

THE PHYSICS OF CHROMATIN SILENCING: BI-STABILITY AND FRONT PROPAGATION

BY MOHAMMAD SEDIGHI

A dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements

for the degree of
Doctor of Philosophy
Graduate Program in Physics

Written under the direction of
Anirvan Sengupta
and approved by

New Brunswick, New Jersey

May, 2007

ABSTRACT OF THE DISSERTATION

The Physics of Chromatin Silencing: Bi-stability and Front Propagation

by Mohammad Sedighi

Dissertation Director: Anirvan Sengupta

A mean-field dynamical model of chromatin silencing in budding yeast is provided and the conditions giving rise to two states: one silenced and another un-silenced, is studied. Based on these conditions, the space of control parameters is divided into two distinct regions of mono-stable and bi-stable solutions (the bifurcation diagram). Then, considering both the discrete and continuous versions of the model, the formation of a stable boundary between the silenced and un-silenced areas on DNA is investigated. As a result, a richer phase diagram is provided. The dynamics of the boundary is also studied under different conditions. Consequently, assuming negative feedback due to possible depletion of silencing proteins, the model explains a paradoxical epigenetic behavior of yeast that happens under some mutation. A stochastic treatment of the model is also considered to verify the results of the mean-field approximation and also to understand the role of intrinsic noise at single cell level. This model could be used as a general guide to discuss chromatin silencing in many organisms.

Acknowledgements

This research was done in joint cooperation with BioMaPS Institute, Rutgers University. With special thanks to James Broach, Bradley, Marc Gartenberg, Vincenzo Pirrotta and John Widom for the informative consultations.

Dedication

To my parents.

Table of Contents

Abstract	ii
Acknowledgements	iii
Dedication	iv
 1. Introduction	 1
1.1. Biological Background	1
1.1.1. DNA and Proteins	1
1.1.2. Higher Degrees of DNA Configuration	2
1.2. Gene Silencing	4
1.2.1. Position Effect Variegation	5
1.2.2. Silencing In Budding Yeast, <i>S. Cerevisiae</i>	6
1.3. A Stepwise Model for Silencing	8
1.4. Experimental Observations	11
1.4.1. Bi-stability and Epigenetic Inheritance in Budding Yeast	11
1.4.2. The Effect of Low Acetylation	12
1.5. The Mathematical Analysis	13
 2. Bifurcation Analysis of a Model for Silencing	 15
2.1. Dynamical Equations	15
2.2. Uniform Solutions	16
2.3. The Hysteresis Behavior	23
2.4. The Bifurcation Diagram	25
2.4.1. The Critical Point	28

2.4.2.	The Hysteresis Behavior in Phase Space	29
2.4.3.	The Role of Non-local Interaction Factor, γ and Degree of Cooperativity, n	30
3.	Non-uniform Solutions and Front Propagation	33
3.1.	The Continuum Limit	34
3.1.1.	Wave Solutions and Zero Velocity Line	35
3.2.	The Discrete Model	38
3.3.	Front Dynamics	40
3.4.	The Role of Finite Supply of Sir Proteins	42
4.	The Role of Stochasticity	48
4.1.	Extrinsic Fluctuations and Front Robustness	49
4.2.	The Intrinsic Stochastic Treatment of the System	53
5.	Alternative Possible Non-linearity in the System	60
6.	Biological Consequences of the Model	63
7.	Discussion	67
8.	Appendix A: Uniform Fixed Points and Their Stability	69
9.	Appendix B: Numerical Methods in Discrete Model Approach	73
10.	Appendix C: Monte Carlo Simulation of the Stochastic Model	75
	References	77
	Vita	80

Chapter 1

Introduction

This chapter is mainly devoted to the preliminary biological knowledge required to understand the phenomenon of *chromatin silencing*. As a matter of fact, most of the material in the following sections, can be found in any introductory *molecular biology* text book [1, 2] unless otherwise referenced.

