tested by experiments.

3.4.3 The role of Fj

Fj has played multiple roles in the Hippo pathway. One is to facilitate cell-cell communication through phosphorylation of both Fat and Ds, which have opposite effects on binding processes between each other. The other is to provide a feedback loop for the signaling pathway. Of course these two effects may contribute significantly to cell polarization as previous studies suggest. Here we concentrate on their impact on cell signaling. The model predicts that either overexpression of Fj or loss-of-function of Fj leads to reduction in growth measured by Yki concentration, and the impact of interfering Fj expression has relatively mild effect on Yki activity in compared with interfering Fat or Ds expression. This is consistent with experimental observation.

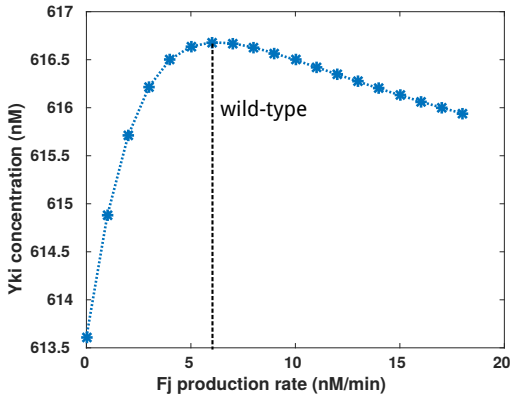


Figure 3.14: The effect of Fj on growth measured by Yki concentration

3.4.4 Signal Propagation

Propagation of the boundary effect, especially the non-autonomous response, is typically seen to reach $2 \sim 4$ cells away from the boundary between clone and wild-type cells. The interaction between Fat and Ds controls signal transport between cells, and the intracellular processes regulate signal transduction within a cell. Together, they determine the spatial response to cell clones and how far the boundary effect propagates.

The interaction between Fat and Ds in our full model is modified by Fj and gets complicated as they also interact with downstream factors. For simplicity without loss of generality, here we present the computational analysis of a general model of reciprocal binding processes between two species (heterodimer formation) and explore the impact of parameters on signal propagation in the disc. The following analysis is restricted to two species that form heterodimers on opposing cell membranes and mediate cell-cell interaction.

Below we will elaborate upon this general view and zoom into specific processes that influence signal transport between cells, and discuss respectively how transport within a cell (diffusion), binding processes (speed and affinity), and degradation processes (decay rate) affect the boundary effect propagation. The predictions of the model show that the boundary effect induced by overexpression of Ds in cell clones, including autonomous and non-autonomous responses, are remarkably stable as the nature of signal propagation does not change when the parameters for different kinetic processes are changed within an order of magnitude, although the degree and extent are affected. That is to say, it maintains a sensitivity over a relatively broad range of parameters but it is not highly sensitive to possible fluctuations.

$$F_c^i \underset{k_F^-}{\overset{k_F^+}{\rightleftharpoons}} F_m^i$$

$$S_c^i \underset{k_S^-}{\overset{k_S^+}{\rightleftharpoons}} S_m^i$$

$$S_{m_1}^i + F_{m_2}^{i-1} \underset{k_{FS}^-}{\overset{k_{FS}^+}{\rightleftharpoons}} \overline{FS}_{m_1}^i$$

$$S_{m_2}^i + F_{m_1}^{i+1} \underset{k_{FS}^-}{\overset{k_{FS}^+}{\rightleftharpoons}} \overline{FS}_{m_2}^i$$

At steady state, the governing equations are the following:

Fat concentration in cytosol:

$$D_F \frac{\partial^2 F_c^i}{\partial x^2} + \alpha_F - \beta_F F_c^i = 0 \tag{3.24}$$

with boundary conditions:

$$-\frac{D_F}{\varepsilon}\frac{\partial F_c^i}{\partial x}|_{x_i} = -k_F^+ F_{c_1}^i + k_F^- F_{m_1}^i$$

$$-\frac{D_F}{\varepsilon}\frac{\partial F_c^i}{\partial x}|_{x_{i+1}} = k_F^+ F_{c_n}^i - k_F^- F_{m_2}^i$$

Ds concentration in cytosol:

$$D_S \frac{\partial^2 S_c^i}{\partial x^2} + \alpha_S - \beta_S S_c^i = 0 \tag{3.25}$$

with boundary conditions:

$$-\frac{D_S}{\varepsilon}\frac{\partial S_c^i}{\partial x}|_{x_i} = -k_S^+ S_{c_1}^i + k_S^- S_{m_1}^i$$

$$-\frac{D_S}{\varepsilon}\frac{\partial S_c^i}{\partial x}|_{x_{i+1}} = k_S^+ S_{c_n}^i - k_S^- S_{m_2}^i$$

Fat concentration on cell membrane:

$$k_F^+ F_{c_1}^i - k_F^- F_{m_1}^i - k_{FS}^+ F_{m_1}^i S_{m_2}^{i-1} + k_{FS}^- \overline{FS}_{m_2}^{i-1} - \gamma_F F_{m_1}^i = 0 \tag{3.26}$$

$$k_F^+ F_{c_n}^i - k_F^- F_{m_2}^i - k_{FS}^+ F_{m_2}^i S_{m_1}^{i+1} + k_{FS}^- \overline{FS}_{m_1}^{i+1} - \gamma_F F_{m_2}^i = 0 \tag{3.27}$$

Ds concentration on cell membrane:

$$k_S^+ S_{c_1}^i - k_S^- S_{m_1}^i - k_{FS}^+ S_{m_1}^i F_{m_2}^{i-1} + k_{FS}^- \overline{FS}_{m_1}^i - \gamma_S S_{m_1}^i = 0 \tag{3.28}$$

$$k_S^+ S_{c_n}^i - k_S^- S_{m_2}^i - k_{FS}^+ S_{m_2}^i F_{m_1}^{i+1} + k_{FS}^- \overline{FS}_{m_2}^i - \gamma_S S_{m_2}^i = 0 \tag{3.29}$$

Fat-Ds concentration on cell membrane:

$$k_{FS}^+ F_{m_2}^{i-1} S_{m_1}^i - k_{FS}^- \overline{FS}_{m_1}^i - \gamma_{FS} \overline{FS}_{m_1}^i = 0 \tag{3.30}$$

$$k_{FS}^+ F_{m_1}^{i+1} S_{m_2}^i - k_{FS}^- \overline{FS}_{m_2}^i - \gamma_{FS} \overline{FS}_{m_2}^i = 0 \tag{3.31}$$

Diffusion

Physically, transport from one location to the other on cell membrane, which allows communication between molecules at different locations on the cell membrane, is essential for the boundary effect propagation. In the model, transport between locations is represented by diffusion. Thus, as expected, faster diffusion leads to further signal propagation from the boundary between wild-type and clone cells shown in Figure 3.15. When one side of the cell detects the abnormality of its neighbors via heterodimer formation on the membrane, the corresponding change on the other side is more sensitive when the diffusion coefficient is increased.

We are aware that in reality, transport between locations could be via either transport in cytosol or movement on the cell membrane, whereas transport via cytosol is the only way in our model since there is no diffusion on the cell membrane. However, in general, movement such as diffusion in the cytosol is much faster than that on the cell membrane, and thereby it is likely that communication via diffusion in the cytosol that discussed in our model dominates. In addition, it should be noted that in reality signal transport in a cell is more complex than simple diffusion. Here diffusion coefficient can be regarded in general as a measure of molecular movement, and we expect that mechanisms that facilitate molecular movement will promote boundary signal propagation.

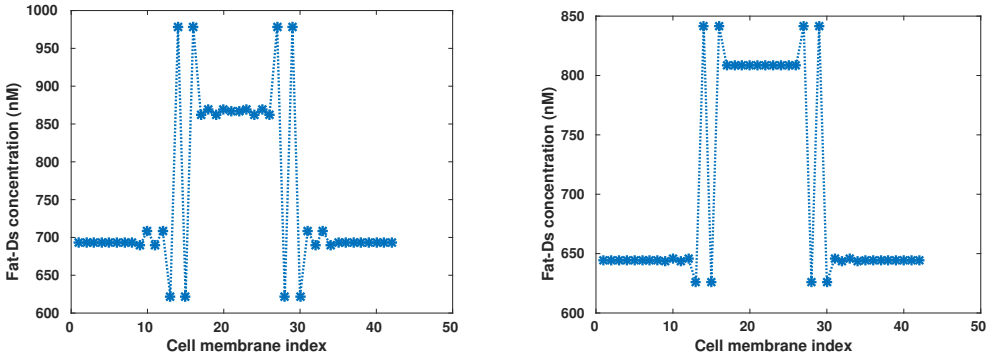


Figure 3.15: Signal propagation in response to the change of diffusion coefficients of Fat and Ds in cytosol on . Left:$D_F = 14\mu m^2/min, D_S = 16\mu m^2/min$; Right: $D_F = 1.4\mu m^2/min, D_S = 1.6\mu m^2/min$

Membrane localization

To explore how membrane localization processes influence signal propagation, we examine two cases: 1) speeding up both on (from cytosol to membrane) and off (from membrane to cytosol) rates at a fixed ratio; and 2) varying the ratio of on to off rates.

Firstly, expediting membrane localization processes without changing the affinity enhances the boundary effect on both sides of the boundary in extent and in amplitude shown in Figure 3.16. Mathematically, it is due to the increase of mass flux from cytosol to membrane.

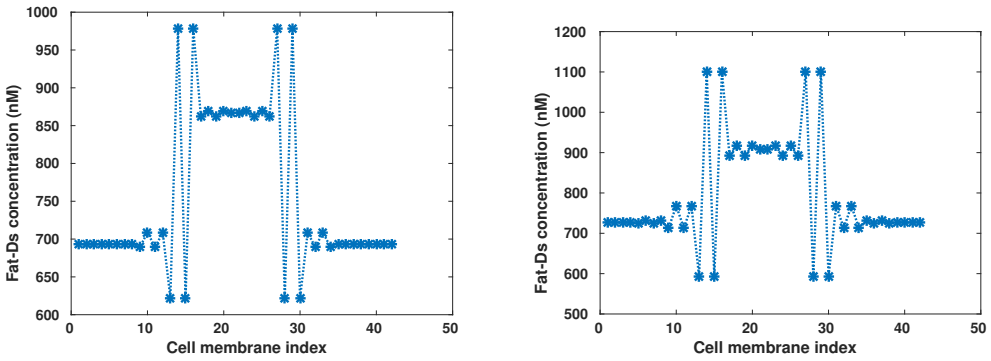


Figure 3.16: Signal propagation in response to the change of the speed of the binding processes from cytosol to membrane. Left: $k_F^+ = k_S^+ = 1min^{-1}, k_F^- = k_S^- = 0.1min^{-1}$; Right: $k_F^+ = k_S^+ = 10min^{-1}, k_F^- = k_S^- = 1min^{-1}$

We then vary the ratio of on to off rates by changing on rates at a fixed off rate, and find that there exists an optimal on rate that results in the most prominent boundary effect, as shown in Figure 3.17. Moreover, the optimal ratio of on to off rates depends on the heterodimer process.

Heterodimer (Fat-Ds) formation

To examine the impact of heterodimer formation, we tested the model in two circumstances: 1) speeding up the process without changing the affinity; 2) using three different binding affinities.

Speeding up forward binding and backward dissociation rates of Fat-Ds heterodimer

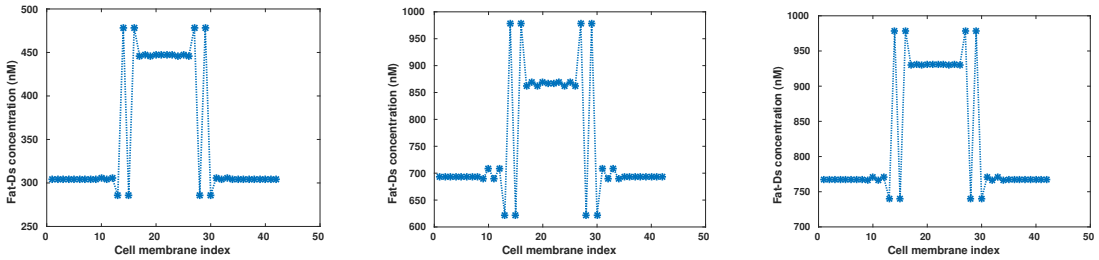


Figure 3.17: Signal propagation in response to the change of the on rates (affinities) from cytosol to membrane. Left: $k_F^+ = k_S^+ = 0.1min^{-1}$; Middle(base): $k_F^+ = k_S^+ = 1min^{-1}$; Right: $k_F^+ = k_S^+ = 10min^{-1}$

formation simultaneously without changing binding affinities slightly enhances the boundary effect (results are not shown here). Meanwhile, similarly, we use three different binding affinities by varying the forward rates at a fixed backward rate, and find that there exists an optimal forward rate which results in the most prominent boundary effect shown in Figure 3.18. The optimal value of k_{FS}^+ depends on the amount of Fat and Ds available shown in Figure 3.19.

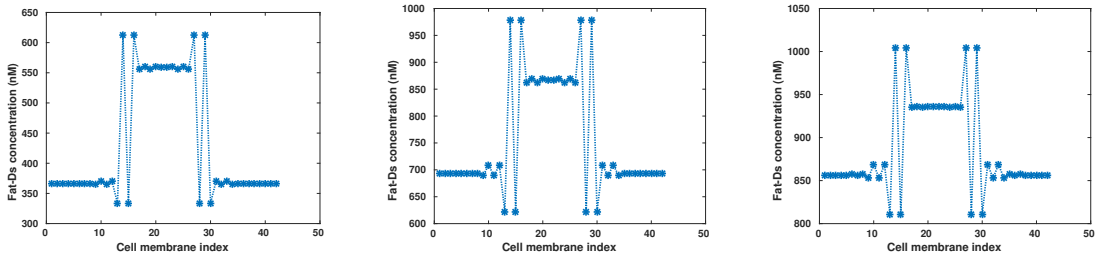


Figure 3.18: Signal propagation in response to the change of forward binding rates (affinities) between Fat and Ds on the membrane. Left: $k_{FS}^+ = 0.02nM^{-1}min^{-1}$; Middle (base): $k_{FS}^+ = 0.2nM^{-1}min^{-1}$; Right: $k_{FS}^+ = 2nM^{-1}min^{-1}$

Degradation processes

We also look at the impact of degradation rates on signal propagation, and find that they affect the boundary effect in a reverse manner, i.e. increasing degradation rates

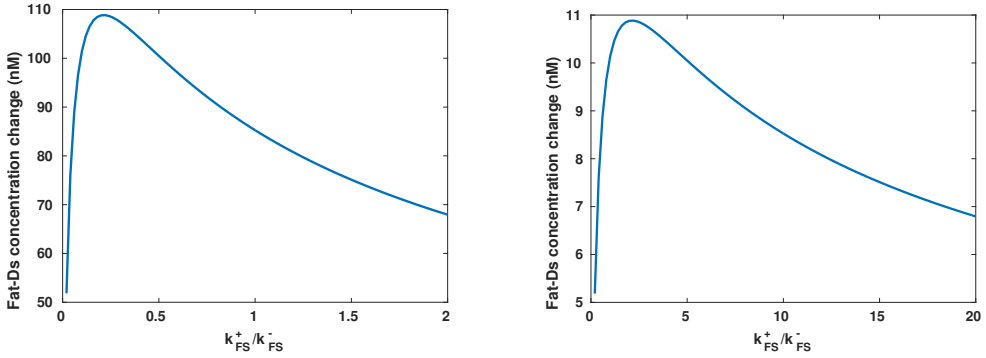


Figure 3.19: Each figure represents how the amplitude of the boundary effect measured by the change of Fat-Ds concentration responds to the change of binding affinities between Fat and Ds on the membrane. Left: $\alpha_F = \alpha_S = 40nM \cdot min^{-1}$ (normal), $\alpha_S = 80nM \cdot min^{-1}$(clone); Right: $\alpha_F = \alpha_S = 4nM \cdot min^{-1}$ (normal), $\alpha_S = 8nM \cdot min^{-1}$(clone)

reduce the degree and/or spatial extent of the non-autonomous response, and vice versa. Illustrations of the predicted responses are shown below.

Increasing the decay rates of Fat and Ds on cell membrane significantly decreases the extent of the boundary effect outside the clone shown in Figure 3.20.

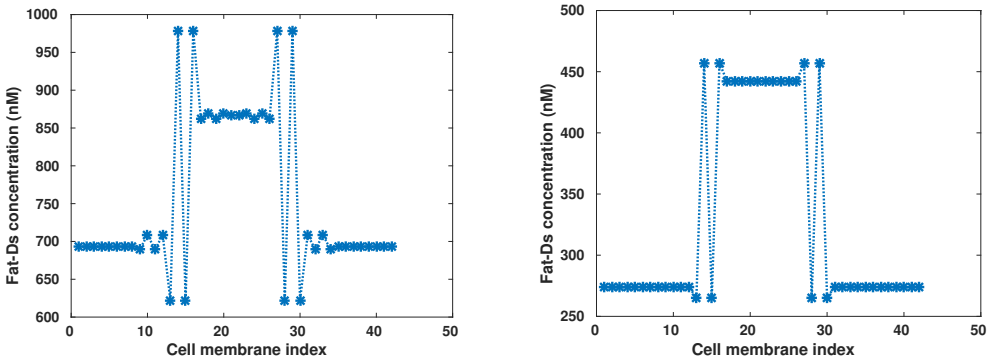


Figure 3.20: Signal propagation in response to the change of decay rates of Fat and Ds on the membrane. Left:$\gamma_F = \gamma_S = 0.02min^{-1}$; Right: $\gamma_F = \gamma_S = 0.1min^{-1}$

Increasing the decay rate of Fat-Ds complex primarily reduces the amplitude of the boundary effect shown in Figure 3.21.

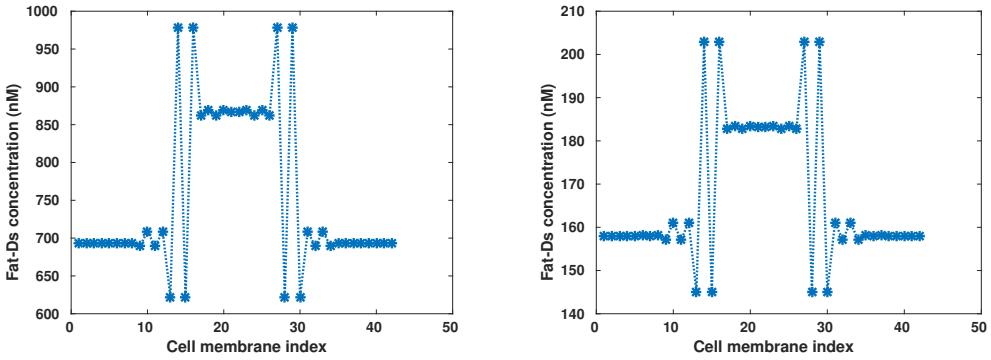


Figure 3.21: Signal propagation in response to the change of decay rates of Fat-Ds complex on the membrane. Left:$\gamma_{FS} = 0.02min^{-1}$; Right: $\gamma_{FS} = 0.1min^{-1}$

Increasing the decay rates of Fat and Ds in cytosol reduces the boundary effects on both sides shown in Figure 3.22.

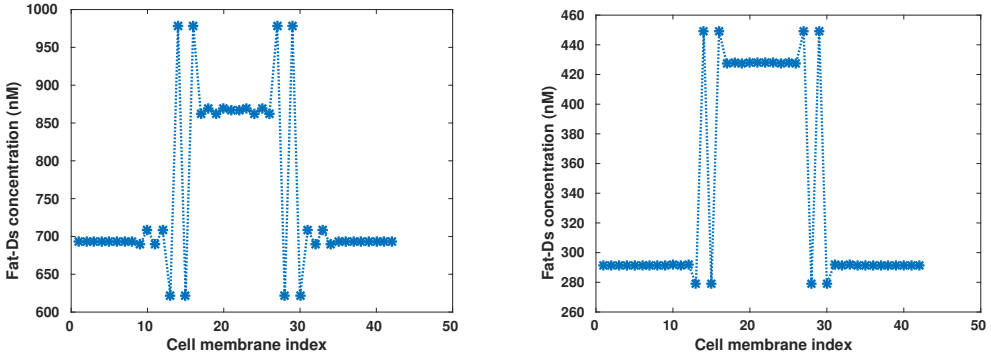


Figure 3.22: Signal propagation in response to the change of decay rates of Fat and Ds in cytosol. Left: $\beta_F = \beta_S = 0.1min^{-1}$; Right: $\beta_F = \beta_S = 1min^{-1}$

3.5 Discussion

Given the complex behaviors of this highly coupled signaling network, qualitative arguments fall short of providing insights, and reliable and testable computational models are required. The goal of our model is to provide a framework that would help to interpret the complex phenotypes associated with the pathway and make testable predictions that can guide further experimental study. With the model we developed, we have reproduced all primary experimental observations, such as non-monotonic effects in disc-wide intervention of Fat and Ds expression, and non-autonomous effects induced by cell clones. The model suggests that those seemingly inexplicable observations result from the perturbation of the delicate balance between positive and negative control from intra- and intercellular signals. Moreover, we have found that the regulation of Dachs and Riq location to the cell membrane plays an important role in both non-autonomous and non-monotonic effects. The model suggests a difference between autonomous and non-autonomous response simulated by clone cells with disrupted Fat/Ds signaling pathway. The impact of various kinetic processes on the boundary effect propagation are also explored. The model also provides a mechanistic explanation for the *fat*, *ds* double mutant phenotype, which supports our hypothesis that Ds interacts with Dachs. The presented work addresses the limitation of previous models due to lack of mechanistic details. The fact that the model predicts all the primary characteristic phenotypes demonstrates the feasibility of the underlying biological model.

The non-monotonic response of Fat on growth and the non-autonomous response induced by overexpression of Ds in cell clones has also been explained by a recent model which assumes mutual inhibition between the opposite orientations of the heterodimers while self-promotion of the same orientations [47]. While it sounds appealing, there is little experimental evidence in support of such assumptions. In contrast, our model does not assume such roles and yet predicts non-monotonic and non-autonomous responses. We make use of the positive regulatory step from Ds via Riq and the potential interaction between Ds and Dachs that is detected in vitro but not yet in vivo, and propose the similar interaction with further hypothesis concerning the effect of interaction on degradation, focusing efforts on elucidating these interactions would generate important mechanistic insights.

A concern following our model and analysis is the role of dimension in model predictions. As all the major experimental observations can be presented in one dimensional model, our model is a feasible and rational start to understand the behaviors of this pathway. This is an important step toward the overall goal of developing a mechanistically-realistic model. Of course, adopting the model into two or three dimensions will be one of our future directions. We suspect that it might not change our conclusions, but rather affect the strength and/or extent of observations that we have discussed.

It is also important to note that the movement of species in the cytosol has been modeled as diffusion while in reality more complex mechanisms may be involved. Although our results are not sensitive to diffusion coefficients, we have pointed out that faster movement for species in cytosol favors further propagation of the boundary effect. There might be a complex mechanism such as active transport involved which allows the big molecules such as Fat and Ds to speed up their movement and hence facilitate signal propagation.

In the future, we will study the mechanism underlying the temporal phenomenon, i.e. the non-autonomous effect diminishes in time and eventually vanishes, as is observed experimentally. Furthermore, the non-autonomous response also occurs in neighbor cells when Dpp signaling is disrupted [54, 91], which implies crosstalk between the Hippo and morphogen pathways. We will integrate different signal pathways and study how they interact and are balanced. Our mathematical model will provide a systematic characterization of the signaling pathways governing tissue patterning and growth in the disc. It will be easy to vary parameters in each pathway to test the predictions of the model against observations. This in turn can direct more targeted experiments, and accelerate progress toward a better integrated understanding of the mechanisms of patterning and growth control. As disruption of the Hippo signaling pathway is associated with many cancers, studies on this pathway might also provide insights into cancer development and potential therapeutics for treatment of the diseases.

Name	Description	Values	Units
Production rates			
α_F	Fat production rate	8	$nM \cdot min^{-1}$
α_S	Ds production rate	4	$nM \cdot min^{-1}$
α_J	Fj production rate	6	$nM \cdot min^{-1}$
α_A	Dachs production rate	5	$nM \cdot min^{-1}$
α_R	Riq production rate	4	$nM \cdot min^{-1}$
α_W	Wts production rate	4	$nM \cdot min^{-1}$
α_Y	Yki production rate	2	$nM \cdot min^{-1}$
Forward binding rates from cytosol to membrane			
k_F^+	Binding $F_c^i \to F_m^i$	1	min^{-1}
k_S^+	Binding $S_c^i \to S_m^i$	1	min^{-1}
Forward binding rates between two membrane species			
k_{FS}^+	Binding $F_m^i + S_m^{i-1} \to \overline{FS}_m^{i-1}$	4×10^{-3}	$nM \cdot min^{-1}$
k_{FS1}^+	Binding $Fp_m^i + S_m^{i-1} \to \overline{FpS}_m^{i-1}$	2×10^{-2}	$nM \cdot min^{-1}$
k_{FS2}^+	Binding $F_m^i + Sp_m^{i-1} \to \overline{FSp}_m^{i-1}$	2×10^{-3}	$nM \cdot min^{-1}$
k_{FS3}^+	Binding $Fp_m^i + Sp_m^{i-1} \to \overline{FpSp}_m^{i-1}$	4×10^{-3}	$nM \cdot min^{-1}$
k_{AS}^+	Binding between A_m and all the forms of Ds on the membrane	1×10^{-2}	$nM \cdot min^{-1}$
k_{AFS}^+	Binding between A_m and all the forms of $Fat - Ds$ complex on the membrane	1×10^{-2}	$nM \cdot min^{-1}$
Forward binding rates between membrane and cytosolic species			
k_{RS}^+	Binding between R_c and all the forms of Ds on the membrane	0.01	$nM \cdot min^{-1}$
k_{RFS}^+	Binding between R_c and all the forms of $Fat - Ds$ complex on the membrane	0.1	$nM \cdot min^{-1}$
k_{WR}^+	Binding between W_c and all the forms of $Ds - Riq$, $Fat - Ds - Riq$on the membrane	0.01	$nM \cdot min^{-1}$
k_{WA}^+	Binding between W_c and all the forms of $Dachs, Ds - Dachs, Fat - Ds - Dachs$ on the membrane	0.02	$nM \cdot min^{-1}$
Forward binding rates between two cytosolic species			
k_{JF}^+	Binding between Fj and Fat	1	$nM \cdot min^{-1}$
k_{JS}^+	Binding between Fj and Ds	1	$nM \cdot min^{-1}$
k_{WY}^+	Binding between W_c and Y_c and Yp_c	1	$nM \cdot min^{-1}$
Diffusion coefficients			
D_F	Diffusion, Fat	14	$\mu m^2 \cdot min^{-1}$
D_{Fp}	Diffusion, Phosphorylated Fat	12	$\mu m^2 \cdot min^{-1}$
D_S	Diffusion, Ds	16	$\mu m^2 \cdot min^{-1}$
D_{Sp}	Diffusion, Phosphorylated Ds	14	$\mu m^2 \cdot min^{-1}$
D_J	Diffusion, Fj	20	$\mu m^2 \cdot min^{-1}$
D_{JF}	Diffusion, Fat-Fj complex	12	$\mu m^2 \cdot min^{-1}$
D_{JS}	Diffusion, Ds-Fj complex	14	$\mu m^2 \cdot min^{-1}$
D_A	Diffusion, Dachs	10	$\mu m^2 \cdot min^{-1}$
D_R	Diffusion, Riq	10	$\mu m^2 \cdot min^{-1}$
D_W	Diffusion, Wts	20	$\mu m^2 \cdot min^{-1}$
D_Y	Diffusion, Yki	20	$\mu m^2 \cdot min^{-1}$
D_{Yp}	Diffusion, Phosphorylated Yki	18	$\mu m^2 \cdot min^{-1}$
D_{WY}	Diffusion, Wts-Yki complex	18	$\mu m^2 \cdot min^{-1}$

Table 3.2: Model parameters. All the backward dissociation rates and dephosphorylation rates are set to be $0.1 min^{-1}$, and the phosphorylation rates are $0.5 min^{-1}$. All the decay rates are $0.1 min^{-1}$ except for Ds-Dachs complexes which have the decay rate of $0.2 min^{-1}$.

Chapter 4

FRAP

To demonstrate the effect of different model assumptions and different ways of utilizing the data from a FRAP experiment, we avoid the above difficulties of unknown mechanisms and other factors by generating data computationally for a known model with known parameters and then testing our recovery of parameters from the data. By using a recovery model identical to the theoretical model, we show that the choice of observation time can significantly affect the estimates. We also show that changing the bleaching region to rebalance the diffusion and binding processes can significantly improve the estimates. By varying the recovery model from the theoretical model, we investigate whether the simplified recovery model can, in some circumstances, appropriately describe the FRAP process, the relationship between parameters in the theoretical model and those in the recovery model, and under what conditions some processes can be neglected in the recovery model. Lastly, we introduce sensitivity analysis as a technique to better understand FRAP data and to improve FRAP model formulation.

In the following section we begin with a simple example in which the parameters of a complex model can be related to parameters in a simplified description. We then develop and solve the evolution equation from which the parameters are estimated in a standard experiment, and we describe the computational setup and the analysis of the data. We provide a detailed analysis of simplified diffusion-reaction models of FRAP, and use these to show how the neglect of processes in FRAP lead to erroneous estimation. For simplicity we assume throughout that diffusion is the only spatial transport process involved (internalization of receptor-ligand complexes is allowed, as

discussed later). We provide the solution for models with influx (or production) and decay. We further restrict attention to geometrically one-dimensional systems, but the method can easily be generalized to 2D or 3D, and can be used to study more complicated questions in FRAP, for example, when binding is nonlinear. We believe it advances our understanding of the limitations of the existing FRAP experiments and models, helps to reconcile the parameter estimates in biological systems, and will direct the improvement of the FRAP technique.

4.1 The mathematical framework for parameter estimation and model testing

We begin the mathematical description of FRAP with a simplified geometric description of a wing disc for the purpose of (i) emphasizing the assumptions implicit, but rarely discussed, in many FRAP analyses, and (ii) showing that coefficients extracted for a simple description may reflect more complicated processes than the usual interpretation of parameters would suggest. A general formulation of the linear reaction-diffusion systems that govern FRAP analysis is given in the Appendix 4.1.1. As shown in Figure 2.1, the geometry of the disc is complex, and morphogen transport in the disc may involve several different mechanisms that are discussed later, but for simplicity we consider a thin fluid layer (Figure 4.1), equivalent to the lumen in the wing disc, and admit diffusion, binding to surface-bound receptor, and internalization and re-expression of receptors to emphasize assumptions implicit in most analyses. We assume that the surface reactions involve only binding to a receptor and decay of the receptor-ligand complex, and to simplify the analysis, suppose that whenever a receptor-ligand complex is internalized it is replaced by a bare receptor [20]. We measure receptor and receptor-ligand concentrations in molecules or moles per unit area, and we assume that the steps by which an occupied receptor is internalized and a free receptor is recycled to the surface reach a steady state rapidly compared with other processes, which implies that the total amount of receptor is constant at every point on the surface $z = 0$, *i.e.,* $R + \overline{RC} = R_T$ where R_T is a constant.

The surface $z = 0$ can be regarded as the outer cell boundary of a sheet of cells covered by a thin layer of fluid, as in a simplified description of the *Drosophila* wing

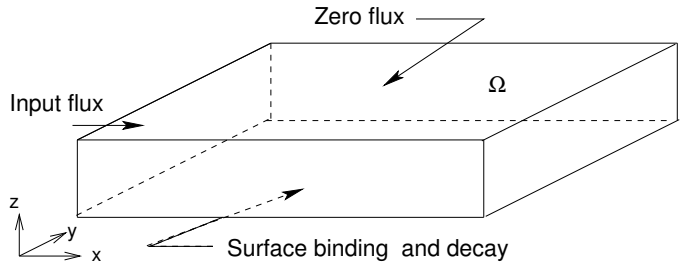


Figure 4.1: The geometry of a thin fluid layer over receptors embedded in a surface. Modified from [14].

disc. The lengths in the x, y and z directions are L_x, L_y, and L_z, respectively, and we let C be the concentration of a morphogen in the fluid and R the concentration of receptor on the surface $z = 0$. Suppose, as in the earlier 1D example, that there is a fixed influx of C on the boundary $x = 0$, and zero flux on the remaining faces except $z = 0$; then the governing equations can be written as follows.

$$\frac{\partial C}{\partial t} = D\Delta C \qquad \text{in} \quad \Omega \tag{4.1}$$

$$\frac{\partial R}{\partial t} = -k^+RC + (k^- + k_e)\overline{RC} \qquad \text{on} \quad z = 0 \tag{4.2}$$

$$\frac{\partial \overline{RC}}{\partial t} = k^+RC - (k^- + k_e)\overline{RC} \qquad \text{on} \quad z = 0 \tag{4.3}$$

$$-D\frac{\partial C}{\partial z} = -k^+RC + k^-\overline{RC} \qquad \text{on} \quad z = 0 \tag{4.4}$$

$$-D\frac{\partial C}{\partial x} = j \qquad \text{on} \quad x = 0 \tag{4.5}$$

$$D\frac{\partial C}{\partial x} = 0 \qquad \text{on} \quad x = L \tag{4.6}$$

where k^+ and k^- are the binding and dissociation rates between ligand and receptor, and k_e is the decay rate of the receptor-ligand complex.

This system can be simplified by defining the dimensionless variables $u = C/C_0$, $v = R/R_T$, $w = \overline{RC}/R_T$, the scaled coordinates $\xi = x/L_x$, $\eta = y/L_y$ and $\zeta = z/L_Z$,

and the dimensionaless time $\tau = t/T$. The system then becomes

$$\begin{aligned}
\frac{\partial u}{\partial \tau} &= \frac{DT}{L_x^2}\left(\frac{\partial^2 u}{\partial \xi^2} + \frac{L_x^2}{L_y^2}\frac{\partial^2 u}{\partial \eta^2} + \frac{L_x^2}{L_z^2}\frac{\partial^2 u}{\partial \zeta^2}\right) && \text{in} \quad \Omega && (4.7)\\
\frac{\partial v}{\partial \tau} &= -Tk^+C_0R_Tuv + T(k^- + k_e)(1-v) && \text{at} \quad \zeta = 0 && (4.8)\\
-\left(\frac{DC_0}{L_z}\right)\frac{\partial u}{\partial \zeta} &= -k^+R_TC_0uv + k^-(1-v) && \text{at} \quad \zeta = 0 && (4.9)\\
-\frac{\partial u}{\partial \xi} &= \frac{jL_x}{DC_0} && \text{at} \quad \xi = 0 && (4.10)\\
\frac{\partial u}{\partial \xi} &= 0 && \text{at} \quad \xi = 1 && (4.11)
\end{aligned}$$

In view of the boundary conditions the solution must be constant in the η direction at steady state, and we assume this for the transient problem as well. Furthermore, since the fluid layer is thin $L_z << L_x, L_y$, and the equations can be averaged over ζ. In this case the equations reduce to

$$\begin{aligned}
\frac{\partial \bar{u}}{\partial \tau} &= \left(\frac{DT}{L_x^2}\right)\frac{\partial^2 \bar{u}}{\partial \xi^2} - Tk^+R_T\bar{u}\bar{v} + \left(\frac{Tk^-R_T}{C_0}\right)(1-\bar{v}) && \text{in} \quad \Omega\\
\frac{\partial \bar{v}}{\partial \tau} &= -Tk^+C_0\bar{u}\bar{v} + T(k^- + k_e)(1-\bar{v}) && \text{at} \quad \zeta = 0\\
-\frac{\partial \bar{u}}{\partial \xi} &= \frac{jL_x}{DC_0} && \text{at} \quad \xi = 0\\
\frac{\partial \bar{u}}{\partial \xi} &= 0 && \text{at} \quad \xi = 1
\end{aligned} \qquad (4.12)$$

where $\bar{u}$ and $\bar{v}$ are the averages over ζ. At steady state the system reduces to

$$\begin{aligned}
\frac{d^2u}{d\xi^2} &= \gamma^2\frac{u}{K+u} && \text{in} \quad \Omega\\
-\frac{du}{d\xi} &= J && \text{at} \quad \xi = 0\\
\frac{du}{d\xi} &= 0 && \text{at} \quad \xi = 1
\end{aligned} \qquad (4.13)$$

where u now stands for the average over ζ, and

$$K = \frac{k^- + k_e}{k^+C_0} \qquad \gamma^2 = \frac{k_eR_TL_x^2}{DC_0L_z} \qquad J = \frac{jL_x}{DC_0}.$$

If $u << K$ then this reduces to

$$\begin{aligned} \frac{d^2u}{d\xi^2} &= \delta^2 u \qquad \text{in} \qquad \Omega \\ -\frac{du}{d\xi} &= J \qquad \text{at} \qquad \xi = 0 \\ \frac{du}{d\xi} &= 0 \qquad \text{at} \qquad \xi = 1 \end{aligned} \tag{4.14}$$

where

$$\delta^2 = \frac{k_e k^+}{k^- + k_e} \frac{R_T L_x^2}{D L_z} \equiv k_s \frac{R_T L_x^2}{D L_z}.$$

The dimensionless solution u is

$$u(\xi) = \frac{j_1}{\delta} \left[\frac{e^{\delta(1-\xi)} + e^{-\delta(1-\xi)}}{e^{\delta} - e^{-\delta}} \right]. \tag{4.15}$$

The stationary distribution is characterized by two dimensionless parameters: δ and J. The first is the square root of the ratio of a diffusion time scale τ_D and a kinetic time scale $\tau_K \equiv k^{-1}$, and the second is the ratio of the input flux j to a characteristic velocity defined by the diffusion constant and the decay rate. The former enters in the shape function ϕ via the exponential terms and determines how rapidly the morphogen concentration decays in space: the larger δ the more rapidly the solution decays from the value at the source. Thus reducing the kinetic scale by reducing the half-life of the morphogen, or increasing the diffusion time scale by decreasing the diffusion constant, leads to sharper, more rapidly-decreasing spatial profiles. It should also be noted that the second term in both the numerator and denominator of (4.15) arises from the finite length of the domain, and can only be neglected if $\delta \gg 1$.

While it is sometimes assumed that δ also controls the approach to the steady state, the following shows that this is not correct. To illustrate this as simply as possible, consider the transient version of (4.13), which reads

$$\begin{aligned} \frac{\partial c}{\partial \tau} &= \frac{\partial^2 c}{\partial \xi^2} - \delta^2 c \qquad \xi \in (0,1) \\ -\frac{\partial c}{\partial \xi} &= J \qquad \xi = 0 \\ \frac{\partial c}{\partial \xi} &= 0 \qquad \xi = 1 \\ c(\xi, 0) &= 0. \qquad \xi \in (0,1) \end{aligned} \tag{4.16}$$

The approach to steady is governed by the evolution of the difference $w \equiv u - u^s$, where u^s is the steady state solution (4.15). This satisfies (4.16) with $J = 0$ and $w(\xi, 0) = -u^s(\xi)$, and the solution is

$$w(\xi, \tau) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n \tau} \cos n\pi\xi, \tag{4.17}$$

wherein the constants a_n are determined by the steady-state solution. The exponential decay rates λ_n are given by

$$\lambda_n = (n\pi)^2 + \delta^2 \qquad n = 1, 2, \ldots \tag{4.18}$$

and the smallest of these, λ_1, defines the relaxation time of the slowest decaying mode $\cos(\pi\xi)$ in the transient solution. The reciprocal of this is a dimensionless relaxation time, and converting it to dimensional form one finds that the relaxation time to the steady state is given by

$$T_R \equiv \frac{1}{\dfrac{D\pi^2}{L^2} + k} = \frac{1}{\dfrac{\pi^2}{\tau_D} + \dfrac{1}{\tau_K}}. \tag{4.19}$$

This shows that either morphogen diffusion or morphogen decay can dominate the relaxation process, and their effect is additive. The relaxation time increases with a decrease in D or an increase in L, while the effect of the morphogen decay is independent of the space scale. In the context of the *Drosophila* wing disc the half life of the morphogen Dpp has been estimated as 45 min, and the diffusion coefficient is estimated to be $0.1\,\mu m^2/$ sec [15]. For a disc of 50μ the diffusion factor in the denominator of (4.19) is 0.0204, while the second factor is 0.022, which leads to a relaxation time of about 24 minutes. On the other hand, if the diffusion coefficient is $20\,\mu m^2/$ sec [28] the relaxation time is reduced to ~ 0.25 minutes. While the basis for the disparity is obvious in this example, the relaxation to a steady state following a perturbation could be used experimentally to gain additional information about the underlying processes.

In case the smaller value of D obtains, diffusion and decay are balanced on the scale of the wing disc, but in other systems the conclusion may be quite different. In the context of *Drosophila* embryonic development, the half-life of the transcription factor Bicoid has been estimated to range from ~ 8 mins [92] to less than ~ 30 mins [93], and if we use 20 mins as an intermediate estimate, $k = 0.05$ min^{-1}. Estimates of the

diffusion coefficient range upward from $0.3\,\mu m^2/$ sec [93], and thus for the lowest D and an embryo length of $L = 500\mu m$, the relaxation time of the slowest decaying mode is $\sim$20 mins, and is determined almost solely by the degradation rate.

Several assumptions are noteworthy. Firstly, u represents the average concentration over the thickness of the fluid layer due to the averaging over ζ. This is an appropriate description for most FRAP experiments, in which averaging over the ROI precedes the parameter estimation, but it must be noted that receptor concentrations have to be defined appropriately, and that the interpretation of binding constants reflects this. Secondly, though the steady-state problem with binding and internalization leads to the simple problem at (4.14), the parameter δ comprises several parameters that describe binding and internalization, and thus interpreting this as a simple decay constant is generally not valid.

4.1.1 The general framework

The general form of the system of reaction-diffusion equations that we shall use hereafter have the following form.