1.1 Biological Background

1.1.1 DNA and Proteins

The genetic information in living cells is stored in a long double-stranded helical macromolecule called DNA. This information, or the *genome*, is encoded along each strand of DNA in a language of four letters represented by four alternating units called *nucleotides*. In a more detailed perspective, each strand of DNA is a chain consisting of a repeating sugar/phosphate backbone and a *base* attached to each sugar unit. The base is chosen from only four different organic compounds: adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T). Base is the only part of each block of DNA that changes along this sequential structure. Then each strand is attached to its complementary strand through Hydrogen bonds between bases. Note that, hydrogen bonds happen between base pairs of (A,T) and (C, G) only. In the common stable double helical conformation of DNA the length between successive base pairs is 1 bp=.34 nm, which is also used as unit of length. As a consequence of this complimentary base pairing the genetic

information kept in each of the strands is doubled in DNA.

This genetic information is mainly used to build proteins, the molecules responsible for vital interactions inside the cell. The part of genome that contains the information is referred to as the *gene* of the protein. Proteins are governing all the interactions inside the cell which include, but are not limited to, extracting genetic information from DNA, building other proteins from their genetic code and also mutual interactions leading the cell to behave in a desired way. As an example, when a particular protein is produced, an agent protein called RNA-polymerase, with the help of other proteins locates the corresponding gene among the whole genome, attaches to the specified sequence on DNA and start making a single stranded copy of the gene called RNA. However RNA is not just the exact copy of some section of the single stranded DNA. There are two main differences in their structures: sugar/phosphate backbone of RNA uses a different type of sugar and uracil (U), replaces thymine (T) for the bases. RNA molecule is then used directly to build the required protein. The process of producing a RNA copy of the gene by RNA-polymerase is called *transcription*.

There are several mechanisms inside the cell to control interactions among proteins during different stages of cell's life or to respond to changes in environment. One of the known mechanisms is regulation of gene transcription. This is done by employing proteins that may aid or stop the attachment of RNA-polymerase to the desired site on DNA. This can totally depend on the environment and the same cell may then demonstrate various behaviors as a consequence of different states of gene activity.

1.1.2 Higher Degrees of DNA Configuration

The double helical configuration of DNA is how genetic material is kept inside single celled species like bacteria. In most of the multi cellular species, however, the DNA is kept inside an enveloped structure called nucleus (about 10

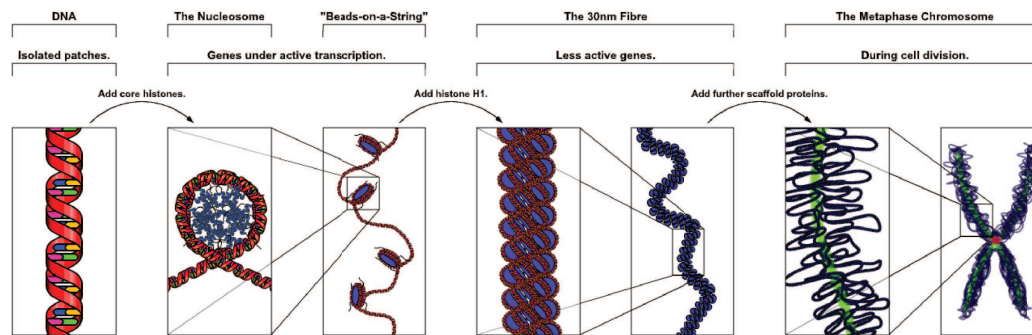


Figure 1.1: Hierarchy of DNA configuration in eukaryotic cells.

micrometers in diameter). A cell that contains a nucleus is called an *eukaryotic cell*, and the species with this form of cells are called *eukaryotes*. In order to compactify the long DNA inside nucleus, eukaryotic cells employ structural and configurational proteins such as *histones*. Eight histone proteins make a cylindrical complex, simply called the histone-octamer (11nm in diameter). In the very basic level, eukaryotic DNA is wrapped around cylindrical histone-octamers like a thread around spools, each for about 146 base pairs. Each histone-octamer along with the DNA around it, is then called a *nucleosome*. At this level, under the microscope these nucleosomes on DNA look like beads on a string (Fig.1.1)¹. However, there are also several higher levels of eukaryotic DNA condensation utilizing different structural proteins. Biologists use the word *chromatin* when they refer to the highly complex mixture of DNA and structural proteins (particularly histones). Note that, higher order organisms like human being have more than one linear DNA molecule, where each single DNA macro-molecule exhibits its own hierarchy of compactification and is called a *chromosome*.