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c + \bar{R}(c, \bar{p}) \qquad \text{in } \Omega \tag{4.20}$$

$$\mathbf{n} \cdot D_c \nabla c = \bar{J} \qquad \text{on } \partial\Omega \tag{4.21}$$

$$c(r, 0) = c_0(r), \tag{4.22}$$

Here the vector $c = (c_1, c_2, \cdots, c_m)$ is the vector of chemical concentrations, D_c is assumed to be a constant diagonal matrix, and J is a prescribed flux on$\partial\Omega$. Ω is a bounded region in $\Re^q, q = 1, 2$ with a smooth boundary and outward normal $\mathbf{n}$. The functions $\bar{R}_i$ give the net rate of production of the i^{th} species and they are herein are always linear or quadratic polynomials in the c_i's. The vector $\bar{p}$ is a parameter vector, which can include the kinetic constants and perhaps species that appear in the kinetic mechanism but do not change significantly on the time scale of interest. As described in the preceding example, we can regard these equations as appropriate for a thin fluid layer over a 1D or 2D domain. In applications to the wing disc the geometry of the hexagonal packing of cells is more complex than the description above and a mathematical description that accounts for the geometric complexity is far more complex. A 2D model of the disc that incorporates this complexity is given in [94].

This system can be nondimensionalized as follows. Let L be a measure of the size of the system, C_i be a reference concentration for species i, and ω^{-1} be a time scale characteristic of the reactions[1] . Define the dimensionless quantities $u_i = c_i/C_i$, $\tau = \omega t$, $D_i = D_{ci}/\omega L^2$, and $\xi = r/L$, where $r \equiv (x_1, \cdots, x_q)$. The dimensionless governing equations are

$$\begin{aligned} &\frac{\partial u}{\partial \tau} = D\nabla^2 u + R(u,p) \quad \text{in } \Omega \\ &\mathbf{n}\cdot D\nabla u = J \qquad\qquad\quad \text{on } \partial\Omega \\ &u(\xi,0) = u_0(\xi), \end{aligned} \tag{4.23}$$

where $D = diag\{D_1, D_2, \cdots, D_m\}$, $J_i = \bar{J}_i/(\omega C_i)$, $R(u,p)$ is the dimensionless form of $\bar{R}(c,\bar{p})$, and Ω is scaled. If there are species that do not diffuse the corresponding D_i and J_i are zero, and unless stated otherwise, we assume that all boundary fluxes are zero, *i.e.,*we impose homogeneous Neumann boundary conditions.

We show later that in many FRAP experiments the kinetics can be linearized, and therefore in the majority of what follows we focus on linear kinetic models, and we write the system (4.23) as

$$\begin{aligned} &\frac{\partial u}{\partial \tau} = D\nabla^2 u + Ku \\ &\mathbf{n}\cdot\nabla u = 0, \qquad \text{on } \partial\Omega \\ &u(\xi,0) = u_0(\xi) \end{aligned} \tag{4.24}$$

This system has solutions of the form

$$u(\xi,\tau) = \sum_{n=0}^{\infty} y_n(\tau)\phi_n \tag{4.25}$$

where ϕ_n is a solution of the scalar eigenvalue problem

$$\begin{aligned} &\nabla^2\phi_n = -\alpha_n^2\phi_n \\ &\mathbf{n}\cdot\nabla\phi_n = 0 \end{aligned} \tag{4.26}$$

and y_n is given by

$$y_n(\tau) = e^{(K-\alpha_n^2 D)\tau} y_n(0). \tag{4.27}$$

[1] In the simulations done later we set $\omega = 1$ since the time scales of diffusion and kinetic processes are unknown beforehand. Thus the units of D_i are time^{-1} and similarly for the rate constants of linear reactions.

For a reasonable boundary the eigenfunctions ϕ_n form a complete orthonornal set under the standard L_2 inner product $[u, v] = \int u(\xi)v(\xi)d\xi$, and the solution of (4.24) can be written as [95]

$$u(\xi, \tau) = \sum_{n=0}^{\infty} e^{(K-\alpha_n^2 D)\tau} y_n(0)\phi_n(\xi.) \tag{4.28}$$

The initial condition can be written

$$u_0(\xi) = \sum_{n=0}^{\infty} y_n(0)\phi_n(\xi) \tag{4.29}$$

and therefore $y_n(0) = \langle u_0(\xi), \phi_n \rangle$, where here and hereafter $\langle \cdot, \cdot \rangle$ denotes the real or complex, as appropriate, Euclidean inner product, taken component-wise when one argument is a vector and the other a scalar.

Remark 1 *If the influx is non-zero,* i.e.,$J \neq 0$*, we let* $w = u - u^s$ *where* u^s *is the steady state solution. Then* w *satisfies*

$$\frac{\partial w}{\partial \tau} = D \bigtriangledown^2 w + Kw \tag{4.30}$$

with zero Neumann boundary conditions. The solution w *has the representation given at (4.28), and* u *is obtained from this. In particular, when* $m = 1$*, we obtain the solution given at (4.17).*

The difficulty in FRAP analysis of multicomponent systems, even when they are linear, stems from the structure of (4.28). To simplify the analysis, suppose that the family of matrices $\{K - \alpha_n^2 D\}$ is semisimple – which means that they can be diagonalized – for all n. Then the matrix exponential has the representation

$$e^{(K-\alpha_n^2 D)\tau} = \sum_{j=1}^{m} e^{\lambda_{jn}\tau} P_{jn} \tag{4.31}$$

where the projections P_{jn} are associated with a given λ_{jn} [96]. Since we assume that the $\{K - \alpha_n^2 D\}$ are semisimple they have the representation

$$P_{jn} = \Psi_{jn} * \Psi_{jn}^* \tag{4.32}$$

wherein Ψ_{jn} is an eigenvector of $K - \alpha_n^2 D$ associated with λ_{jn} and Ψ_{jn}^* is the corresponding adjoint eigenvector. The action of any P on a vector u is defined by

$$Pu = \langle \Psi_{jn}^*, u \rangle \Psi_{jn}.$$

The eigenvector $\Psi_{jn} = (\Psi_{1jn}, \Psi_{2jn}, \cdots, \Psi_{mjn})^T \in \Re^m$ is a solution of the algebraic eigenvalue problem

$$(K - \alpha_n^2 D - \lambda_{jn} I)\Psi_{jn} = 0, \qquad j = 1, 2, \cdots, m, \quad n = 0, 1, \ldots. \tag{4.33}$$

and the eigenvalues λ_{jn} are solutions of the characteristic equation

$$det(K - \alpha_n^2 D - \lambda I) = 0. \tag{4.34}$$

The problem simplifies significantly when K and D commute, for then the eigenvalues of $K - \alpha_n^2 D$ are simply $\lambda_{jn} = \lambda_j^K - \alpha_n^2 \lambda_j^D$.

The characteristic equation is an m^{th} degree polynomial for an m-component system. Thus, while the representation at (4.31) has the apparently simple form of a sum of exponentials, the eigenvalues λ_{jn} and the eigenfunctions Ψ_{jn} are complicated functions of the kinetic rate parameters and the diffusion constants. The case $m = 1$ is as done previously, and only if $m = 2$ or 3 can one make analytical progress toward understanding how the eigenvalues and eigenfunctions depend on the rate parameters [95], and only in these cases can one hope to gain analytical insights into the problem of extracting rate parameters from FRAP data. We turn to these cases in the following sections.

If $R(u, p) = Ku + F(u, \xi, \tau)$ the general solution of (4.23) can be written as

$$u(\xi, \tau) = \int_\Omega G(\xi - \xi', \tau) u_0(\xi') d\xi' + \int_0^t \int_\Omega G(\xi, \xi', \tau - \tau') F(u, \xi, \tau') d\xi' d\tau' \tag{4.35}$$

where $G(\xi, \xi', t)$ is the Green's function for the linear operator $L = D\nabla^2 + K$. This has the representation

$$G(\xi - \xi', \tau - \tau') = \sum_{n=0}^{\infty} \sum_{j=1}^{m} e^{\lambda_{jn}(\tau - \tau')} P_{jn} \phi_n(\xi) \phi_n(\xi') \tag{4.36}$$

This form (4.35) can be used when diffusion during bleaching and waiting periods are incorporated in the analysis.

4.1.2 A special case – diffusion and binding only

The general framework allows for first-order reactions of any type (of which there are four –*cf.* [97]), but when the only processes are diffusion of the fluorescent species and

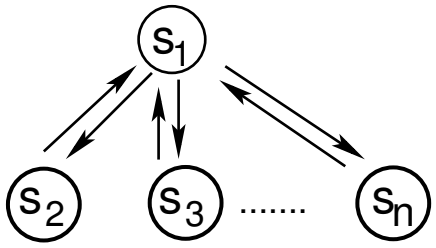


Figure 4.2: The notation for m-1 binding sites

binding to one or more independent immobile sites more can be said about the solutions. Suppose there are $m-1$ independent types of binding sites, as shown in Figure 4.2. In Appendix 4.1.3 we re-derive the known fact [79] that for a single binding site the recovery process can be modeled as a linear process, even though the binding step is nonlinear, and point out that the on- and off-rates in the resulting equations are composites of parameters in the original equations [79]. The analysis given there can be applied to the case of $m-1$ independent types of binding sites, although the amount of unbound fluorescent molecules is the solution of a higher-order polynomial when there is more than one type of site. This leads to the linear system

$$\begin{aligned} \frac{\partial u_1}{\partial \tau} &= D\nabla^2 u_1 - \sum_{k=2}^{m} B_k(u_1, u_k, \tau) \qquad \text{for} \quad \xi \in (0,1) \\ \frac{du_k}{d\tau} &= B_k(u_1, u_k, \tau) \equiv k_{k1}u_1 - k_{1k}u_k \qquad k = 2, \cdots, m \\ u(\xi, 0) &= u_0(\xi) \end{aligned} \tag{4.37}$$

Thus the matrices K and D in (4.24) take the following form.

$$K = \begin{bmatrix} -\sum_{j=2}^{m} k_{j1} & k_{12} & k_{13} & \cdots & k_{1m} \\ k_{21} & -k_{12} & 0 & \cdots & 0 \\ \vdots & & & & \\ k_{m1} & 0 & 0 & \cdots & -k_{1m} \end{bmatrix} \tag{4.38}$$

and $D = diag(D_1, 0, 0, \cdots, 0)$. The columns of the matrix K sum to zero and therefor K is singular, and if the k_{ij} are all non-zero it is irreducible and an application of the Perron-Froebenius theorem shows that the zero eigenvalue is simple. The fact that the kinetic steps are mass-preserving implies that the left eigenvector of K corresponding

to the zero eigenvalue is $(1, 1, 1, \cdots, 1)^T$, and therefore

$$\frac{\partial}{\partial \tau}(u_1 + u_2 + \cdots + u_m) = D_1 \nabla^2 u_1, \tag{4.39}$$

which reflects the fact that the total local concentration only changes due to diffusion.

The matrix K can be symmetrized, and therefore the eigenvalues are all real and non-positive. The kinetic interactions involved in binding are restricted to the reaction simplex defined by the relation $\sum_{i=1}^{m} u_i(t) = \sum_{i=1}^{m} u_i(0)$. The foregoing properties of K also imply that there is a unique steady state of the equation

$$\frac{du}{d\tau} = Ku$$

on the simplex defined by the initial condition. Eq. (4.39) shows how the simplices vary in space due to diffusion, but two useful extreme cases arise when either diffusion is slow relative to the binding or when it is rapid relative to binding. The first case arises when either the on-rate or the off-rate of every binding step is large relative to $\tau_D^{-1} = D_1/L^2$. Then a singular-perturbation analysis shows that to leading order in the small parameter $\epsilon \equiv \tau_R/\tau_D$, the kinetics reach the steady state on the reaction simplex defined by the initial data at every point in space, and these simplices then evolve slowly on the diffusion time scale.

In this limit the free and bound forms are related by

$$u_i = \frac{k_{i1}}{k_{1i}} u_1 \equiv K_i u_1, \qquad i = 2, \cdots, m. \tag{4.40}$$

to within correction terms proportional to ϵ.[2] As a result, (4.39) becomes

$$\frac{\partial}{\partial \tau}(1 + \sum_{i=2}^{m} K_i) u_1 = D_1 \nabla^2 u_1 \tag{4.41}$$

which leads to a diffusion equation for u_1 with the effective diffusion coefficient

$$D_{1,eff} = \frac{D_1}{1 + \sum_{i=2}^{m} K_i} \tag{4.42}$$

This reduction of the diffusion coefficient in the presence of rapid binding was apparently first observed by Crank [98].

[2] This can be obtained directly only if the on- and off-rates are comparable, and otherwise some additional scaling steps of the free or bound forms may be needed.

When $\tau_D << \tau_R$, *i.e.,* D_1/L^2 is much larger than the largest kinetic rate constant, one can show that to lowest order the spatial distribution of fluorescent molecules relaxes to a uniform state given by

$$\overline{u}_1(\xi,t) \sim \overline{u}_1(0) - \sum_{k=2}^{m} \int_0^t \overline{B_k}(u_1,u_k,\tau)d\tau \tag{4.43}$$

where the overbar denotes the spatial average and $\overline{B_k}(u_1,u_k,\tau) \equiv k_i^+\overline{u}_1 - k_i^-\overline{u}_i$. This solution can then be used in the binding equations to obtain an explicit expressions for the $\bar{u}_k(\tau), k = 2,\cdots,m$. Thus the evolution on the slow time scale is simply the readjustment of the fractions of bound fluorescent molecules. Since the diffusion time scale depends on the length scale of the domain, the balance between diffusion and binding can be controlled by altering the size of the bleaching region.

4.1.3 Justification of the assumption of linear kinetics in FRAP modeling

Suppose that there is one diffusing species that can bind to an immobile receptor, and let (u_1,v_1) and (u_2,v_2) be the free and bound concentrations of fluorescent and bleached molecules, respectively. These satisfy the following equations.

$$\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1 - k^+ u_1[R] + k^- u_2$$

$$\frac{\partial u_2}{\partial \tau} = k^+ u_1[R] - k^- u_2 \tag{4.44}$$

$$\frac{\partial v_1}{\partial \tau} = D_1 \bigtriangledown^2 v_1 - k^+ v_1[R] + k^- v_2$$

$$\frac{\partial v_2}{\partial \tau} = k^+ v_1[R] - k^- v_2 \tag{4.45}$$

where [R] is the concentration of free binding sites.

The sums $w_1 = u_1 + v_1$ and $w_2 = u_2 + v_2$ satisfy

$$\frac{\partial w_1}{\partial \tau} = D_1 \bigtriangledown^2 w_1 - k^+ w_1[R] + k^- w_2 \tag{4.46}$$

$$\frac{\partial w_2}{\partial \tau} = k^+ w_1[R] - k^- w_2. \tag{4.47}$$

If adequate time has elapsed before bleaching begins, one can assume that the system is at steady state prior to bleaching, and further assume that bleaching does not perturb the binding reactions, it merely substitutes a bleached for a fluorescent molecule. Then throughout the bleaching process $w_1 = w_1^s$, the steady-state value of the free concentration, and the total free concentration of binding sites is given by

$$R = R_T \frac{\hat{K}}{\hat{K} + w_1^s}, \tag{4.48}$$

where $R_T = R + w_2$ is the total concentration (free and occupied) of binding sites and $\hat{K} = k^-/k^+$. Similarly, the concentration of bound sites is

$$w_2^s = R_T \frac{w_1^s}{\hat{K} + w_1^s}.$$

Let $W = w_1 + w_2$ be the total concentration of all forms of the molecules, which here is assumed constant pointwise in space and time. Then it follows that w_1^s is the solution of a quadratic equation whose coefficients involve K, W and R_T, and the solution of this quadratic then leads to the free site concentration (4.48), and (4.44) can be written as

$$\begin{aligned} \frac{\partial u_1}{\partial \tau} &= D_1 \bigtriangledown^2 u_1 - k^+ u_1 + k^- u_2 \\ \frac{\partial u_2}{\partial \tau} &= k^+ u_1 - k^- u_2 \end{aligned} \tag{4.49}$$

where $k^+ = k^+([R_T] - w_2^s)$ and $k^- = k^-$. This is the linear system widely-used in this context, and that data **what data** is used to estimate the coefficients involved. However, the constant k^+ is a composite constant the involves three separate constants in an experiment, and two other independent measurements are needed to determine the three independent constants.

To define the initial conditions, let $I_0(\xi) = u_{10}(\xi) + u_{20}(\xi)$ be the initial concentration of the fluorescent molecules in space at the onset of the recovery phase of an experiment, which is usually measured directly from the first postbleaching image. If we assume that bleaching in the ROI is instantaneous (*i.e.* $T_1 = 0$), then the initial condition for equation (4.49) is

$$u_{10} = \frac{I_0(\xi)}{1 + K_d^{-1}} \qquad \text{and} \qquad u_{20} = \frac{I_0(\xi)}{1 + K_d}.$$

where $K_d \equiv k^-/k^+$ is the dissociation constant. In the simulations described in the text these are piecewise constant functions. If bleaching is not instantaneous then diffusion alters the initial condition, and the initial conditions for the recovery period must be computed using the equations applicable to the interval $[T_0, T_0 + T_1]$ given in Section **??**.

4.1.4 Analysis of eigenvalues and eigenvectors for a two-component system

In the following, we use eigenvalues and eigenvectors to analyze how FRAP data is determined by the parameters in the model. Let

$$Q = K - \alpha_n^2 D = \begin{pmatrix} -k^+ - \alpha_n^2 D_1 & k^- \\ k^+ & -k^- \end{pmatrix}$$

and

$$\gamma = k^+ - k^- + \alpha_n^2 D_1$$

Then the eigenvalues of the matrix Q are:

$$\lambda_{1n} = -\frac{1}{2}(k^+ + k^- + \alpha_n^2 D_1) + \sqrt{\gamma^2 + 4k^+k^-})$$

$$\lambda_{2n} = -\frac{1}{2}(k^+ + k^- + \alpha_n^2 D_1) - \sqrt{\gamma^2 + 4k^+k^-})$$

$$\lambda_{10} = -(k^+ + k^-), \lambda_{20} = 0.$$

The averaged fluorescent intensity can be written as

$$\overline{u_1} + \overline{u_2} \equiv (\overline{u_1}_0 + \overline{u_2}_0) + \sum_{n=1}^{\infty} g(n)(\overline{u_1}_n + \overline{u_2}_n) = y_{10} + y_{20} + \sum_{n=1}^{\infty} g(n)(d_{1n}e^{\lambda_{1n}t} + d_{2n}e^{\lambda_{2n}t}) \tag{4.50}$$

where

$$u_{10} = \frac{k^+y_{10} - k^-y_{20}}{k^+ + k^-}e^{\lambda_{10}t} + \frac{k^-y_{10} + k^-y_{20}}{k^+ + k^-}$$

$$u_{20} = \frac{k^-y_{20} - k^+y_{10}}{k^+ + k^-}e^{\lambda_{10}t} + \frac{k^+y_{10} + k^+y_{20}}{k^+ + k^-}$$

Note that

$$\lambda_{1n}\lambda_{2n} = \alpha_n^2 D_1 k^-$$

$$d_{1n} = \frac{y_{1n}+y_{2n}}{2} - \frac{\alpha_n^2 D_1(y_{2n}-y_{1n})}{2\sqrt{\gamma^2+4k^+k^-}} - \frac{(k^++k^-)(y_{1n}+y_{2n})}{2\sqrt{\gamma^2+4k^+k^-}}$$

$$d_{2n} = \frac{y_{1n}+y_{2n}}{2} + \frac{\alpha_n^2 D_1(y_{2n}-y_{1n})}{2\sqrt{\gamma^2+4k^+k^-}} + \frac{(k^++k^-)(y_{1n}+y_{2n})}{2\sqrt{\gamma^2+4k^+k^-}}$$

Since

$$y_{10}+y_{20} = \frac{1}{N}\sum_{k=1}^{N} I_0(k)$$

$$y_{1n}+y_{2n} = \frac{1}{N}\sum_{k=1}^{N} I_0(k)(cos\frac{n\pi}{N}(k-1) + cos\frac{n\pi}{N}k)$$

Note that it is positive and independent of the parameters

$$y_{2n}-y_{1n} = \frac{k^+-k^-}{k^++k^-}\frac{1}{N}\sum_{k=1}^{N} I_0(k)(cos\frac{n\pi}{N}(k-1) + cos\frac{n\pi}{N}k) = \frac{k^+-k^-}{k^++k^-}(y_{1n}+y_{2n})$$

If we define

$$f(n) = g(n)\frac{y_{1n}+y_{2n}}{2}$$

we have

$$\overline{u_1}+\overline{u_2} = \frac{1}{N}\sum_{k=1}^{N} I_0(k) + \sum_{n=1}^{\infty} f(n)(c_{1n}e^{\lambda_{1n}t} + c_{2n}e^{\lambda_{2n}t}) \tag{4.51}$$

where

$$c_{1n} = 1 - \frac{\alpha_n^2 D_1 \frac{k^+-k^-}{k^++k^-}}{\sqrt{\gamma^2+4k^+k^-}} - \frac{k^++k^-}{\sqrt{\gamma^2+4k^+k^-}}$$

$$c_{2n} = 1 + \frac{\alpha_n^2 D_1 \frac{k^+-k^-}{k^++k^-}}{\sqrt{\gamma^2+4k^+k^-}} + \frac{k^++k^-}{\sqrt{\gamma^2+4k^+k^-}}$$

Note that $c_{1n}, c_{2n} > 0$, $\lambda_{1n}, \lambda_{2n} < 0$.

As n increases, c_{1n} increases, c_{2n} decreases, both λ_{1n} and λ_{2n} decrease (algebraically). Moreover, as $n \to \infty$, $\lambda_{1n} \to -\alpha_n^2 D_1$, $\lambda_{2n} \to -k^-$, and if $k^+ > k^-$,

$$c_{1n} \to 2\frac{k^-}{k^++k^-}, c_{2n} \to 2\frac{k^+}{k^++k^-}$$

if $k^+ < k^-$,

$$c_{1n} \to 2\frac{k^+}{k^+ + k^-}, c_{2n} \to 2\frac{k^-}{k^+ + k^-}$$

and if $k^+ = k^-$,

$$c_{1n} \to 1, c_{2n} \to 1.$$

One may first estimate the eigenvalues and corresponding eigenvectors from FRAP data based on equation (4.51). Then by using the relationship between paramters and eigenvalues and eigenvectors, diffusion coefficients and kinetic parameters may be estimated. Moreover, the quantitative analysis of eigenvalues may lead to more insights into parameter estimation. These aspects can be explored in the future.

4.1.5 The computational FRAP setup

To investigate different approaches to the analysis of FRAP data, we use a computational model to generate the FRAP data, which facilitates evaluation of the effect of experimental parameters such as the size of the ROI and the time of observation. In our simulations the FRAP data is generated by using the dimensionless form of equations as (4.24) described except that real time is used, *i.e.,* without scaling.

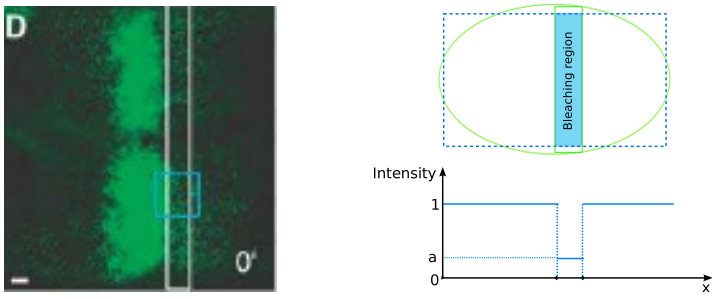


Figure 4.3: Left: A region of the wing disc that is scanned (from [15] with permission). Green indicates GFP-labelled Dpp, the white box is the ROI, and the scale bar is $10\mu m$, Right: (top) The computational approximation of the disc as an ellipse and the rectangular ROI, (bottom) the initial data along a one-dimensional cross-section of the region.

The geometry of the tissue can be described as an approximately rectangular compartment, and when the bleaching is done in a stripe as in Figure 4.3 ([15, 28]), the concentration of fluorescent molecules varies primarily in one direction (x above), variations are negligible in the y direction under no-flux boundary conditions in that direction, and are typically recorded as the maximum projection in the z direction. Under these conditions the data analysis can be reduced to a one dimensional problem. Accordingly, in this paper the mathematical formulation and all simulations are done only in 1D. However, the conclusions and numerical procedure in our 1D system can be applied and easily extended to 2D and 3D systems.

FRAP recovery data is the spatial average of the sum of free and bound fluorescence in the observation region, which may be smaller that the ROI, so suppose that u_1 represents the free molecule and other species $u_i, i > 1$ are bound species. Then the fluorescence intensity as a function of time τ is described as

$$\overline{u_i(\tau)} = \frac{1}{l_R - l_L} \int_{l_L}^{l_R} u_i(\xi, \tau) d\xi \qquad frap(\tau) = \sum_{i=1}^{m} \overline{u_i(\tau)} \tag{4.52}$$

where $1 > l_R = \frac{L_R}{L} > l_L = \frac{L_L}{L} > 0$. Here L is the overall size of the system[3] , and L_L and L_R define the observation region. The problem of parameter estimation can be considered as an inverse problem or optimization process, and the algorithm underlying our method is shown in Figure 4.4, and the mathematical details are given in the next section. In practice, the number of terms (M) that are retained in the eigenfunction expansion is determined by setting a threshold for changes and increasing the number of terms until the parameter estimates do not change within the threshold.

Another issue that arises in parameter estimation concerns the initial condition for the recovery equation. It is well known that during the bleaching phase bleached and unbleached molecules diffuse in and out of the ROI, and this results in a transition region with various levels of photobleaching between the bleached and unbleached region. The size of the intermediate region depends on the molecular diffusivity, size of the bleaching region and the bleaching time. It may have substantial effects on the estimation of molecular motility and binding kinetics. In this paper, we assume that the bleaching

[3] The computations that follow are based on a diffusion coefficient of $10\mu^2/sec$ and $L = 200\mu$, unless otherwise noted.

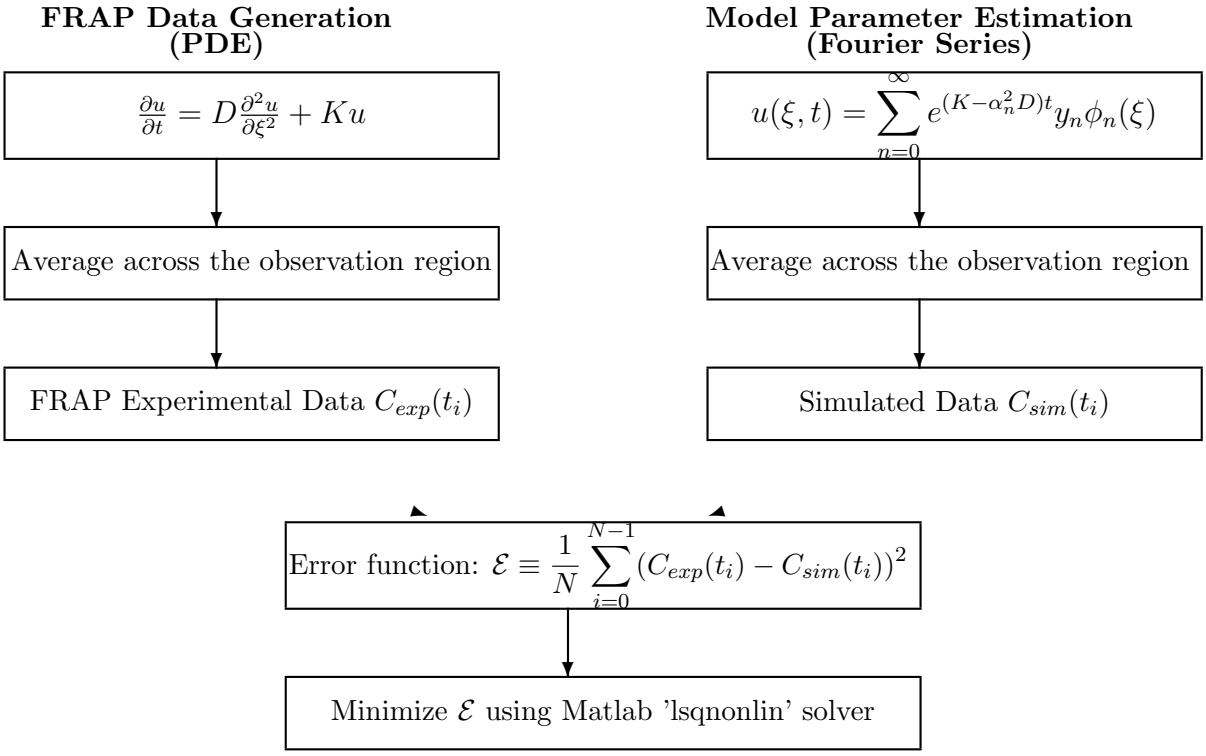


Figure 4.4: The computational algorithm used throughout the paper.

process is instantaneous and use piece-wise constant data as shown in Figure 4.3 as the initial data for simulations unless specified otherwise.

4.2 Theoretical models for FRAP data generation

We examine two classes of theoretical models – which are summarized in Table 4.1 – one in which the system is closed and a second one in which there is a specified flux at the boundary . We briefly present the governing equations and some basic results in this section to make it easier for the reader to compare the models. In the following section we use the models in computational studies to show how parameter estimates depend on the model used and how they can be improved with different protocols.

	FRAP models for closed systems	FRAP models with boundary fluxes
1-component model	**Model B1 (4.2.1)** $\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=0} = 0$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=1} = 0$	**Model B2 (4.2.3)** $\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1 - k^d u_1$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=0} = J$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=1} = 0$
2-component model	**Model 1 (4.2.1, 4.1.1-4.1.5, 4.2.1, 4.2.2)** $\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1 - k^+ u_1 + k^- u_2$ $\frac{\partial u_2}{\partial \tau} = k^+ u_1 - k^- u_2$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=0} = 0$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=1} = 0$	**Model 3 (4.2.2, 4.1.6, 4.2.3)** $\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1 - k^+ u_1 + k^- u_2$ $\frac{\partial u_2}{\partial \tau} = k^+ u_1 - k^- u_2 - k^d u_2$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=0} = J$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=1} = 0$
3-component model	**Model 2 (4.2.1, 4.1.6, 4.2.2)** $\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1 - k_1^+ u_1 + k_1^- u_2 - k_2^+ u_1 + k_2^- u_3$ $\frac{\partial u_2}{\partial \tau} = k_1^+ u_1 - k_1^- u_2$ $\frac{\partial u_3}{\partial \tau} = k_2^+ u_1 - k_2^- u_3$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=0} = 0$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=1} = 0$	**Model 4 (4.2.2, 4.1.6, 4.2.3)** $\frac{\partial u_1}{\partial \tau} = D_1 \bigtriangledown^2 u_1 - k^+ u_1 + k^- u_2$ $\frac{\partial u_2}{\partial \tau} = k^+ u_1 - k^- u_2 - k^i u_2$ $\frac{\partial u_3}{\partial \tau} = k^i u_2 - k^d u_3$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=0} = J$ $- D_1 \frac{\partial u_1}{\partial \xi}\vert_{\xi=1} = 0$

Table 4.1: Summary of the models for the following analysis and simulations. Numbers in parentheses refer to the subsections in which the corresponding model is analyzed (3.x.x) and the computational results are given (4.x.x) .

4.2.1 FRAP models for closed systems

Model 1: one mobile species and one type of binding site

In the simplest model there is one diffusible fluorescent species U_1 that can bind to an immobile receptor R to produce the complex U_2, according to the reaction $U_1 + R \underset{k^-}{\overset{k^+}{\rightleftharpoons}} U_2$. As stated earlier, the recovery process can be modeled as a linear process and this leads to a solution of the form (4.28), where u_1, u_2 denote the concentration of unbound and bound fluorescent molecules,resp.,[4] , and

$$K = \begin{pmatrix} -k^+ & k^- \\ k^+ & -k^- \end{pmatrix} \qquad D = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}$$

and $\phi_n = cos(n\pi\xi)$, and $\alpha_n^2 = (n\pi)^2$. Both K and D are singular, the former due to the conservation condition[5] .

We denote by I_0 the total initial concentrations and show in Appendix 4.1.3 that the initial fractions are

$$u_{10} = \frac{I_0}{1 + K_d^{-1}} \quad \text{and} \quad u_{20} = \frac{I_0}{1 + K_d}.$$

where $K_d \equiv k^-/k^+$ is the dissociation constant. In addition, $k^+ = k^+([R_T] - u_2^{ts})$ and $k^- = k^-$, where R_T, u_2^{ts} denote the total binding sites and the total concentration of bound molecules, respectively. Note that these two quantities are constant throughout the FRAP experiment. The initial conditions are

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}_{\tau=0} = \sum_{n=0}^{\infty} \begin{pmatrix} y_{1n} \\ y_{2n} \end{pmatrix} cos(n\pi\xi) = \begin{pmatrix} u_{10} \\ u_{20} \end{pmatrix} \tag{4.53}$$

where u_{10}, u_{20} are the initial concentrations of unbound and bound fluorescent molecules after photobleaching, respectively. The coefficients y_{1n}, y_{2n} are given by

$$y_{10} = \int_0^1 u_{10}(x)dx \qquad y_{1n} = 2\int_0^1 u_{10}(x)cos(nx)dx$$

and

$$y_{20} = \int_0^1 u_{20}(x)dx \qquad y_{2n} = 2\int_0^1 u_{20}(x)cos(nx)dx.$$

The average fluorescence intensities of bound and unbound molecules across the ROI during the recovery phase are given by

$$\begin{aligned} \begin{pmatrix} \overline{u}_1 \\ \overline{u}_2 \end{pmatrix} &= \frac{1}{\delta^-} \int_{l_L}^{l_R} \begin{pmatrix} u_1(\tau,\xi) \\ u_2(\tau,\xi) \end{pmatrix} d\xi \\ &= e^{K\tau} y_0 + \frac{1}{\pi\delta^-} \sum_{n=1}^{\infty} \frac{1}{n} [sin(n\pi l_R) - sin(n\pi l_L)] e^{[K - D(n\pi)^2]\tau} y_n \qquad (4.54) \\ &= e^{K\tau} y_0 + \frac{1}{\pi\delta^-} \sum_{n=1}^{\infty} \frac{1}{n} [sin\frac{(n\pi\delta^-)}{2} cos\frac{(n\pi\delta^+)}{2}] e^{[K - D(n\pi)^2]\tau} y_n \qquad (4.55) \end{aligned}$$

[4] Since the total concentration of fluorescent molecules is constant one can alternatively regard the u_i as fractions

[5] We remind the reader that τ represents real time, and thus the diffusion coefficients and the kinetic constants for linear steps have dimensions $time^{-1}$.

where $y_n = (y_{1n}, y_2)^T$, and $\delta^{\pm} = l_R \pm l_L$. Finally, the average fluorescence intensity in time, *i.e.,* the FRAP data, is given by

$$frap(\tau) = \overline{u_1}(\tau) + \overline{u_2}(\tau). \tag{4.56}$$

The parameter space for this two-component diffusion-binding model can be divided into several regimes that reflect different balances between the component processes. There are three different

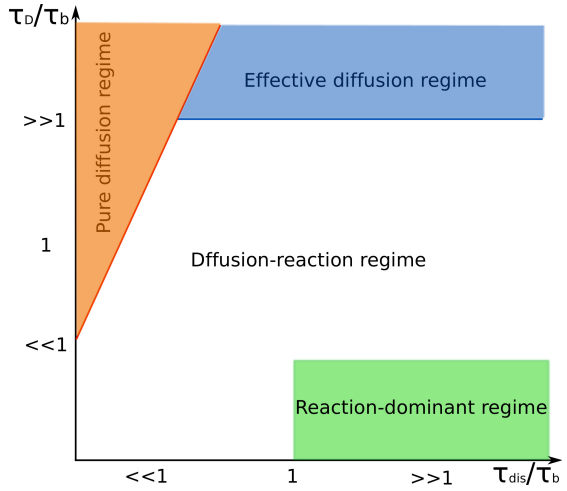


Figure 4.5: Different regimes in the parameter space for a diffusion binding model

characteristic time scales in the system: the diffusion time $\tau_D = (\delta^-)^2/D$ where δ^- is the width of the bleaching region, the binding time $\tau_b = 1/k^+$, and the dissociation time $\tau_{dis} = 1/k^-$. In the initial stages of recovery the primary effect of diffusion is from a small region adjacent to the bleached region into the bleached region, and therefore we use a characteristic diffusion time based on the width of the ROI. When binding is weak, *i.e.,* $\tau_{dis}/\tau_b << 1$, and the diffusion is not much faster than binding, *i.e.,* above $\tau_D/\tau_b << 1$, the recovery process can be described by pure diffusion, in which case the parameters are located in the pure diffusion regime of Figure 4.5. When binding is tighter, *i.e.,* $\tau_{dis}/\tau_b >> 1$, and much faster than diffusion, *i.e.,* $\tau_D/\tau_b >> 1$, the recovery process can be approximated as a diffusion process but with an effective diffusion coefficient $D_{eff} = D/(1 + K_d^{-1})$, and the parameters are located in the effective diffusion regime. When the binding is tighter, i.e. right $\tau_{dis}/\tau_b \geq 1$, and diffusion is much faster than binding $\tau_D/\tau_b << 1$, the recovery time is determined by the reaction process when the parameters are located in reaction-dominant regime. Outside the three special regimes, the remainder of the parameter space is called diffusion-reaction regime. For the upcoming discussion, it is worth noticing that the diffusion characteristic time can be manipulated by the spatial scale of the bleaching region, which could lead to a shift of the location of parameters from one regime to another.

Model 2: one mobile species and multiple types of binding sites

To illustrate the different scales that arise with multiple binding sites, consider two sites, for which K and D can be obtained from Appendix (4.1.2). When the binding steps are much faster than diffusion this is the case treated earlier, and the effective diffusion coefficient given at (4.42) takes the form $D_{eff} = D_1/(1 + K_2 + K_3)$, where $K_2 = k_1^+/k_1^-, K_3 = k_2^+/k_2^-$.

Suppose however that binding to the second type reaches a quasi-steady state rapidly compared to binding to the first and to diffusion. Then local equilibrium gives rise to $u_3 = K_3 u_1$, and by adding the governing equations, we have

$$\frac{\partial}{\partial \tau}(u_1 + u_3) = D \bigtriangledown^2 u_1 - k_1^+ u_1 + k_1^- u_2$$

By dividing the equations for u_1 and u_2 by $1 + K_3$ one can reduce to a model with a single binding site, which reads

$$\frac{\partial u_1}{\partial \tau} = D_{rd} \bigtriangledown^2 u_1 - k^+ u_1 + k^- u_2' \tag{4.57}$$

$$\frac{\partial}{\partial \tau}(u_2') = k^+ u_1 - k^- u_2' \tag{4.58}$$

where

$$u_2' = \frac{u_2}{1 + K_3}, \quad D_{rd} = \frac{D_1}{1 + K_3}, \quad k^+ = \frac{k_{21}}{1 + K_3}, \quad \text{and} \quad k^- = \frac{k_{12}}{1 + K_3}.$$

These parameters are what are obtained in the parameter estimation when the two-site model is reduced to a model with a single binding site as above, and illustrate again that the estimated parameters may be complex functions of the more fundamental parameters.

4.2.2 FRAP models with boundary fluxes

Model 3: Influx, diffusion, binding and decay

The parameters measured from different time and space scales may reflect different integration of biological processes such as production, internalization and decay in addition to diffusion and binding discussed above. In biological systems at tissue level in the long run, the FRAP recovery is amalgamation of these processes. Our approach for parameter estimation can be extended to the model with more than diffusion and binding. Here we derive the analytical solution for these models, and show that more parameters such as internalization rate and decay rate can be estimated from FRAP in addition to diffusion coefficients and binding/unbinding rates.

The time-dependent solution formula (4.28) can be used directly for Model 3 with influx, diffusion, binding and decay, wherein

$$K = \begin{pmatrix} -k^+ & k^- \\ k^+ & -(k^- + k^d) \end{pmatrix} \qquad D = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}$$

For the steady state of u_1^s , the solution formula (4.15) applies with $\delta = \sqrt{k^d/D_1}$ where $k^d = k^+ - k^+ k^-/(k^- + k^d)$, and $u_2^s(\xi) = k^+/(k^- + k^d) u_1^s(\xi)$.

Model 4: Influx, diffusion, binding, internalization and decay

When the model contains more processes, as in Model 4, the solution form (4.28) and (4.15) can be used. In those equations, k^i is the internalization rate constant and k^d is the decay rate constant, $\delta = \sqrt{\widetilde{k^d}/D_1}$ where $\widetilde{k^d} = k^+k^i/(k^- + k^i)$, and

$$K = \begin{pmatrix} -k^+ & k^- & 0 \\ k^+ & -(k^- + k^i) & 0 \\ 0 & k^i & -k^d \end{pmatrix} \qquad D = \begin{pmatrix} D_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

4.3 Recovery models for parameter estimation

The parameter estimates are extracted from FRAP by fitting a specified model with FRAP data, using the algorithm described earlier. However, even if a good data fitting is achieved, little biological information can be inferred from the estimates of parameters without careful examination and analysis. The potential problems behind a good fit to the recovery curve can be explored from two distinct aspects. The first one is how accurate the estimates are, assuming that an appropriate model has been used for data fitting. To address this we propose several methods used later, such as choosing the appropriate observation time, reducing the bleaching size and using spatial FRAP to improve parameter estimation when the model used for estimation is identical to the one used for FRAP data generation. The other problem in FRAP is how well the model reflects the actual processes involved in a FRAP experiment. We will show the impact of the reduced model on estimates of parameters, and explain how to relate the complex steps in the realistic model to the higher level description in the reduced model. In addition, we also show how reducing the bleaching size could help to formulate an appropriate FRAP model. The models used in our simulations are summarized in Table 4.1.

4.3.1 Identical recovery model - Methods to improve parameter estimation

In the six parts of this subsection the model to estimate the parameters is the same as that to generate FRAP data. With complete knowledge of the parameters, we will show that even though the recovery curve fit is good, the estimates may not be accurate. Moreover, we will discuss the ways to improve parameter estimation via the simple diffusion-binding model (4.1.1-4.1.5), and show that these methods can be extended to parameter estimation in other models which are common and widely used in studying the mechanism of pattern formation(4.1.6). In the following simulations, we use a piece-wise constant initial condition for FRAP recovery, which is widely-used in existing FRAP models and also valid in reality if there is little fluorescence recovery in the bleaching region before observation. For the record, a wide range of initial guesses including the true values of parameters have been used, and they result in similar estimations, which provides a foundation for our conclusions.