The degree of DNA packaging varies as the cells goes through different stages

¹Courtesy of Richard Wheeler (Zephyris), wikipedia.com; the creator of this work, hereby grants the permission to copy, distribute and/or modify this document under the terms of the *GNU Free Documentation License*, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

of its life cycle. Moreover, different regions of DNA inside nucleus may exhibit different degrees of complexity and condensation. As a matter of fact, during cell divisions, chromatin is in its utmost level of compactification (Fig.1.1). However, during *interphase*, the period between nuclear divisions, eukaryotic chromatin can be divided into two distinct regions based on their degree of condensation. *Heterochromatin*, which refers to highly condensed and packed areas and *euchromatin* referring to lightly condensed and dispersed parts.

1.2 Gene Silencing

In condensed heterochromatin domains, nucleosomes are so packed in high order structures that they are not normally accessible to proteins for transcription, thus not transcriptionally *active*. The size of these regions can be from several kilo bases to several hundred kilo bases and can even cover the whole chromosome. In contrast, euchromatin regions are less compact and transcriptionally active.

Therefore, the formation of heterochromatin can be also considered as a way of *silencing* the expression of a number of adjacent genes. In particular, although all cells in a multi-cellular organism contain the same copy of DNA and the exact same genetic information, since variant regions of genome may be silenced in different cells, distinct functional identities exist throughout the body of organism. It is said in this case that, all the cells are of the same *genotype* but different *phenotypes*. Genotype refers to the genetic makeup, whereas phenotype points to the appearance of the organism such as color, size, behavior, etc.

Furthermore, in many circumstances, the structural organization of chromatin will be inherited to new cells generated through cell divisions. As a result, silencing also plays a crucial role in multi-cellular development by stabilizing gene

expression patterns in specialized cells at early stages and maintaining their identities throughout their life. One example of this, is the cell type dependent silencing of Hox genes, important in development of body plans, by the Polycomb group of proteins [4].

As it was discussed above, one of the interesting aspects of developmental processes is that one could get multiple heritable cell fates without irreversible changes to the genetic information. Heritable differences in cellular behavior or phenotype, despite having the same genetic information, is called *epigenetic phenomenon*. Apart from its fundamental role in development, epigenetic effects are of great importance in certain diseases like cancer [3]. Note that there are several mechanisms that lead to epigenetic effects and only one of these mechanisms is transcriptional silencing.

1.2.1 Position Effect Variegation

As we discussed, whether a gene is expressed or silenced, depends on its position along the DNA in eukaryotic cells. In other words, if one relocates an active gene, experimentally, from euchromatin regions to heterochromatin regions, it gets silenced. The opposite also holds when a gene is moved from a silenced region to an un-silenced region.

Since the *position effects*, mentioned above, play a crucial rule in the behavior and epigenetic identity of a cell; it is important to understand how the boundaries between euchromatin and heterochromatin regions are determined on DNA. It is known, that in many cases this boundary is pinpointed by some *boundary elements* along chromosome. The boundary element is therefore, any structural factor that hinders the *spreading* of silencing at the desired position on DNA. In some cases, however, there are no special boundary elements to precisely determine the border between two regions. In these circumstances, the boundary is not fixed and silenced region can expand into or retract from active region on a seemingly

random basis. However when the formation of boundaries has been settled at early stages of development it will be preserved during the cell's life and through cell divisions. Note that, there is always a low chance of random shift in the boundary in either direction at any point of cells life.

As a consequence of this dynamical behavior of boundary, it is observed that genes that are located near the boundary can switch states from active to silenced and vice versa at small frequencies. This phenomenon is called *position effect variegation*. This effect has been detected in many organisms such as *Drosophila*, the fruit fly. There is a gene in *Drosophila* which is responsible for the red color of its eyes. In other words, if this gene is not active, the eyes will look white. Now, for flies when this gene is active but has been positioned near heterochromatin, the eyes include patches of both red and white colors rather than being entirely red. The red regions represent cells with the active pigment gene and white spots represent cells with silenced pigment gene. In other words, when the boundary between chromatin regions is first being formed, there is a chance for the boundary to shift and heterochromatin covers the pigment gene. This configuration then will be inherited stably through many generations resulting in contiguous patches of red or white.