The observation region versus the bleaching region - Model 1

We find that taking a subdomain within the bleaching region (ROI) as the observation region (OR) considerably improves the estimation, because of the Gibbs effect that results from the piece-wise constant initial condition [6] . Hereafter we fix the number of terms M = 1000 for minimization (Figure 4.6), and we also fix the ROI at $l_R - l_L = .07$ and let d be the distance between the boundary of the ROI and that of the OR, as shown in Figure 4.7. The quality of estimates is good even though d is small (Table 4.2). On the other hand, good estimates can be obtained with a small number of terms by using a smaller subdomain of the bleaching region as the observation region (results not shown here), which make the computation more efficient, especially in 2D or 3D cases.

Because the initial condition is set as piece-wise constant in all the following simulations, the default observation region will always be set smaller than the bleaching region – specifically we set $d = 0.2(l_R - l_L)$. Thus the size of the bleaching region is 1.4 times that of observation region, and the truncation is at $M = 1000$ terms in order to eliminate the Gibbs effect on estimation.

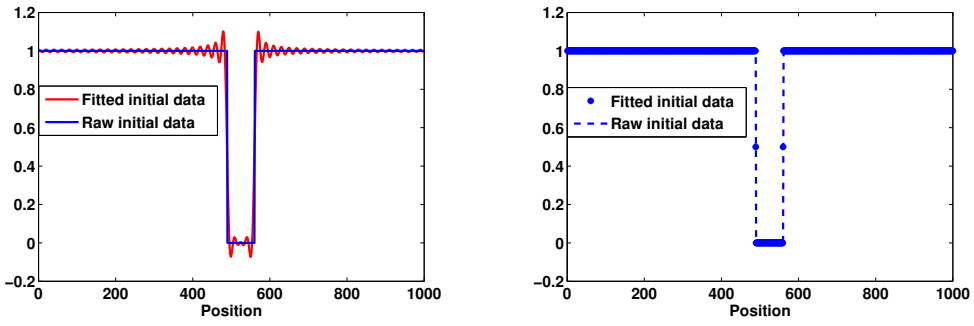


Figure 4.6: The Gibbs effect in representing the initial data. The sum is truncated at M terms left: M=100, right: M=1000

Estimation of parameters in different regimes - Model 1

To test our approach for parameter estimation, the diffusion coefficient, binding and unbinding rates have been estimated in different regimes shown in Figure 4.5: pure diffusion regime (diffusion and weak binding), effective diffusion regime (diffusion and fast binding), reaction-dominant regime (fast diffusion and binding), and diffusion-reaction regime. We can accurately estimate the diffusion coefficient and the binding and release rates in both reaction-dominant regime and diffusion-reaction regime (Table 4.3)[7]

[6] Other basis functions that require fewer terms could be used, but the eigenfunction expansion is most commonly used and we do so here.

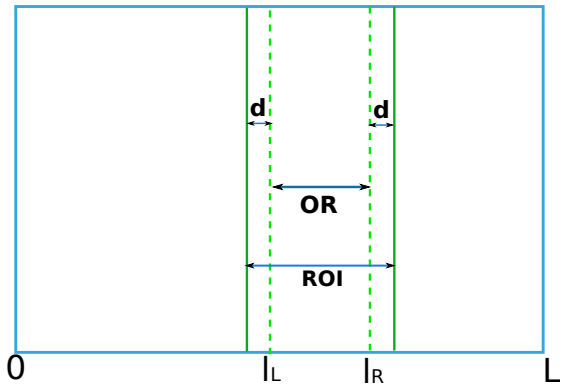


Figure 4.7: The relationship between the observation region and the bleaching region

Distance	Time	Estimates for M=100			Estimates for M=1000		
d	T	D	k^+	k^-	D	k^+	k^-
0	100	4.9848e-5	1.4840e-2	1.3079e-3	2.2217e-4	1.0431e-2	1.0140e-3
	1000	2.4006e-4	1.8935e-2	1.0870e-3	2.3858e-4	1.0527e-2	1.0103e-3
0.01	100	2.8206e-4	9.7154e-3	9.8591e-4	2.5236e-4	1.0019e-2	1.0000e-3
	1000	2.6560e-4	9.5765e-3	9.8642e-4	2.5092e-4	9.9987e-3	9.9964e-4

Table 4.2: The influence of the choices of the observation region within the bleaching region and the observation time on the estimates of parameters in the diffusion-binding model. The centered observation region has a fixed width of 0.05, and the bleaching region is enlarged by increasing d. The FRAP data is generated by the same model with parameters $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1 \times 10^{-2}$ sec^{-1}, $k^- = 1 \times 10^{-3}$ sec^{-1}.

. However, when the parameters are in either the pure diffusion regime or effective diffusion regime, because it has been proven that the FRAP curve can be well described by the diffusion only [79], we find that it is difficult to estimate all the three parameters accurately at the same time. How to improve the parameter estimation in these regimes will be discussed later.

[7] Here and hereafter we use an error tolerance of 10^{-10}, but larger tolerances produce very similar results.

Regime	True values of parameters			Estimates of parameters		
	D	k^+	k^-	D	k^+	k^-
Diffusion-reaction	2.5e-4	1e-2	1e-3	2.5236e-4	1.0019e-2	1.0000e-3
Pure-diffusion	2.5e-4	1e-3	1e-1	2.4999e-4	6.0222e-4	7.3957e-2
Effective-diffusion	2.5e-5	1	1e-1	3.9024e-6	4.6544e-2	6.4715e-2
Reaction-dominant	1e-2	1e-5	1e-6	1.0005e-2	1.0008e-5	1.0051e-6

Table 4.3: In conventional FRAP, the estimates of parameters are accurate in the diffusion-reaction and the reaction-dominant regimes, but not in the pure-diffusion and effective-diffusion regimes. All the results are simulated by using the observation time of $T = 1000$ sec. Default values are used for the size of the bleaching and observation regions.

Appropriate observation time - Model 1

One of our most intriguing discoveries is the importance of the choice of observation time in parameter estimation, especially when the parameters are located in the regime where effective diffusion applies. When the binding process is relatively faster than diffusion process, it is more difficult to estimate all three parameters, because FRAP data can be well interpreted by effective diffusion. We find that using FRAP data collected in an appropriate observation time period gives rise to quality estimates (Table 4.4). Moreover, the appropriate time is determined by the characteristic time of the dissocation process.

Mathematically, the effective diffusion is based on the assumption that the binding and dissociation processes are much faster and equilibrate before diffusion plays a significant role. Physically, the fluorescence recovery in FRAP is essentially produced by the unbound fluorescent molecules diffusing into the bleaching region and binding, which cannot happen until the bound bleached molecules in the bleached region are dissociated from the binding sites. Therefore, if the observation time is shorter than the characteristic time of the dissociation process, *i.e.,* the recovery data is obtained before dissociation ensues, the binding and dissociation processes haven't reached equilibrium, and only diffusion and binding contributes to the dynamics of recovery data, which is distinguished from the effective diffusion process. After that, the process of recovery is reduced to an effective diffusion process, in which case it is more difficult to estimate all

the three parameters from the standard recovery curve. In real FRAP experiments, the time interval for data collection might not be small enough as it is in simulation. The data also contains noise at some level. Thus, the appropriate observation time might be different from what it is in theory.

True values of parameters			Observation time	Estimates of parameters		
D	k^+	k^-	$T(sec)$	D	k^+	k^-
2.5e-4	1	0.1	100	1.9132e-4	0.7542	0.1013
			10	2.5689e-4	0.9684	0.09649
			5	2.6533e-4	1.0010	0.09653
2.5e-4	1	1	100	2.0315e-4	0.5236	0.8367
			10	2.1076e-4	0.5543	0.8183
			5	2.3477e-4	0.7100	0.8379
2.5e-4	0.1	1	100	2.3841e-4	0.02947	0.6054
			10	2.4728e-4	0.04320	0.5689
			5	2.5934e-4	0.08219	0.6980

Table 4.4: Choosing an appropriate observation time results in better estimation in the effective diffusion regime. The FRAP data is generated with the default sizes of the bleaching and observation regions.

Size of the bleaching region - Model 1

In reality, because the parameters such as the dissociation rate are unknown, it is difficult to determine *a priori* what the appropriate observation time is for parameter estimation. However, the time scale of diffusion is determined by the spatial scale of the bleaching region. Therefore, by changing the size of bleaching region, we can change the balance of diffusion and binding processes, *i.e.,* we can change the relative location of parameters in parameter space. Since binding and dissociation are local activities that are independent of spatial scale, we expect that reducing the size of the bleaching region to make the characteristic diffusion time smaller than the binding time. That is, relocating the parameters from the effective-diffusion regime to the reaction-diffusion regime in Figure 4.5 can improve the estimates when binding is faster than diffusion with the default bleaching size, and this is validated by our simulation results in Table 4.5. The estimates of parameters are more accurate, and do not depend on the length

of observation time when the size of the bleaching region is reduced [8] .

True values of parameters			Bleaching region	Observation T	Estimates of parameters		
D	k^+	k^-	BR	$T(sec)$	D	k^+	k^-
2.5e-4	1	0.1	Default	100	1.9132e-4	0.7542	0.1013
				10	2.5689e-4	0.9684	0.09649
			Reduced size	100	2.4228e-4	0.9666	0.09999
				10	2.4364e-4	0.9666	0.09990
2.5e-4	1	1	Default	100	2.0315e-4	0.5236	0.8367
				10	2.1076e-4	0.5543	0.8183
			Reduced size	100	2.4891e-4	0.9912	0.9997
				10	2.4888e-4	0.9916	0.9999
2.5e-4	0.1	1	Default	100	2.3841e-4	0.02947	0.6054
				10	2.4728e-4	0.04320	0.5689
			Reduced size	100	2.4987e-4	0.09936	0.9995
				10	2.4988e-4	0.09935	0.9994

Table 4.5: The estimates are improved by reducing the size of the bleaching region so as to change the time scale of diffusion relative to that of binding.

Moreover, our approach to improve estimates is also applied to the case where the binding is weak. Because weak binding makes little contribution to the FRAP recovery, the diffusion-binding model can be approximated by pure diffusion when the binding is very weak, which makes it very difficult to estimate the association and dissociation rates for weak binding. However, our results suggest that by reducing the bleaching region to make the diffusion time smaller than the binding time, we can achieve good estimates of binding/unbinding rates as well as the diffusion coefficient (Table 4.6). In addition, the way of reducing the size of the bleaching region affects the estimation in the same way as increasing the diffusion coefficient as their characteristic time of diffusion is similar (Table 4.6).

To further explore how either reducing the size of the bleaching region or increasing the diffusion coefficient improves parameter estimation, first the recovery of unbound and bound fluorescent molecules are observed separately (Figure 4.8) and compared when the size of the bleaching region is reduced. We find that reducing the size of

[8] The values 0.007 (0.005) for the bleaching (observation) regions are used hereafter whenever the size is reduced, unless stated otherwise.

True values of parameters			Bleaching size	Observation T	Estimates of parameters		
D	k^+	k^-	BR	$T(sec)$	D	k^+	k^-
2.5e-4	1e-3	1e-1	Default	1000	2.4999e-4	6.0222e-4	7.3957e-2
				100	2.5098e-4	6.5301e-4	6.9036e-2
				10	2.5844e-4	3.9193e-3	1.5256e-1
			Reduced size	1000	2.5003e-4	9.9063e-4	9.9615e-2
				100	2.5007e-4	9.9101e-4	9.9574e-2
				10	2.5049e-4	1.0051e-3	9.9828e-2
1e-2 (Increased)	1e-3	1e-1	Default	1000	1.0006e-2	1.0045e-3	1.0021e-1
				100	1.0006e-2	1.0045e-3	1.0021e-1
				10	1.0013e-2	1.0176e-3	1.0085e-1

Table 4.6: Estimates are better when the size of the bleaching region is smaller or the diffusion coefficient is larger.

the bleaching region speeds up the recovery of unbound fluorescence more than that of bound fluorescence, as shown by comparing (b) to (a) and (d) to (c) in Figure 4.8, which makes the recovery of unbound fluorescence relatively quicker in the bleaching region in comparision to bound fluorescence. Therefore, the recovery is dominated by diffusion initially, and by the binding process later in time, which make the estimation of three parameters more feasible. Similarly, the FRAP recovery in the bleaching region is more uniform when the diffusion coefficient is increased since the spatial nonuniformity relaxes rapidly (results not shown here), which suggests that averaging the data over the bleaching region has less impact on estimation for fast diffusion than slow diffusion.

Exploiting spatial information in FRAP – Model 1

The effect of either reducing the size of the bleaching region or increasing the diffusion coefficient on parameter estimation in FRAP suggests that the standard way to average data across the whole bleaching region to get FRAP data loses information, which is more prominent when the bleaching region is large or the diffusion coefficient is small. Therefore, we propose a new way to use FRAP data. Insteading of averaging the data, first we try to use the FRAP data in space for parameter estimation, which turns out to greatly improve the estimates as reducing the size of the bleaching region does as

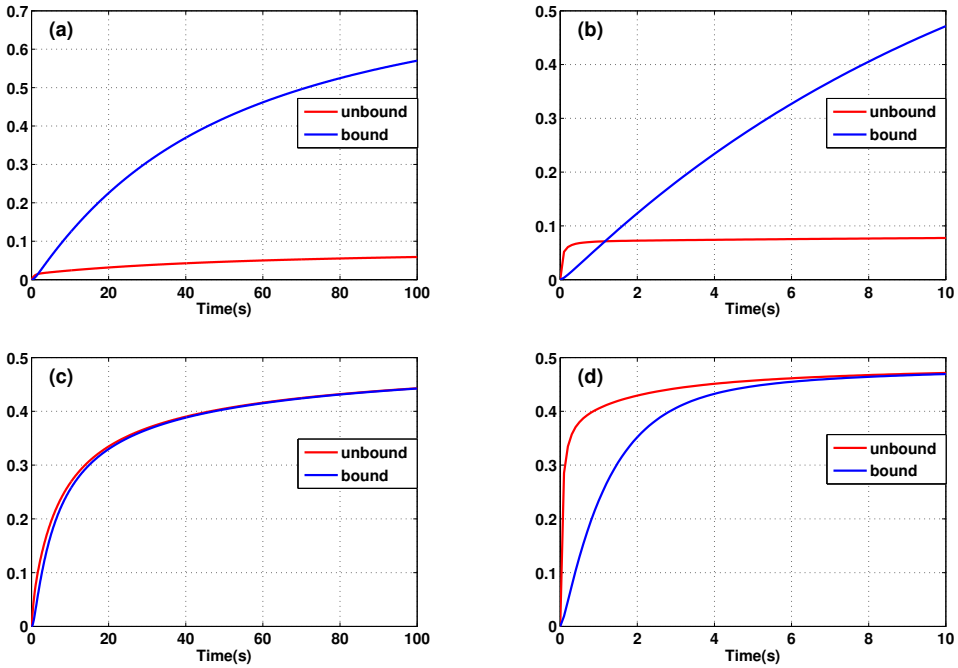


Figure 4.8: The effect of reducing the size of the bleaching region on the recovery of bound and unbound molecules. (a) and (b): The recovery curves are generated with $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1$ sec^{-1}, $k^- = 0.1$ sec^{-1}; (c) and (d): The recovery curves are generated with $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1$ sec^{-1}, $k^- = 1$ sec^{-1}; (a) and (c): with default sizes of the bleaching and observation regions; (b) and (d) with the reduced sizes of the bleaching and observation regions.

shown in Figure 4.7. The error function when using FRAP data in space is given by

$$\mathcal{E} \equiv \frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (C_{exp}(t_i, x_j) - C_{sim}(t_i, x_j))^2.$$

In conventional FRAP, the recovery data is obtained by averaging the fluoresence across the entire observation region. Our results suggest the spatial information of FRAP data, which is lost in the average process, can contribute to parameter estimation. In reality, the spatial data might contain noise and cannot be used directly. If it is the case, local average across several pixels instead of average across the entire observation region can be used to retain part of the spatial information which still help improve parameter estimation.

True values of parameters			Bleaching size	Observation T	Estimates of parameters		
D	k^+	k^-	Method	$T(sec)$	D	k^+	k^-
2.5e-4	1	0.1	Default	1000	1.8661e-4	0.7167	9.9335e-2
				100	1.9132e-4	0.7542	0.1013
				10	2.5689e-4	0.9684	0.09649
			Reduced size	100	2.4228e-4	0.9666	0.09999
				10	2.4364e-4	0.9666	0.09990
			Spatial FRAP	1000	2.5411e-4	1.0176	9.9956e-2
				100	2.5397e-4	1.0169	0.09995
				10	2.5093e-4	1.004	0.09992

Table 4.7: Spatial FRAP improves parameter estimation as much as reducing the size of the bleaching region does.

Applications to Model 2, 3 and 4

The conclusions about how to improve parameter estimation via reducing the size of the bleaching region and/or using spatial FRAP data that are obtained above from the diffusion-binding model can also be applied to other models. For the simulations of the models that involve both influx and decay, the initial condition is equal to the fluorescence profile at the steady state of the corresponding system outside the bleachig region and is equal to zero in the bleaching region. In addition, the estimation results are all obtained for a fixed influx. The following results, which are obtained with different models, *i.e.,* a model with multi-binding sites (Table 4.8), a model with influx, binding, diffusion and decay (Table 4.9), and a model with influx, diffusion, binding, internalization and decay (Table 4.10), will show that both reducing the size of the bleaching region and using spatial FRAP data can improve the parameter estimation, and that they have synergistic effect on improvements.

4.3.2 The effect of a reduced recovery model

In addition to inaccurate parameter estimation, which is difficult to ascertain when a visually good curve fit for recovery is obtained, another issue concerns the appropriateness of the model in terms of whether it reflects the biological processes involved in the system. We will show that the estimates of parameters in the reduced model can be

Method	D	k_1^+	k_1^-	k_2^+	k_2^-
Default	1.9784e-4	7.7553e-2	9.9400e-2	4.8572e-1	8.1734e-1
Reduced size	2.4809e-4	9.8993e-2	9.9982e-2	9.8786e-1	1.0003
Spatial	2.5259e-4	1.0124e-1	1.0008e-1	1.0251	1.0055
Spatial and Reduced Size	2.5091e-4	1.0041e-1	9.9992e-2	1.0060	9.9934e-1

Table 4.8: Reducing the size of the bleaching region and/or using spatial FRAP improves the estimates when there are multiple binding sites (Model 2). All the results are based on $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = k^- = 0.1$ sec^{-1}, $k_2^+ = k_2^- = 1$ sec^{-1} and an observation time of 100 sec.

True values of parameters				Method	Estimates of parameters			
D	k^+	k^-	k^d		D	k^+	k^-	k^d
2.5e-4	1	1	1e-2	Default	2.3349 e-4	8.7891e-1	1.0087	1.0705e-2
				Reduced size	2.4806e-4	9.7373e-1	9.8147e-1	1.0014e-2
				Spatial	2.5048e-4	1.0189	1.0126	9.9287e-3
				Spatial and Reduced size	2.5019e-4	1.0150	1.0043	9.9329e-3
2.5e-5	1	0. 1	1e-3	Default	3.3815e-6	1.0022e-1	1.9562e-1	3.5639e-3
				Reduced size	3.0862e-5	1.4288	1.0568e-1	1.0226e-3
				Spatial	2.8780e-5	1.1685	1.0003e-1	9.7550e-4
				Spatial and Reduced Size	2.6132e-5	1.0459	9.9101e-2	9.6777e-4

Table 4.9: Reducing the size of the bleaching region and/or sptial FRAP improve estimates when there are influx, diffusion, binding and decay (Model 3). The observation time is 100 sec , and the influx J is given for parameter estimation.

dramatically different from those from the theoretical model for data generation, and thus lead to distinct conclusions about the transport and kinetic processes involved in FRAP. We will also show that the method of reducing the bleaching size can help to evaluate the appropriateness of the model used for estimation.

Reduction from Model 1 to Model B1

As it has been discussed above, the diffusion and binding model could be described by a diffusion-only model when the parameters are in the effective-diffusion regime. By reducing the bleaching size, the parameters can be moved to the reaction-diffusion or

Method	D	k^+	k^-	k_{in}	k^d
Default	5.5847e-5	3.2081e-1	1.5657e-1	3.3183e-3	1.3302e-2
Reduced size	2.6681e-4	1.0845	9.9709e-2	2.5652e-3	1.0016e-2
Spatial	2.4723e-4	9.8682e-1	9.9630e-2	2.5437e-3	9.9927e-3
Spatial and Reduced size	2.4897e-4	1.0011	9.9895e-2	2.5461e-3	9.9722e-3

Table 4.10: Reducing the size of the bleaching region and/or spatial FRAP improves estimates when there is influx, diffusion, binding, internalization and decay (Model 4). The true parameter values are $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1$ sec^{-1}, $k^- = 0.1$ sec^{-1}, $k_{in} = 2.5641 \times 10^{-3}$ sec^{-1}, $k^d = 1 \times 10^{-2}$ sec^{-1}, $J = 1 \times 10^{-2}$ sec^{-1}. The observation time is $100sec$, and the influx J is fixed for parameter estimation.

reaction-dominant regime, where all the three parameters can be estimated accurately. Therefore, the method of reducing the size of the bleaching region could be used to determine whether binding should be included in FRAP modeling given by experimental data. In reality, the experimentalist has to specify the model before data fitting in FRAP, but as was discussed earlier, it can be difficult to distinguish a diffusion-binding scenario and an effective diffusion regime. Moreover, if the binding is tight, orders of magnitude difference in the estimate of diffusion coefficient might be produced due to an inappropriate model.

Therefore, we suggest that reducing the bleaching region can help to distinguish the diffusion-binding case from the effective diffusion case. When the bleaching region is large, it is possible to obtain very good data fitting by using a diffusion model, while in reality there is also a binding process involved (Figure 4.9, left). However, when the bleaching region is reduced, the diffusion process will equilibrate more quickly, which makes the binding process more dominant in the experiment. Thus it is less likely to fit data by a diffusion model when in reality there is an additional binding process (Figure 4.9, right).

Reduction from Model 2 to Model 1

In the case of multiple binding sites (Model 3), without loss of generality, the second binding process is assumed to the faster one between two binding processes, i.e. $\tau_{b1} = 1/k_1^+ >> 1/k_2^+ = \tau_{b2}$. As long as $\tau_D > \tau_{b2}$, the curve fitting is good even though

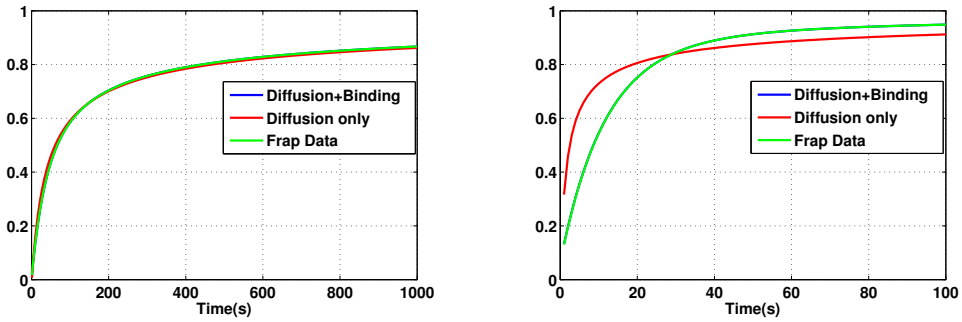


Figure 4.9: Reducing the size of the bleaching region helps to identify the appropriate model. The FRAP data is generated using $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1$ sec^{-1}, $k^- = 0.1$ sec^{-1}. The blue curve lies under the green curve in both panels.

the model for parameter estimation is the single binding site model reduced from the theoretical one. Theoretically, the reduced model gives rise to the estimates of diffusion coefficient and binding/unbinding rates as follows. $D_{rd} = D/(1+k_2^+/k_2^-)$, $k_+ = k_1^+/(1+k_2^+/k_2^-)$, $k^- = k_1^-$. Practically, the simulation results with normal bleaching size also support the conclusion. Moreover, the longer time the data collected for parameter estimation, the closer the simulation results are to the theoretical conclusions.

Theoretical Model (Model 2)							Reduced Model (Model 1)			
Diffusion + Two binding sites					Bleaching size	Time	Diffusion + Single binding site			Curve fitting
D	k_1^+	k_1^-	k_2^+	k_2^-	δ^-	T (sec)	D	k^+	k^-	
2.5e-4	1	0.1	10	1	Default	100	2.2629e-5	8.2322e-2	9.2352e-2	good
2.5e-4	1e-3	1e-3	10	1	Default	1000	2.2122e-5	8.4915e-5	9.8877e-4	good
2.5e-5	1e-3	1e-3	1	0.1	Default	1000	2.2413e-6	1.3014e-4	1.3478e-3	good
					Reduce Size	1000	1.1746e-6	6.2724e-1	1.5113	bad

Table 4.11: Parameter estimation in reduced models when there are multiple binding sites.

From both the analytical and simulation results, we can see that when one of the binding process is fast and tight, it will not only lead to underestimate the diffusion in orders of magnitude, but also underestimate the binding rate in the same order that may change the conclusion about binding affinity fundamentally, i.e., the tight binding may be misinterpreted as loose binding.

In addition, when $\tau_D << \tau_{b2}$, it is difficult to get good curve fitting if the model is

reduced from the theoretical one. Reducing the size of the bleaching region that decrease the characteristic time of diffusion will help to distinguish the multi-binding site model from the single-binding site model, and thus help to formulate the appropriate model for FRAP in order to get meanful estimates of transport and kinetic parameters.

Reduction of Model 4 to Model 3 and Model B2

In the case of FRAP with influx, the data is generated by Model 4 with influx, diffusion, binding, internalization and decay. The first reduced model is Model 2, i.e. it neglects the internalization process. The second reduced model is Model B2, i.e. it neglects all the intermediate processes and only includes influx, diffusion and decay.

	Theoretical Model	Reduced Model (Model 3)	Reduced Model (Model B2)
D	2.5e-4	4.54e-5	1.51e-5
k^+	1	3.38e-2	
k^-	0.1	2.28e-2	
k^i	2.6e-3		
k^d	0.01	3.19e-3	1.85e-3

Table 4.12: Parameter estimation in reduced models when there is diffusion, binding, internalization, decay and influx in the theoretical model. The data is generated using Model 4 with parameters $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1$ sec^{-1}, $k^- = 0.1$ sec^{-1}, $k^i = 2.6 \times 10^{-3}$ sec^{-1}, $k^d = 1 \times 10^{-2}$ sec^{-1}. The flux J is fixed during parameter estimation.

Although the models used for parameter estimation are reduced from the theoretical model by which the FRAP data is generated, the estimates of parameters in these reduced model can still give rise to good curve fittings (results not shown). However, from the results in reduced model 2, in addition to the diffusion coefficient is underestimated in an order of magnitude due to tight binding, the slow internalization rate could lead to the estimate of the decay rate much smaller when the reduced model is missing the intermediate processes and involves only influx,diffusion and decay. In addition, in reduced model 1, the slow internalization process that is missed also has an impact on the estimates of binding and unbinding rates. Therefore, the estimates of parameters in a FRAP model only through curve fitting provides little even wrong information about the actual processes or mechanism involved in biological system.

4.4 Application of sensitivity analysis

In the preceding sections we have analyzed the consequences of various hypotheses about the model and the effect they have on the accuracy of parameter determination. The analysis showed, amongst other things, that different time intervals of observation could significantly affect the parameter estimation. Of course one usually has experimental data rather than computer-generated data, and the question arises as to how one can identify and quantify sensitivity of estimated parameters, other than by the minimization techniques used earlier. The following example illustrates the limitations of the minimization and how sensitivity analysis can give further insights.

Example 2 *Consider a two parameter system, and suppose that the graph of the error function $\mathcal{E}$ to be minimized has either the form of a paraboloid (Figure 4.10(a)) or a parabolic cylinder (Figure 4.10(b)).*

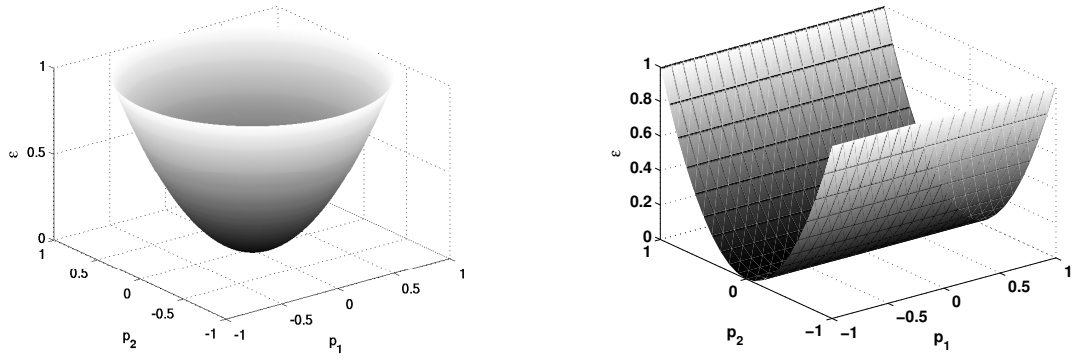


Figure 4.10: Two minimization functions that may give the same minimum, but very different parameter sensitivities. (a) A paraboloid, and (b) a parabolic cylinder.

Clearly the paraboloid has a well-defined minimum and the slope of the function is the same along all directions in the $p_1 - p_2$ plane. By bending the parabolic cylinder slightly upward along the p_2 axes one can guarantee that the minimimun is at zero, as for the paraboloid, but clearly the sensitivity of the error function with respect to variations in the two parameters is very different. In the remainder of this section we show how scatter plots and sensitivity analysis can be used to detect such differences.

4.4.1 The use of scatter plots

In this approach one postulates a model, computes the solutions for a wide-range of the parameters, and then compares the difference between the predictions and the experimental results. With first illustrate this with an example of pure diffusion (Model B1) so as to demonstrate the utility as clearly as possible.

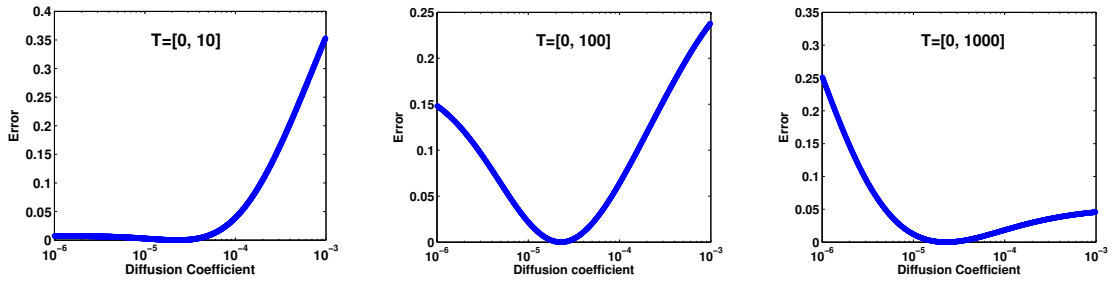


Figure 4.11: Scatterplots of the errors between the model output and the FRAP data versus the diffusion coefficient for different time intervals. The FRAP data is generated by the pure-diffusion model with $D = 2.25 \times 10^{-5}$ sec^{-1} to match the data in previous simulations by using the effective diffusion coefficient. The scatterplots are calculated with parameters uniformly distributed on a logarithmical scale $D \in [1 \times 10^{-6}, 1 \times 10^{-3}]$ sec^{-1}. $N = 1000$ is the number of sample points.

The structure of the scatter plots shown in Figure 4.11 can be understood as follows. When the time interval is short [0,10] the recovery is small and the error is very small for small diffusion coefficients since the error as defined in Figure 4.4 is the distance between the actual and predicted recovery curves (*cf.* Figure 4.12 (right)). When the diffusion coefficient is significantly larger than the true value the predicted recovery curve rises much faster than the true recovery curve and the error increases with the diffusion rate. In an intermediate interval [0,100] the error is significant for both too small and too large a diffusion coefficient. When the observation time period is long [0,1000], the error for larger diffusion coefficients is less significant than for smaller ones because the predicted recovery curve lies close to the true curve at large times, where the error is small (*cf.* Figure 4.12 (left)). These results indicate that the intermediate time interval [0,100] is optimal for this problem, since the true diffusion coefficient is most clearly defined at the minimum of the graph of the error.

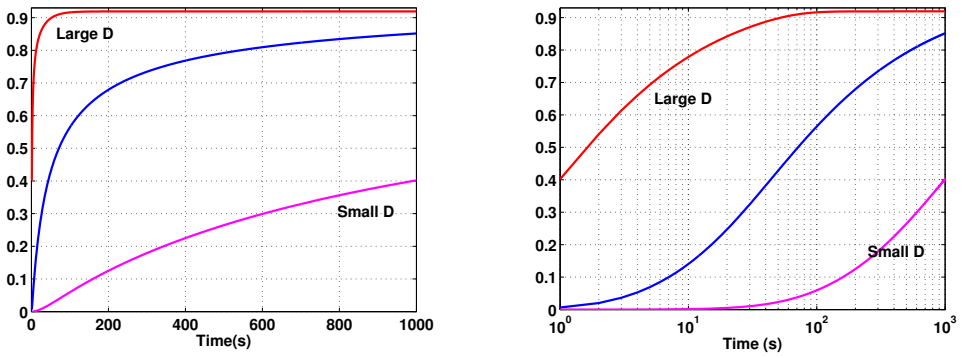


Figure 4.12: FRAP recovery data is generated by the pure-diffusion model with $D = 2.25 \times 10^{-5}$ sec^{-1}. The large D and small D refer to the upper and lower limits of the diffusion coefficients used for the scatterplots, respectively. The figure on the left is plotted on a linear scale, and the one on the right is plotted on a logarithmic scale.

Scatterplots can also give insight for more complicated models such as the diffusion-binding models (Model 1 and Model 2) . The top row in Figure 4.13 shows that the addition of binding to the diffusion-only model of Figure 4.12 has little effect on the error for a short time interval and small diffusion coefficients, and that the error is dominated by diffusion even for large diffusion coefficients. However the effect of variations in the binding parameters is more pronounced for longer time intervals, but in all cases the role of diffusion remains as shown in Figure 4.12.

The center row of Figure 4.13 displays the scatter plots as a function of the binding affinity k^+/k^-, for which the true value is 10. For the short time interval T = [0,10] the recovery is small and the effect of diffusion on the error is negligible for large affinities because the fluorescent molecules are tightly bound in the unbleached region and the flux into the bleached region is small. When the binding affinity is small, the influx is larger and diffusion plays a larger role, which leads to larger errors. As the observation time period increases, the difference between these two upper limits diminishes. When the observation time period is long enough, e.g. $T \in [0, 1000]$, the situation is reversed. Therefore, using a scatterplot for variable observation times T will suggest what the true affinity is when a T that produces the pattern of errors similar to that in the first figure in the middle panel is found.

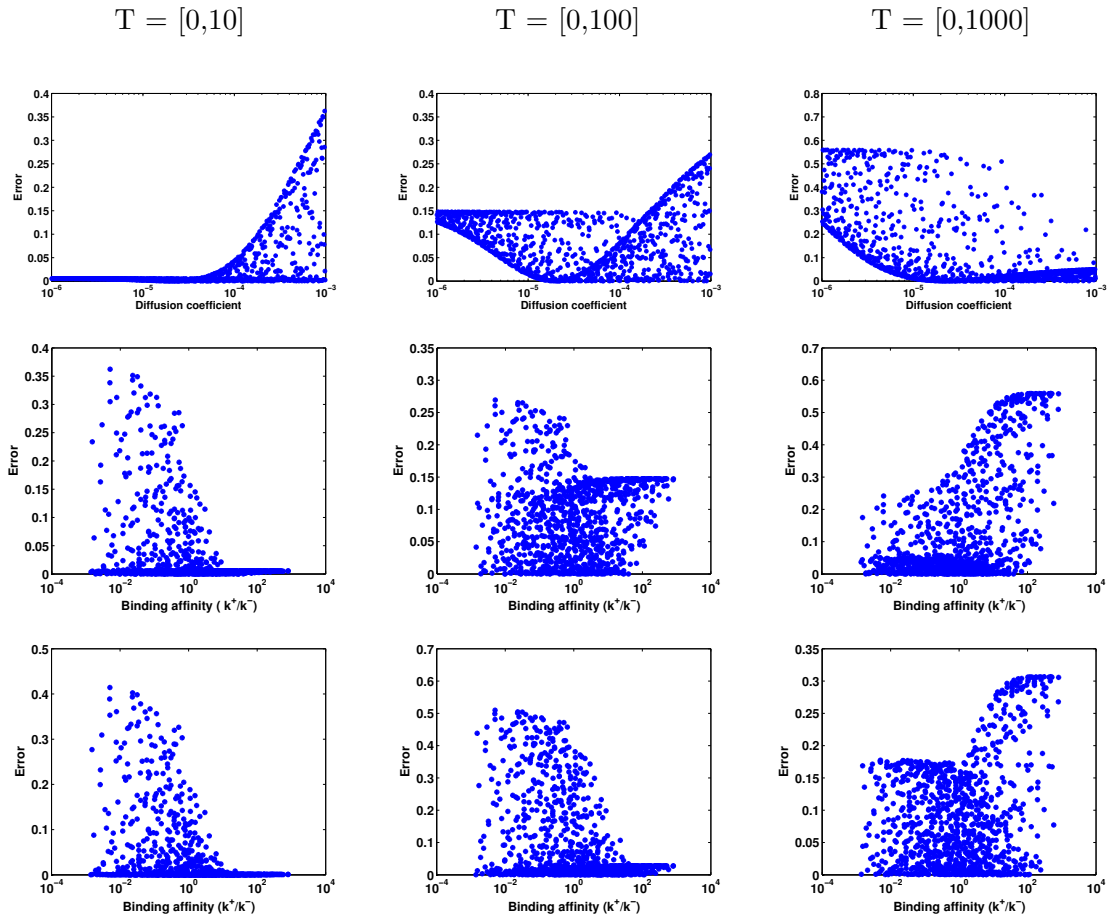


Figure 4.13: Scatterplots of the errors between the model output and the FRAP data versus diffusion coefficients and binding affinities for different time intervals . Top and middle panels: FRAP data is generated with $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1$ sec^{-1}, $k^- = 0.1$ sec^{-1}. In this and the panels below the parameters are log-uniformly distributed – using $D = \in [1\times10^{-6}, 1\times10^{-3}]$ sec^{-1}, $k^+ \in [1\times10^{-2}, 10$ sec$^{-1}]$, $k^- \in [1\times10^{-2}, 10]$ sec^{-1}. Bottom panel: The FRAP data is generated by the model with diffusion and two binding processes with different rates and affinities $D = 2.5\times10^{-4}$ sec^{-1}, $k_1^+ = 5$ sec^{-1}, $k_1^- = 0.5$ sec^{-1}, $k_2^+ = 0.1$ sec^{-1}, $k_2^- = 5 \times 10^{-3}$ sec^{-1}. The parameters are log-uniformly distributed: $D \in [1\times10^{-6}, 1\times10^{-3}]$ sec^{-1}, $k^+ \in [5\times10^{-3}, 5]$ sec^{-1}, $k^- \in [5\times10^{-3}, 5]$ sec^{-1}. $N = 1000$ is the number of sample points for all.

These two rows suggest that the scatterplots of errors against diffusion coefficient with different observation time periods identify whether the parameters are located in effective-diffusion regime and if so, what the effective diffusion coefficient is. If in effective-diffusion regime, by combining the scatterplots against binding affinity and diffusion coefficient, the binding affinity and the effective diffusion coefficient, and thereby the true diffusion coefficient can be obtained.

The bottom row in Figure 4.13 shows scatterplots of the errors for a one-site recovery model when the true model has two binding sites. These scatterplots with different observation time periods not only can provide the experimentalists with the binding affinities but also can suggest how many different types of binding sites are involved when they have different binding affinities. A binding process with affinity equal to ~ 10 appears at the scatterplot with the observation time period $T \in [0, 10]$. And the other one with affinity equal to ~ 20 appears at the scatterplot with $T \in [0, 100]$. Although in reality it may be hard to tell the exact values of binding affinities, the scatterplots with different observation time periods at least help indicate the possiblity of multiple binding sites. They only suggest how many types of binding sites with significantly different rates and affinities are involved.

4.4.2 Variance-based sensitivity analysis

The scatter-plot-based procedure in the preceding section gives qualitative information to parameters, but more precise tests about where the parameter sensitivity lies can be applied after parameter estimation using other methods of sensitivity analysis. The objective of this analysis is to obtain insight as to how the $\mathcal{E}$ varies with parameter variation in a neighborhood of the computed minimum. The non-local analysis described below is more informative than simply computing the derivatives of $\mathcal{E}$ at the minimum because parameters can be varied over large intervals around the minimum. In applying the general technique to the FRAP problem we define $Y = \mathcal{E}$, the error between the observed and predicted recovery as defined earlier, and the factors X are the parameters that are estimated from the data. Thus the model equation is rewritten as

$$\mathcal{E} = f(p_1, p_2, ...p_k) = f(P), \tag{4.59}$$

where $P = (p_1, p_2, ...p_k)$. For the purpose of the sensitivity analysis that follows we assume that parameters are distributed uniformly.

As we showed previously, the FRAP recovery data generated by a theoretical model can be fit very well with a reduced model in some circumstances. For instance, data generated by a diffusion-binding model can be described with a pure-diffusion model when the binding process is much faster than diffusion as shown in the left panel of Figure 4.14. We proposed that changing the balance between different processes by changing the size of the bleaching region gives rise to a different recovery curve and helps to detect missing processes and to formulate a more appropriate model. When all the processes in the model that generates FRAP data are balanced, *i.e.,* they all occur on comparable time scales, it is unlikely to fit the data with a reduced simple model, as shown in the center panel of Figure 4.14.

However, when a complex model that includes more processes than are important is used to fit the data, it may be difficult to detect this from the fit of the recovery curve. As shown in the right panel of Figure 4.14, even though all the processes in the intermediate model that generates the data are well balanced, we can still fit the data very well with a complex model, and this is where the sensitivity analysis can be useful. In this case, when the sensitivity analysis is implemented using the intermediate (theoretical) model, the sensivity indices of all the parameters are comparable, and none of the total indices is very small. However, when the sensitivity analysis is applied to the complex model, the total-order indices of some parameters, such as k^t and k^{d2} shown in (c) and (d) of Figure 4.15. This suggests the possibility of over-parameterization, *i.e.,* the more detailed model might include non-influential processes that are not detectable using the available data.