1.2.2 Silencing In Budding Yeast, *S. Cerevisiae*

Observations on silenced areas of DNA in budding yeast, *Saccharomyces Cerevisiae* have played an important role in understanding how chromatin silencing works. We devote this section to an introduction on *S. Cerevisiae*, since our research is based on a model of gene silencing in this organism.

Budding yeast can be found in two forms: *haploid* or *diploid*. Haploid cells simply contain only one set of chromosomes. There are two types of haploid cells, type a and type α . Diploids, on the other hand, are made by conjugation of the two different types of haploid cells; hence contain two sets of chromosomes. Types

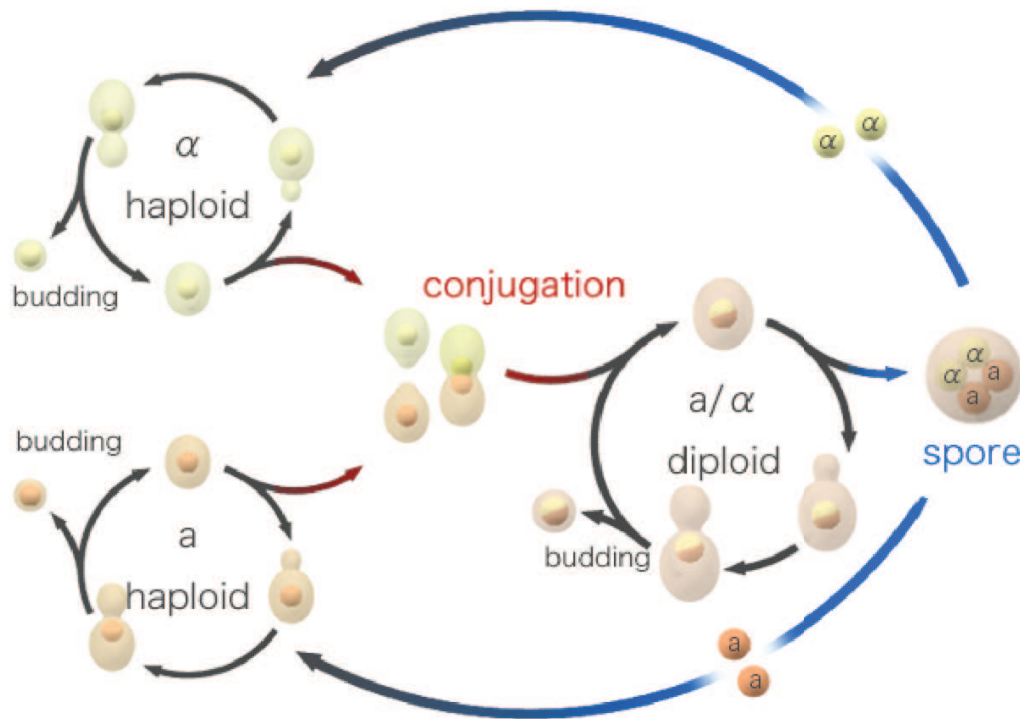


Figure 1.2: The life cycle of budding yeast, *S. Cerevisiae*.

a and α of haploid cells can also be considered as two opposite sex types; and their fusion is also called *mating*. Wild-type haploid cells have a high probability of switching their *mating type* after each cell division (via budding), from a to α and vice versa. In starving situation, haploid cells usually die, but diploid cells *sporulate* and generate four spores. Spores are dormant state of haploid cells and resistant to harsh conditions. When there is availability of nutrients, spores germinate and grow to normal haploid cells (Fig.1.2)².

There are three sections on chromosome III in yeast which are responsible for switching and stabilizing mating types. Two of these sections, each located near one end of the chromosome, are always silenced. They are called *silent mating*

²Courtesy of wikipedia.com; the creator of this work, hereby grants the permission to copy, distribute and/or modify this document under the terms of the *GNU Free Documentation License*, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

loci and noted as *HML* and *HMR* loci and contain copies of genes that decide α -type and a-type identity, respectively. There is a locus in the middle, called *MAT* (mating type locus), which is always active and decides the mating type identity of the haploid cell. Through generations of yeast, the genes from *HML α* and *HMRa* loci are alternatively transferred to the *MAT* locus, causing the change in its mating type. In other words, when *MAT* locus contains *HML α* genetic sequence, yeast acts as an α -type cell and when it contains *HMRa* sequence, it behaves as an a-type cell. This process happens through a recombination mediated process called *mating type switching*.

It is believed that the silencing of *HML* and *HMR* loci originates from DNA sequences next to them called, *silencers*. So if the silencer sequences are experimentally removed, both *HML α* and *HMRa* genes are active. In this case, the haploid cells behave like an α/a diploid cell and are not able to mate. This defective behavior can be used in experiments to detect any deficiency in repression of *HML* and *HMR* sequences on yeast DNA.

Other than the silent mating loci, there are other types of regions on DNA that are silenced. These regions include the *telomeres*, which are the highly condensed ends of the chromosome. Both the telomeres and mating-type loci demonstrate the same features as one expects from the silenced heterochromatin in higher eukaryotes.

1.3 A Stepwise Model for Silencing

Many mechanisms has been proposed for silencing in diverse organisms [5], however, one can always find similar features between these models. In the general model, there is usually a region that *nucleates* silencing by recruiting a silencing complex incorporating a histone modifying enzyme. Modification of histones