It is worth noting that when parameters are estimated using either the reduced simple model [15] or the complex model [28] with the same FRAP data generated by the intermediate model, estimates of diffusion coefficients in both recovery models (4.26×10^{-4} sec^{-1} vs. 1.41×10^{-6} sec^{-1}) differ by a factor of 60, which is similar to the large difference in the measurement of the diffusion coefficient of Dpp in [15] and [28]. However, in our simulations, neither of the estimates are close to the true value for data generation even though two models can fit the steady-state (results not shown here) and recovery data.

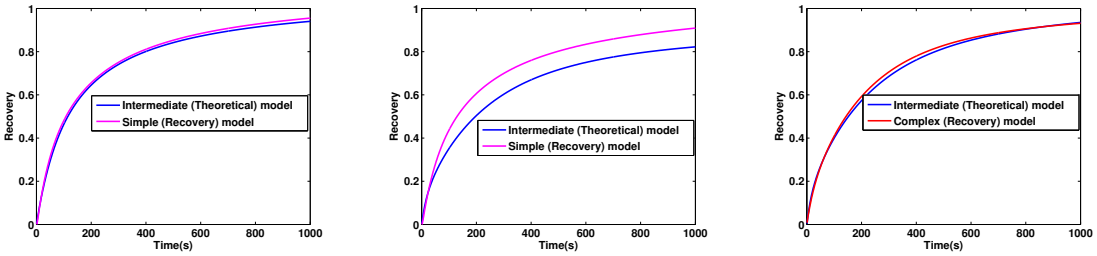


Figure 4.14: FRAP data is generated with the intermediate (theoretical) model (Model 2 in Table 4.1) with the unbalanced processes (left) and the balanced processes (center and right). The true values of the parameters for the unbalanced processes are $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1 \times 10^{-1}$ sec^{-1}, $k^- = 5 \times 10^{-2}$ sec^{-1}, $k^d = 2 \times 10^{-3}$ sec^{-1}, and the parameters for the balanced processes are $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1 \times 10^{-2}$ sec^{-1}, $k^- = 5 \times 10^{-3}$ sec^{-1}, $k^d = 2 \times 10^{-3}$ sec^{-1}. Parameters are estimated by using the simple (recovery) model (Model B2 in Table 4.1, which is also the same as that in [15]) for the left and center panels, and using the complex (recovery) model (Model **??** in Appendix) in the right panel. The estimates are (left) $D = 7.8322 \times 10^{-6}$ sec^{-1}, $k_d = 1.2698 \times 10^{-3}$ sec^{-1}; (center) $D = 7.0290 \times 10^{-6}$ sec^{-1}, $k_d = 9.9313 \times 10^{-4}$ sec^{-1}; (right) $D = 4.4680 \times 10^{-4}$ sec^{-1}, $k^+ = 1.7633$ sec^{-1}, $k^- = 1.1396 \times 10^{-1}$ sec^{-1}, $k^i = 1.1131 \times 10^{-2}$, $k^o = 6.0198 \times 10^{-3}$ sec^{-1}, $k_t = 1.7574 \times 10^{-7}$ sec^{-1}, $k_{d1} = 2.1856 \times 10^{-3}$ sec^{-1}, $k_{d2} = 1.1563 \times 10^{-3}$ sec^{-1}.

4.5 Discussion

In an experimental context the standard approach to use the FRAP technique is to measure the experimental data and then fit a model to the recovery curve. While informative, it is difficult to analyze the model and evaluate the quality of estimates because the parameters underlying physical processes in reality are unknown. Our aim here was to approach this problem by using a theoretical model to generate FRAP data and postulating a recovery model to estimate the parameters, and knowing both *a priori* enabled us to quantitatively assess the quality of estimates and find ways to improve them. Firstly, by using a recovery model identical to the theoretical model, we showed that good fitting of the data may be misleading in some circumstances, in that it does not always indicate high quality estimates. We identified factors that lead

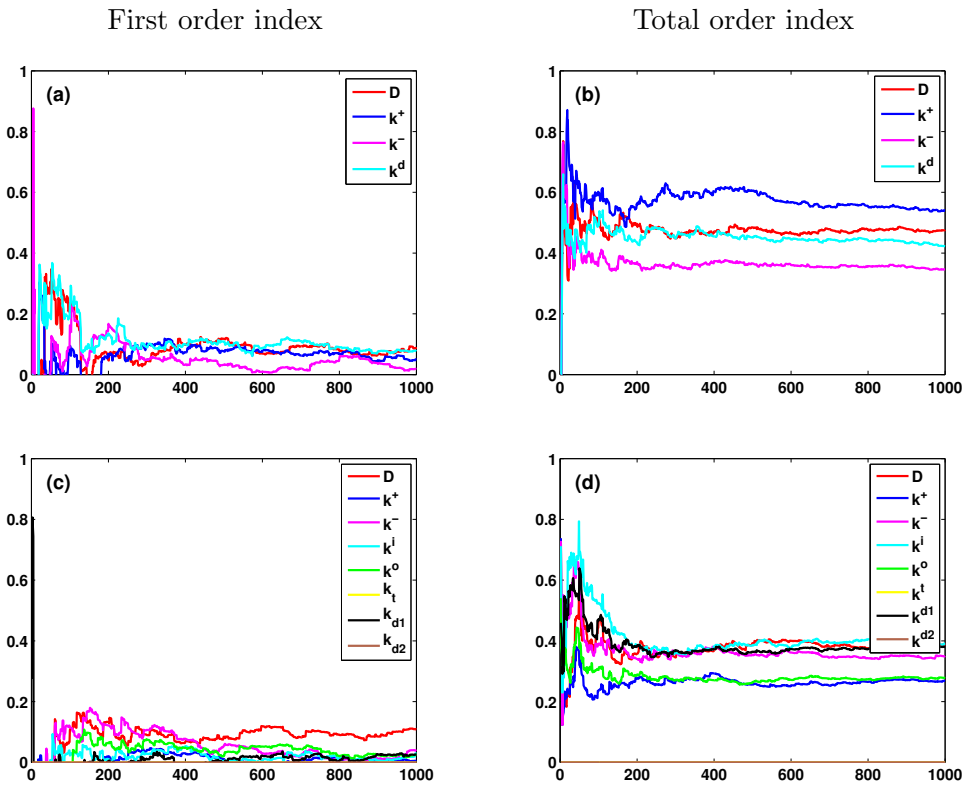


Figure 4.15: FRAP data is generated with the intermediate model (Model 2 in Table 4.1) with parameters $D = 2.5 \times 10^{-4}$ sec^{-1}, $k^+ = 1 \times 10^{-2}$ sec^{-1}, $k^- = 5 \times 10^{-3}$ sec^{-1}, $k^d = 2 \times 10^{-3}$ sec^{-1}. (a) and (b): The first order and total order sensitivity indices are calculated by using the same intermediate model with parameters with uniform linear distribution $D \in [0.5 \times 10^{-5}, 4.5 \times 10^{-5}]$ sec^{-1}, $k^+ \in [0.2 \times 10^{-2}, 1.8 \times 10^{-2}]$ sec^{-1}, $k^- \in [1 \times 10^{-3}, 9 \times 10^{-3}]$ sec^{-1}, $k^d \in [0.4 \times 10^{-3}, 3.6 \times 10^{-3}]$ sec^{-1}. (c) and (d): The first order and total order sensitivity indices are calculated by using the complex model **??** with parameters with uniform linear distribution around the estimates ($[0.2 \times Estimate, 1.8 \times Estimate]$) $D \in [0.8510 \times 10^{-4}, 7.6594 \times 10^{-4}]$ sec^{-1}, $k^+ \in [0.3526, 3.1739]$ sec^{-1}, $k^- \in [0.2849 \times 10^{-1}, 2.5641 \times 10^{-1}]$ sec^{-1}, $k^i \in [0.2226 \times 10^{-2}, 2.0036 \times 10^{-2}]$ sec^{-1}, $k^o \in [1 \times 10^{-3}, 9 \times 10^{-3}]$ sec^{-1}, $k^t \in [0.3515 \times 10^{-7}, 3.1633 \times 10^{-7}]$ sec^{-1}, $k^{d1} \in [0.3868 \times 10^{-3}, 3.4816 \times 10^{-3}]$ sec $^{-1}$, $k^{d2} \in [0.2313 \times 10^{-3}, 2.0813 \times 10^{-3}]$ sec^{-1}.

to poor parameter estimation from FRAP data, and suggested three new, feasible ways in which the estimation can be improved - using the FRAP data in an appropriate observation time period, changing the size of the bleaching region to rebalance the diffusion and kinetic processes, and using the spatial information of FRAP data. Then, by varying the recovery model from the theoretical model, we showed that a simplified recovery model can adequately describe the FRAP processes in some circumstances, and established the relationship between parameters in the theoretical model and those in the recovery model. Finally, we introduced variance-based parameter sensitivity into FRAP analysis, and suggested that the important kinetic processes might be detected by sensitivity analysis before estimation, and the over-parameterization problem in a FRAP model can be perceived by doing sensivity analysis after estimation.

In using FRAP, it is important to determine which processes should be included in the recovery model. For example, ignoring binding processes that are present in the system may lead to underestimation of the diffusion coefficient by an order of magnitude. Given FRAP data, we proposed two different ways that can facilitate identification of the appropriate model. We found that changing the size of the bleaching region gives rise to different FRAP recovery curves and can provide insight into the relative effects of diffusion and binding kinetics. In particuar, reducing the size of the bleaching region to speed up the diffusion process relative to the kinetic processes can help uncover a hidden binding process in the recovery curve, which might be neglected when the bleaching region is large. In addition, we showed that preliminary sensitivity analysis using scatterplots with physically reasonable ranges of parameters may also help detect multiple binding processes. Using the two methods will reduce the chance of neglecting important processes in the model. In addition, sensitivity analysis after estimation using the first and total order indices can suggest over-parameterization problem in a FRAP model, *i.e.,* the model contains non-influential processes. If some of the parameters have very low total-order indices, the model is more complex than is justified by the data available, and has to be reduced by eliminating the non-influential processes. This step can be repeated until none of the parameters in the model have extremely low total-order indices. The ideal scenario is that the corresponding sensitivity index around the estimate of each parameter in the model is comparable, which indicates all the processes are well balanced. Incorporating these methods into the FRAP analysis can

greatly increase the probability of formulating appropriate models and thereby also increase the accurary of parameter estimation.

Although we showed that reducing the size of the bleaching region can help to formulate a more appropriate model and improve parameter estimation in some circumstances, it is difficult to decide what the size of the bleaching region should be at the outset of a FRAP experiment. However the knowledge of the effect of a reduction can be used in the following way. After the estimates of parameters are obtained from a first FRAP experiment, one can calculate the characteristic time scales of diffusion and kinetic processes, and depending on the results, the experiment can be re-done with a bleaching region that leads to a better balance of the processes. Similar remarks apply to the use of spatial information in FRAP data. Averaging spatial FRAP data over the whole bleaching region looses some information which might be useful for parameter estimation. However, for a realistic FRAP experiment, the spatial FRAP data always contains noise which is greatly tenuated by averaging FRAP data. We suggest that it is possible to benefit from spatial FRAP data if local averaging rather than global averaging is implemented. In addition to the quality of parameter estimation, the advantage of using spatial FRAP data may be explored in many other aspects of model identification, such as estimation of more parameters in FRAP models.

In summary, we suggest that the procedure shown in Figure 4.16 to better use FRAP data in the process of model formulation and parameter estimation. This method can be used for general FRAP modeling and analysis not discussed here. For instance, the assumptions used here, such as an instantaneous and homogeneous bleaching process, may not be valid in some circumstances, and it would be interesting to apply our method to model the whole FRAP processes as described in Appendix **??**, and to study how different assumptions affect the estimates quantitatively.

The establishment of the morphogen profiles in the wing disc is a very complex process that may involve several distinct morphogen transport processes, endo- and exocytosis of the morphogen, and intracellular sequestration of it. As a result, parameter estimates derived from fitting of FRAP recovery curves are not likely to bear a close relationship to true parameter values, since the FRAP data is inadequate to extract the true parameters in a complex model. Thus such tissue-level applications of FRAP must be supplemented with other techniques in order to identify the processes and

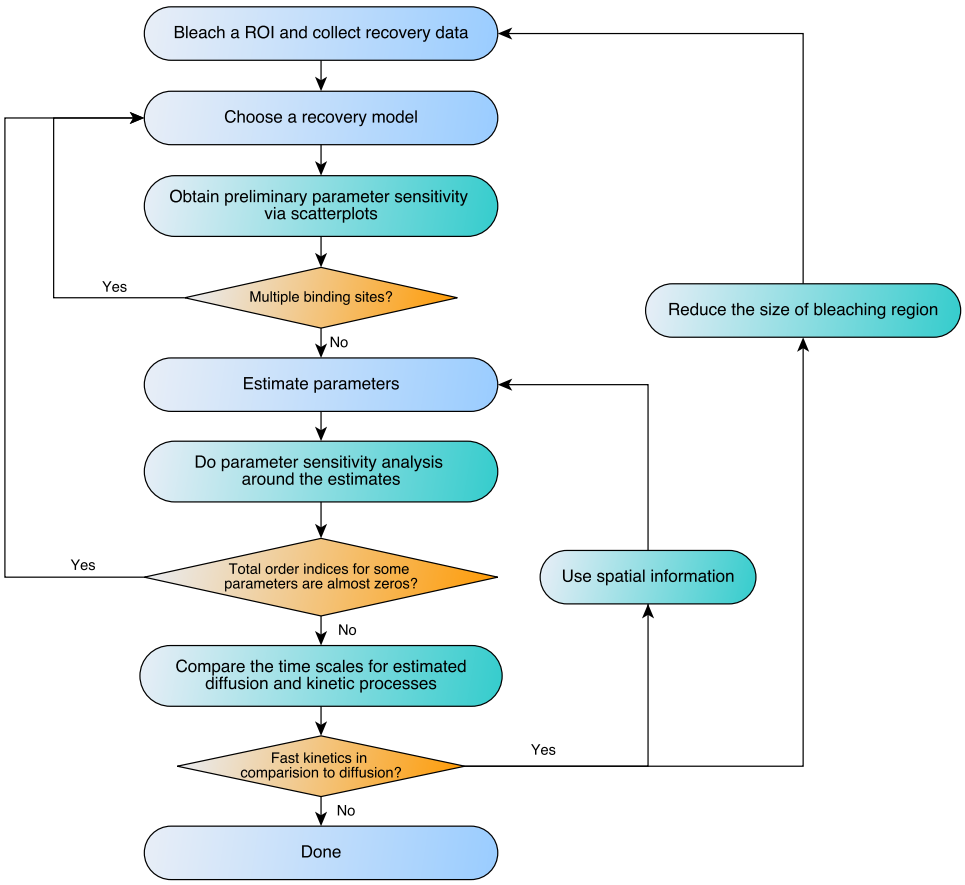


Figure 4.16: A suggested procedure for improving model identification and parameter estimation.

the attendant parameters. This remains as a significant challenge in the context of developmental biology.

The simple model described by (4.16) has been used to estimate parameters from experimental FRAP data in several systems, but it may have very limited applicability in identifying the parameters that govern the *in vivo* dynamics, since other kinetic or transport processes are usually involved. Even in the simple case of Bicoid dynamics in the early *Drosophila* embryo, processes other than diffusion and reaction, including binding, localization of Bicoid in the nuclei, and non-homogeneities in the cytosol

and cortical layer of the syncytium, are certainly involved. Moreover, in a cellularized system such as the *Drosophila* wing disc there may be several other modes of cell-cell communication.

The complexity involved in the interactions of multiple transport processes with binding, internalization and kinetic transformation is hidden by analyzing the spread of morphogens using a simple reaction-diffusion system such as (4.16), and how one relates the 'real' parameters in individual steps to parameters in such a high-level description is usually difficult to determine analytically. This has lead to dramatic differences (*cf.* [15, 28] in the estimates of parameters for a simplified description such as (4.16), and raises the question as to what those parameters represent.

Chapter 5

Conclusion and future direction

The complexity of patterning and growth control demands both experimental and modeling work to address different aspects of problems. Our ultimate goal is to formulate a model that integrates different regulatory mechanisms of patterning and growth control in the disc so as to be able to understand existing experimental results, to make experimentally-testable predictions, and to provide a platform for integrating new results and testing new hypotheses. We are interested in the Hippo pathway, which various regulatory mechanisms and a myriad of signaling pathways impinge on, and in this phase we focus on understanding how the Hippo pathway tranduces Fat/Ds signaling to mediate cell-cell interaction and to coordinate growth on the tissue level, keeping in mind how it is coupled with other pathways and interacts with other regulatory mechanisms.

The first and laborious step in the course of developing a model for the signaling network involves the delineation of all the known individual biochemical processes that comprise the network. After we thoroughly evaluated experimental evidence in the literature and identified the proteins within the network that are expressed in cells, the balance between the scope and level of detail was considered, and the extent to which species included in the model evolve independently of species excluded from the model was carefully examined. Finally, our model encompasses a relative small number of distinct species and all reactions among them taking place in cells and between cells in order to retain the predictive power of the model without overwhelming uncertainty. As little investigation has been done for reactions in this network, in our opinion, properties are more interesting when discovered in the course of analyzing the model based only on

elementary reactions and mass action kinetics. The model advanced our understanding of Fat/Ds induced responses via the Hippo pathway and made predictions that can be tested experimentally at the levels of genes and proteins.

Two hypotheses are generated and predicted in the mechanistic model that we developed. We posit the interaction between Ds and Dachs, and incorporate the effect of Ds on Riq, emphasizing the importance of Ds. This adds another aspect to efforts aiming to elucidate the mechanistic basis of genetic phenotypes associated with the Hippo pathway. One of the hypotheses has been confirmed by experiments, suggesting the reliability of the model for simulating this complex pathway. The model integrates the primary signaling components, captures the interactions between them, and quantitatively investigates how they are balanced to regulate the growth of the disc. A major advance of our work over previous modeling is the level of detail that is incorporated for both signal transduction within cells and signal transport across cells. This newly developed model demonstrates its sufficiency to explain existing experimental observations, and predicts new mechanisms that can be tested experimentally.

The interplay between models and experimental data leads to the question of how the parameters in the model can be obtained from available experimental data, which led us to take a new look at FRAP and to investigate how to improve model identification and parameter estimation given FRAP data. We used a theoretical model to simulate the dynamics of a FRAP experiment and generate the FRAP data that is used in various recovery models to estimate the corresponding parameters. By postulating a recovery model identical to the theoretical model, we proposed three feasible ways to improve parameter estimation in the circumstances that the model is appropriate. We first established that utilization of FRAP data in an appropriate observation time period, as opposed to using the whole data set, can significantly improve the quality of estimates, especially when the diffusion and binding kinetics are not well balanced. Secondly, we found that changing the balance between diffusion and binding kinetics by changing the size of the bleaching region enhances the accuracy of parameter estimation. In addition, we showed that the use of the spatial information in FRAP provides better parameter estimation, because the standard way of averaging FRAP data loses too much information and leads to errors in parameter estimation. By varying the recovery model from a fixed theoretical model, we showed that a simplified recovery model can

adequately describe the FRAP process in some circumstances, and establish the relationship between parameters in the theoretical model and those in the recovery model. We also showed that changing the size of the bleaching region and global sensitivity analysis can be used to improve model identification and to help formulate appropriate FRAP models.

It has been found that Dachs in the Hippo pathway exerts mechanical force at cellcell junctions to alter the geometry of cell shapes and the orientation of the mitotic spindle. As cell mechanics have been proposed to play a role in growth control, it is possible that the Hippo pathway might also regulate growth via mechanical stress. Moreover, it is involved in crosstalk with the morphogen and hormone signaling pathways, for which parameters involved in signal transport across the disc and the entire organism can be measured using FRAP and other techniques. Therefore, making use of our work on parameter estimation and following our work on modeling the Hippo pathway, we will integrate it with the morphogen and hormone regulatory mechanisms, and the effects of mechanical stress on growth to produce an integrated model that can be used for developing new, experimentally-testable hypotheses. The integrated model will enable us to evaluate the importance of various modes of transport and signaling and to make testable predictions as to what processes dominate under various experimental interventions. This will result in a comprehensive computational platform for use in understanding the spatio-temporal control of growth and patterning in wing disc development, and will provide significant new insights into the complex interactions that govern tissue size, shape and patterning.

References

[1] Y. Funakoshi, M. Minami, and T. Tabata. mtv shapes the activity gradient of the dpp morphogen through regulation of thickveins. *Development*, 128(1):67, 2001.

[2] Lewis I Held Jr, Lewis I Held Jr, and Jonathan BL Bard. *Imaginal discs: the genetic and cellular logic of pattern formation*, volume 39. Cambridge University Press, 2005.

[3] M. Fujise, S. Takeo, K. Kamimura, T. Matsuo, T. Aigaki, S. Izumi, and H. Nakato. Dally regulates dpp morphogen gradient formation in the drosophila wing. *Development*, 130(8):1515–1522, 2003.

[4] Isabel Rodríguez. The dachsous gene, a member of the cadherin family, is required for wg-dependent pattern formation in the drosophila wing disc. *Development*, 131(13):3195–3206, 2004.

[5] Michael A Crickmore and Richard S Mann. The control of size in animals: insights from selector genes. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 30(9):843, 2008.