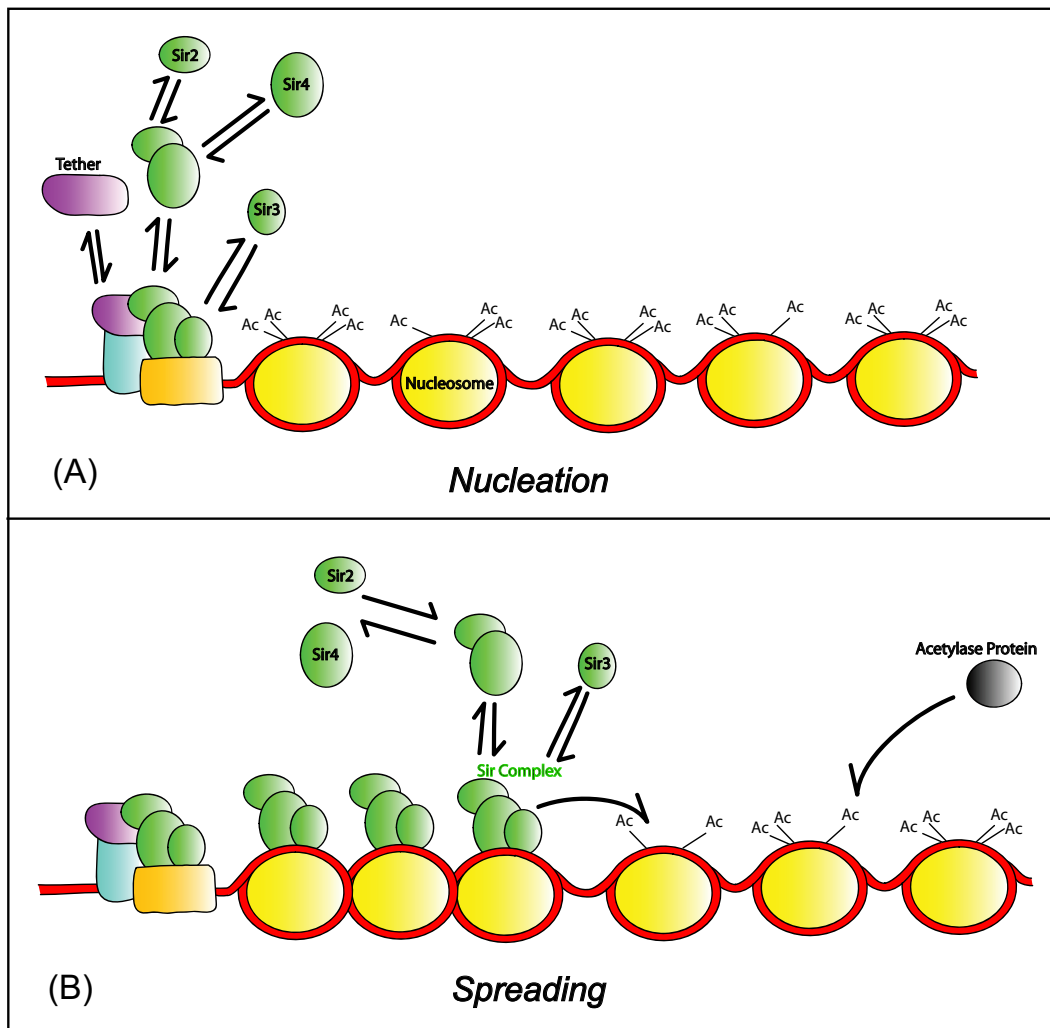


Figure 1.3: A model for nucleation and spreading of silencing in budding yeast, *S. cerevisiae*.

makes that region of chromosome more prone to binding to components of silencing complex, which, in turn, recruits further histone modifying enzymes. That is how the process propagates till it meets some boundary element (or the system reaches a stationary state due to exhaustion of one of the components of the silencing complex, as will be discussed later) (Fig. 1.3).

The mechanism by which silencing nucleates and spreads in budding yeast is relatively well investigated [5, 6] and provides a concrete example of the more

general model mentioned above. It is known that the *Silenced Information Regulator* (SIR) proteins are the main players in gene silencing at telomeres and silent mating loci in yeast. There are four Sir proteins involved in this process, simply called Sir1, Sir2, Sir3 and Sir4. The role of Sir1 is different as it only cooperates in *nucleation* of silencing at the silent mating loci, however the rest of Sir family have also important roles in *spreading* of silencing at both the telomeres and at silent mating loci.

To be concrete, from now on we focus our discussion to the silencing at the silent mating loci. As it was discussed, the model for step-wise gene silencing in *S. cerevisiae*, also posits that silencing happens in two distinct steps: *nucleation* and *spreading*. In nucleation, with the help of site-specific DNA binding proteins (like Rap1) and with Sir1 as a tether, Sir2, Sir3 and Sir4 will form a Sir Complex on the nucleation site (Fig. 1.3.A.). Sir2, a crucial member of this process, works as a *de-acetylase* enzyme. In other words, it can help removing *acetyl* groups from certain parts of nearby histones. Consequently, de-acetylation of the neighboring histones will make binding of Sir3/Sir4 sub-complex easier in the neighborhood of the original nucleation site. Sir3/Sir4 sub-complex, in turn, would recruit more Sir2. Hence, the spreading starts. More de-acetylation of histones improves the recruitment of other Sir proteins and formation of more stable complexes on neighboring sites. If histone de-acetylation is transferred further on, it will result in spreading of silencing to even distal sites (Fig. 1.3.B.). Note that, although the nucleation step is different in telomeric silencing, the process of spreading seems to be very similar [8].

1.4 Experimental Observations

1.4.1 Bi-stability and Epigenetic Inheritance in Budding Yeast

In the wild type budding yeast, the regions that are silenced are, typically, always silenced. As it was discussed in last section, Sir1 proteins are some of the main elements in *nucleation* of silencing at *HMR/HML* loci. As a matter of fact, if Sir1 is missing, *nucleation* effect at the silent mating loci is either absent or very weak. In an experiment in 1989 by L. Pillus and J. Rine [9], it was found that in *sir1* mutants (where the nucleation effect is defective if not absent), there is still a chance for *HMR/HML* loci to be repressed. In other words, silencing can also happen without efficient nucleation. Then, the individual yeast cells would represent two distinguishable type of cells (phenotypes). In this case, in a large number of yeast cells, on average 20% of the cells exhibit repression and can mate like a normal haploid cell, whereas the other 80% cells are de-repressed at *HMR/HML* loci and behave like diploids. It was also found that, both epigenetic states are stable to small fluctuations and are conserved through many generations of yeast. In fact, it was observed that switching from un-silenced to silenced state occurs approximately once in every 250 consecutive cell divisions, or with the small probability of 4×10^{-3} [9]. The observation and results suggest that the system is actually in a *bi-stable* regime, where two opposite stable states can exist under the same conditions.