[6] Yi Arial Zeng, Maryam Rahnama, Simon Wang, Wendy Lee, and Esther M Verheyen. Inhibition of drosophila wg signaling involves competition between mad and armadillo/beta-catenin for dtcf binding. *PLoS One*, 3(12):e3893–e3893, 2008.

[7] Richelle Sopko and Helen McNeill. The skinny on fat: an enormous cadherin that regulates cell adhesion, tissue growth, and planar cell polarity. *Current opinion in cell biology*, 21(5):717–723, 2009.

[8] Kieran F Harvey and Iswar K Hariharan. The hippo pathway. *Cold Spring Harbor perspectives in biology*, 4(8):a011288, 2012.

[9] Maria Willecke, Fisun Hamaratoglu, Leticia Sansores-Garcia, Chunyao Tao, and Georg Halder. Boundaries of dachsous cadherin activity modulate the hippo signaling pathway to induce cell proliferation. *Proceedings of the National Academy of Sciences*, 105(39):14897–14902, 2008.

[10] Hitoshi Matakatsu and Seth S Blair. Separating the adhesive and signaling functions of the fat and dachsous protocadherins. *Development*, 133(12):2315–2324, 2006.

[11] Abhijit A Ambegaonkar, Guohui Pan, Madhav Mani, Yongqiang Feng, and Kenneth D Irvine. Propagation of dachsous-fat planar cell polarity. *Current Biology*, 22(14):1302–1308, 2012.

[12] Mohit Kumar Jolly, Mohd Suhail Rizvi, Amit Kumar, and Pradip Sinha. Mathematical modeling of sub-cellular asymmetry of fat-dachsous heterodimer for generation of planar cell polarity. 2014.

[13] B.L. Sprague and J.G. McNally. Frap analysis of binding: proper and fitting. *Trends in cell biology*, 15(2):84–91, 2005.

[14] David Umulis, Michael B O'Connor, and Hans G Othmer. Robustness of embryonic spatial patterning in drosophila melanogaster. *Current topics in developmental biology*, 81:65–111, 2008.

[15] A. Kicheva, P. Pantazis, T. Bollenbach, Y. Kalaidzidis, T. Bittig, F. Julicher, and M. González-Gaitán. Kinetics of morphogen gradient formation. *Science's STKE*, 315(5811):521, 2007.

[16] Stanislav Y Shvartsman, Cyrill B Muratov, and Douglas A Lauffenburger. Modeling and computational analysis of egf receptor-mediated cell communication in drosophila oogenesis. *Development*, 129(11):2577–2589, 2002.

[17] Keith Amonlirdviman, Narmada A Khare, David RP Tree, Wei-Shen Chen, Jeffrey D Axelrod, and Claire J Tomlin. Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science*, 307(5708):423–426, 2005.

[18] N Yakoby, CA Bristow, I Gouzman, MP Rossi, Y Gogotsi, T Schüpbach, and SY Shvartsman. Systems-level questions in drosophila oogenesis. *IEE Proceedings-Systems Biology*, 152(4):276–284, 2005.

[19] Lea A Goentoro, Gregory T Reeves, Craig P Kowal, Luigi Martinelli, Trudi Schüpbach, and Stanislav Y Shvartsman. Quantifying the gurken morphogen gradient in drosophila oogenesis. *Developmental cell*, 11(2):263–272, 2006.

[20] David M Umulis, Mihaela Serpe, Michael B OConnor, and Hans G Othmer. Robust, bistable patterning of the dorsal surface of the drosophila embryo. *Proceedings of the National Academy of Sciences*, 103(31):11613–11618, 2006.

[21] Theodore J Perkins, Johannes Jaeger, John Reinitz, and Leon Glass. Reverse engineering the gap gene network of drosophila melanogaster. *PLoS Comput Biol*, 2(5):e51, 2006.

[22] Mihaela Serpe, David Umulis, Amy Ralston, Jun Chen, David J Olson, Andrei Avanesov, Hans Othmer, Michael B O'Connor, and Seth S Blair. The bmp-binding protein crossveinless 2 is a short-range, concentration-dependent, biphasic modulator of bmp signaling in drosophila. *Developmental cell*, 14(6):940–953, 2008.

[23] Dali Ma, Keith Amonlirdviman, Robin L Raffard, Alessandro Abate, Claire J Tomlin, and Jeffrey D Axelrod. Cell packing influences planar cell polarity signaling. *Proceedings of the National Academy of Sciences*, 105(48):18800–18805, 2008.

[24] Myriam Zecca and Gary Struhl. A feed-forward circuit linking wingless, fat-dachsous signaling, and the warts-hippo pathway to drosophila wing growth. *PLoS biology*, 8(6):e1000386, 2010.

[25] Floris Bosveld, Isabelle Bonnet, Boris Guirao, Sham Tlili, Zhimin Wang, Ambre Petitalot, Raphaël Marchand, Pierre-Luc Bardet, Philippe Marcq, François Graner, et al. Mechanical control of morphogenesis by fat/dachsous/four-jointed planar cell polarity pathway. *Science*, 336(6082):724–727, 2012.

[26] Tinri Aegerter-Wilmsen, Maria B Heimlicher, Alister C Smith, Pierre Barbier de Reuille, Richard S Smith, Christof M Aegerter, and Konrad Basler. Integrating force-sensing and signaling pathways in a model for the regulation of wing imaginal disc size. *Development*, 139(17):3221–3231, 2012.

[27] Maja Matis and Jeffrey D Axelrod. Regulation of pcp by the fat signaling pathway. *Genes & development*, 27(20):2207–2220, 2013.

[28] Shaohua Zhou, Wing-Cheong Lo, Jeffrey L Suhalim, Michelle A Digman, Enrico Gratton, Qing Nie, and Arthur D Lander. Free extracellular diffusion creates the dpp morphogen gradient of the¡ i¿ drosophila¡/i¿ wing disc. *Current Biology*, 22(8):668–675, 2012.

[29] K. Braeckmans, L. Peeters, N.N. Sanders, S.C. De Smedt, and J. Demeester. Three-dimensional fluorescence recovery after photobleaching with the confocal scanning laser microscope. *Biophysical journal*, 85(4):2240–2252, 2003.

[30] Hans G Othmer, Kevin Painter, David Umulis, and Chuan Xue. The intersection of theory and application in elucidating pattern formation in developmental biology. *Mathematical modelling of natural phenomena*, 4(04):3–82, 2009.

[31] Peter Anthony Lawrence et al. *The making of a fly: the genetics of animal design.* Blackwell Scientific Publications Ltd, 1992.

[32] K Basler and Gary Struhl. Compartment boundaries and the control of drosophila limb pattern by hedgehog protein. *NATURE-LONDON-*, pages 208–208, 1994.

[33] Markus Affolter and Konrad Basler. The decapentaplegic morphogen gradient: from pattern formation to growth regulation. *Nature Reviews Genetics*, 8(9):663–674, 2007.

[34] Laurent Holtzer, Anna Kicheva, Marcos Gonzalez-Gaitan, and Thomas Schmidt. Morphogen gradient formation unraveled using in vivo three-dimensional single molecule microscopy. *Biophysical Journal*, 96(3):33a, 2009.

[35] Thomas Lecuit, William J Brook, Medard Ng, Manuel Calleja, Henry Sun, and Stephen M Cohen. Two distinct mechanisms for long-range patterning by decapentaplegic in the drosophila wing. *Nature*, 381(6581):387–393, 1996.

[36] Peleg Hasson, Bruno Müller, Konrad Basler, and Ze'ev Paroush. Brinker requires two corepressors for maximal and versatile repression in dpp signalling. *The EMBO journal*, 20(20):5725–5736, 2001.

[37] María F Organista and Jose F De Celis. The spalt transcription factors regulate cell proliferation, survival and epithelial integrity downstream of the decapentaplegic signalling pathway. *Biology open*, page BIO20123038, 2012.

[38] Makoto Umemori, Masahiko Takemura, Kousuke Maeda, Keisuke Ohba, and Takashi Adachi-Yamada. ¡ i¿ drosophila¡/i¿ t-box transcription factor optomotor-blind prevents pathological folding and local overgrowth in wing epithelium through confining hh signal. *Developmental biology*, 308(1):68–81, 2007.

[39] Gerald Schwank, Sascha Dalessi, Schu-Fee Yang, Ryohei Yagi, Aitana Morton de Lachapelle, Markus Affolter, Sven Bergmann, and Konrad Basler. Formation of the long range dpp morphogen gradient. *PLoS-Biology*, 9(7):1549, 2011.

[40] Andrea Penton and F Michael Hoffmann. Decapentaplegic restricts the domain of wingless during drosophila limb patterning. 1996.

[41] Myriam Zecca and Gary Struhl. Recruitment of cells into the drosophila wing primordium by a feed-forward circuit of vestigial autoregulation. *Development*, 134(16):3001–3010, 2007.

[42] Eunjoo Cho, Yongqiang Feng, Cordelia Rauskolb, Sushmita Maitra, Rick Fehon, and Kenneth D Irvine. Delineation of a fat tumor suppressor pathway. *Nature genetics*, 38(10):1142–1150, 2006.

[43] Yaopan Mao, Cordelia Rauskolb, Eunjoo Cho, Wei-Li Hu, Heather Hayter, Ginny Minihan, Flora N Katz, and Kenneth D Irvine. Dachs: an unconventional myosin that functions downstream of fat to regulate growth, affinity and gene expression in drosophila. *Development*, 133(13):2539–2551, 2006.

[44] Joffrey L Degoutin, Claire C Milton, Eefang Yu, Marla Tipping, Floris Bosveld, Liu Yang, Yohanns Bellaiche, Alexey Veraksa, and Kieran F Harvey. Riquiqui and minibrain are regulators of the hippo pathway downstream of dachsous. *Nature cell biology*, 15(10):1176–1185, 2013.

[45] Hiroyuki O Ishikawa, Hideyuki Takeuchi, Robert S Haltiwanger, and Kenneth D Irvine. Four-jointed is a golgi kinase that phosphorylates a subset of cadherin domains. *Science*, 321(5887):401–404, 2008.

[46] Michael A Simon, Aiguo Xu, Hiroyuki O Ishikawa, and Kenneth D Irvine. Modulation of fat: dachsous binding by the cadherin domain kinase four-jointed. *Current Biology*, 20(9):811–817, 2010.

[47] Madhav Mani, Sidhartha Goyal, Kenneth D Irvine, and Boris I Shraiman. Collective polarization model for gradient sensing via dachsous-fat intercellular signaling. *Proceedings of the National Academy of Sciences*, 110(51):20420–20425, 2013.

[48] Yongqiang Feng and Kenneth D Irvine. Processing and phosphorylation of the fat receptor. *Proceedings of the National Academy of Sciences*, 106(29):11989–11994, 2009.

[49] Dragana Rogulja, Cordelia Rauskolb, and Kenneth D Irvine. Morphogen control of wing growth through the fat signaling pathway. *Developmental cell*, 15(2):309–321, 2008.

[50] Justin A Bosch, Taryn M Sumabat, Yassi Hafezi, Brett J Pellock, Kevin D Gandhi, and Iswar K Hariharan. The drosophila f-box protein fbxl7 binds to the protocadherin fat and regulates dachs localization and hippo signaling. *Elife*, 3:e03383, 2014.

[51] Hitoshi Matakatsu and Seth S Blair. Separating planar cell polarity and hippo pathway activities of the protocadherins fat and dachsous. *Development*, 139(8):1498–1508, 2012.

[52] Flavio Garoia, Daniela Guerra, Maria Cristina Pezzoli, Ana López-Varea, Sandro Cavicchi, and Antonio Garcı&a-Bellido. Cell behaviour of¡ i¿ drosophila¡/i¿¡

i¿ fat¡/i¿ cadherin mutations in wing development. *Mechanisms of development*, 94(1):95–109, 2000.

[53] Laura A Johnston and Peter Gallant. Control of growth and organ size in drosophila. *Bioessays*, 24(1):54–64, 2002.

[54] Dragana Rogulja and Kenneth D Irvine. Regulation of cell proliferation by a morphogen gradient. *Cell*, 123(3):449–461, 2005.

[55] Luis Alberto Baena-Lopez, Isabel Rodríguez, and Antonio Baonza. The tumor suppressor genes dachsous and fat modulate different signalling pathways by regulating dally and dally-like. *Proceedings of the National Academy of Sciences*, 105(28):9645–9650, 2008.

[56] Leonie Enderle and Helen McNeill. Hippo gains weight: Added insights and complexity to pathway control. *Science signaling*, 6(296):re7, 2013.

[57] Peter A Lawrence, Gary Struhl, and José Casal. Do the protocadherins fat and dachsous link up to determine both planar cell polarity and the dimensions of organs? *Nature cell biology*, 10(12):1379–1382, 2008.

[58] Georg Halder and Randy L Johnson. Hippo signaling: growth control and beyond. *Development*, 138(1):9–22, 2011.

[59] Kieran F Harvey, Xiaomeng Zhang, and David M Thomas. The hippo pathway and human cancer. *Nature Reviews Cancer*, 13(4):246–257, 2013.

[60] Binnaz Kucuk Staley and Kenneth D Irvine. Hippo signaling in drosophila: recent advances and insights. *Developmental Dynamics*, 241(1):3–15, 2012.

[61] Felix A Grusche, Helena E Richardson, and Kieran F Harvey. Upstream regulation of the hippo size control pathway. *Current Biology*, 20(13):R574–R582, 2010.

[62] Patrick Müller, Katherine W Rogers, R Yu Shuizi, Michael Brand, and Alexander F Schier. Morphogen transport. *Development*, 140(8):1621–1638, 2013.

[63] Ainhoa Callejo, Aphrodite Bilioni, Emanuela Mollica, Nicole Gorfinkiel, Germán Andrés, Carmen Ibáñez, Carlos Torroja, Laura Doglio, Javier Sierra, and Isabel

Guerrero. Dispatched mediates hedgehog basolateral release to form the long-range morphogenetic gradient in the drosophila wing disk epithelium. *Proceedings of the National Academy of Sciences*, 108(31):12591–12598, 2011.

[64] Pascal P Thérond. Release and transportation of hedgehog molecules. *Current opinion in cell biology*, 24(2):173–180, 2012.

[65] Hans G Othmer and John A Aldridge. The effects of cell density and metabolite flux on cellular dynamics. *Journal of mathematical biology*, 5(2):169–200, 1978.

[66] Bonnie L Bassler. How bacteria talk to each other: regulation of gene expression by quorum sensing. *Current opinion in microbiology*, 2(6):582–587, 1999.

[67] Steve Atkinson, Chien-Yi Chang, R Elizabeth Sockett, Miguel Cámara, and Paul Williams. Quorum sensing in yersinia enterocolitica controls swimming and swarming motility. *Journal of bacteriology*, 188(4):1451–1461, 2006.

[68] AB Goryachev, DJ Toh, and T Lee. Systems analysis of a quorum sensing network: design constraints imposed by the functional requirements, network topology and kinetic constants. *Biosystems*, 83(2):178–187, 2006.

[69] Lars Hufnagel, Aurelio A Teleman, Hervé Rouault, Stephen M Cohen, and Boris I Shraiman. On the mechanism of wing size determination in fly development. *Proceedings of the National Academy of Sciences*, 104(10):3835–3840, 2007.

[70] Matthew C Gibson, Dara A Lehman, and Gerold Schubiger. Lumenal transmission of decapentaplegic in drosophila imaginal discs. *Developmental cell*, 3(3):451–460, 2002.

[71] Felipe-Andrés Ramírez-Weber and Thomas B Kornberg. Cytonemes: cellular processes that project to the principal signaling center in drosophila imaginal discs. *Cell*, 97(5):599–607, 1999.

[72] Patricia Rojas-Ríos, Isabel Guerrero, and Acaimo González-Reyes. Cytoneme-mediated delivery of hedgehog regulates the expression of bone morphogenetic proteins to maintain germline stem cells in drosophila. 2012.

[73] James Briscoe and Pascal P Thérond. The mechanisms of hedgehog signalling and its roles in development and disease. *Nature Reviews Molecular Cell Biology*, 14(7):416–429, 2013.

[74] Marcus Bischoff, Ana-Citlali Gradilla, Irene Seijo, Germán Andrés, Carmen Rodríguez-Navas, Laura González-Méndez, and Isabel Guerrero. Cytonemes are required for the establishment of a normal hedgehog morphogen gradient in drosophila epithelia. *Nature cell biology*, 15(11):1269–1281, 2013.

[75] J Beaudouin, Mario S Mommer, Hans Georg Bock, and Roland Eils. Experiment setups and parameter estimation in fluorescence recovery after photobleaching experiments: A review of current practice. In *Model Based Parameter Estimation*, pages 157–169. Springer, 2013.

[76] M. Kang, C.A. Day, K. Drake, A.K. Kenworthy, and E. DiBenedetto. A generalization of theory for two-dimensional fluorescence recovery after photobleaching applicable to confocal laser scanning microscopes. *Biophysical journal*, 97(5):1501–1511, 2009.

[77] K. Braeckmans, K. Remaut, R.E. Vandenbroucke, B. Lucas, S.C. De Smedt, and J. Demeester. Line frap with the confocal laser scanning microscope for diffusion measurements in small regions of 3-d samples. *Biophysical journal*, 92(6):2172–2183, 2007.

[78] D. Mazza, K. Braeckmans, F. Cella, I. Testa, D. Vercauteren, J. Demeester, S.S. De Smedt, and A. Diaspro. A new frap/frapa method for three-dimensional diffusion measurements based on multiphoton excitation microscopy. *Biophysical journal*, 95(7):3457–3469, 2008.

[79] B.L. Sprague, R.L. Pego, D.A. Stavreva, and J.G. McNally. Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophysical journal*, 86(6):3473–3495, 2004.

[80] P. Hinow, C.E. Rogers, C.E. Barbieri, J.A. Pietenpol, A.K. Kenworthy, and E. DiBenedetto. The dna binding activity of p53 displays reaction-diffusion kinetics. *Biophysical journal*, 91(1):330–342, 2006.

[81] F. Mueller, P. Wach, and J.G. McNally. Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophysical journal*, 94(8):3323–3339, 2008.

[82] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

[83] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, 2010.

[84] Guohui Pan, Yongqiang Feng, Abhijit A Ambegaonkar, Gongping Sun, Matthew Huff, Cordelia Rauskolb, and Kenneth D Irvine. Signal transduction by the fat cytoplasmic domain. *Development*, 140(4):831–842, 2013.

[85] Maria Willecke, Fisun Hamaratoglu, Madhuri Kango-Singh, Ryan Udan, Chiaolin Chen, Chunyao Tao, Xinwei Zhang, and Georg Halder. The fat cadherin acts through the hippo tumor-suppressor pathway to regulate tissue size. *Current Biology*, 16(21):2090–2100, 2006.

[86] Yongqiang Feng and Kenneth D Irvine. Fat and expanded act in parallel to regulate growth through warts. *Proceedings of the National Academy of Sciences*, 104(51):20362–20367, 2007.

[87] Caroline Badouel, Laura Gardano, Nancy Amin, Ankush Garg, Robyn Rosenfeld, Thierry Le Bihan, and Helen McNeill. The ferm-domain protein expanded regulates hippo pathway activity via direct interactions with the transcriptional activator yorkie. *Developmental cell*, 16(3):411–420, 2009.

[88] Fisun Hamaratoglu, Maria Willecke, Madhuri Kango-Singh, Riitta Nolo, Eric Hyun, Chunyao Tao, Hamed Jafar-Nejad, and Georg Halder. The tumour-suppressor genes nf2/merlin and expanded act through hippo signalling to regulate cell proliferation and apoptosis. *Nature cell biology*, 8(1):27–36, 2005.

[89] Douglas A Lauffenburger and Jennifer Linderman. *Receptors: models for binding, trafficking, and signaling*. Oxford University Press, 1993.

[90] Boris N Kholodenko. Cell-signalling dynamics in time and space. *Nature reviews Molecular cell biology*, 7(3):165–176, 2006.

[91] LA Baena-Lopez and Antonio García-Bellido. Control of growth and positional information by the graded vestigial expression pattern in the wing of drosophilamelanogaster. *Proceedings of the National Academy of Sciences*, 103(37):13734–13739, 2006.

[92] Alexander Spirov, Khalid Fahmy, Martina Schneider, Erich Frei, Markus Noll, and Stefan Baumgartner. Formation of the bicoid morphogen gradient: an mrna gradient dictates the protein gradient. *Development*, 136(4):605–614, 2009.

[93] Oliver Grimm, Mathieu Coppey, and Eric Wieschaus. Modelling the bicoid gradient. *Development*, 137(14):2253–2264, 2010.

[94] David M Umulis and Hans G Othmer. The role of mathematical models in understanding pattern formation in developmental biology. *Bulletin of mathematical biology*, 77(5):817–845, 2014.

[95] Hans George Othmer and LE Scriven. Interactions of reaction and diffusion in open systems. *Industrial & Engineering Chemistry Fundamentals*, 8(2):302–313, 1969.

[96] Tosio Kato. Wave operators and similarity for some non-selfadjoint operators. In *Contributions to Functional Analysis*, pages 258–279. Springer, 1966.

[97] Chetan Gadgil, Chang Hyeong Lee, and Hans G Othmer. A stochastic analysis of first-order reaction networks. *Bulletin of mathematical biology*, 67(5):901–946, 2005.

[98] John Crank. *The Mathematics of Diffusion: 2d Ed.* Clarendon Press, 1975.