This kind of epigenetic switches between bi-stable states has received much scientific attention in prokaryotes. Multiple phenotypes are usually represented as multiple stable equilibrium points in deterministic descriptions of the biochemical dynamics. For instance, computational modeling of lambda phage [10] has played a crucial role in the development of systems biology [11, 12]. From the response of lac operon in the presence of TMG [13, 14, 15] to synthetic genetic networks

like the toggle switch [16], mathematical analysis has been an integral part of understanding such phenomena. In particular, the biological model, in each of these examples, provides a mechanism of positive feedback. However, positive feedback is not sufficient to guarantee multi-stability, but essential for giving rise to non-trivial epigenetic states.

1.4.2 The Effect of Low Acetylation

As it was discussed in the stepwise model of silencing in section 1.2, Sir2 has de-acetylating enzymatic activity. This is a crucial part of the silencing process since less acetylated nucleosomes are better places for Sir complex to attach. Then, apparently, the presence of acetyl groups on parts of histones in nucleosomes makes DNA less amenable to binding of Sir proteins. It is observed that near the silenced regions in yeast, because of Sir2 activity, more histones are de-acetylated as opposed to distal nucleosomes which are more acetylated. As a matter of fact, there is an increasing gradient of acetylation as one moves along DNA, away from the heterochromatin region [25, 26]. The acetylation activity, however, is referred to an acetylase protein called, Sas2 protein. In particular, Sas2 along with Sas4 and Sas5 proteins comprise a protein complex called SAS-I which is linked to histone acetylation in yeast [27]. Note that, Sas4 and Sas5 proteins are both required for maximal SAS-I acetylation activity.

One would normally expect that when *SAS2* gene is mutated (where acetylation is defective or absent) the silencing process should prevail. However, researchers have observed many different and opposing effects of *sas2* mutations. In particular, deletion of *SAS2* gene from DNA decreases the silencing process at regions near telomere and at silent mating loci rather than helping it [28]. In contrast, when SAS-I acetylation activity is eliminated, the cell loses its bi-stability at mating-type loci and demonstrates an *intermediate* state which is not either silenced or un-silenced [22]. This intermediate state, can be considered as

a *porous* heterochromatin, where there are many random un-silenced spots inside a silenced region on DNA. This is especially interesting when we compare it to *SIR1* gene mutants when the cells can exhibit only one of the two stable states, silenced or un-silenced.

1.5 The Mathematical Analysis

The crucial aspect of analysis of a mathematical model of epigenetic switches is computing the *bifurcation diagram*, which tells us what region in the space of control parameters is actually associated with bi-stability. The bifurcation diagram also indicates the qualitative behavior of the system when perturbed (or mutated) in a particular manner as in the low acetylation cases discussed before. In contrast to prokaryotic epigenetic switches mentioned at the end of section 1.3.1, modeling eukaryotic epigenetic silencing involves studying a spatially extended bi-stable system. Hence, the system shows interesting phenomena, like front propagation, allowing for a richer bifurcation diagram.

In this dissertation, we introduce a mathematical model of step-wise heterochromatin silencing. A mean-field description of the dynamics explains many features of the real system. Epigenetic states, in the absence of nucleation, can be explained as a consequence of the existence of two stable uniform static solutions: the un-silenced/hyper-acetylated state and silenced/hypo-acetylated states on DNA. Studying the conditions under which the positive reinforcement inherent in the proposed silencing mechanism is strong enough to give rise to bi-stability and to cause stable inheritance of chromatin configurations of the two phenotypes is one of the main goals of this dissertation. In addition, the conditions required for static fronts will set additional constraints on the model.

Moreover, a stochastic treatment of the model is also considered. Fluctuations in bio-molecular networks has been the subject of many research activities

recently [31]. To analyze single cell data, one needs not only how the deterministic model behaves but also how noise in various quantities affects expression. A stochastic version of the model, a lattice model with local states of acetylation, and Sir occupancy, will be studied by direct simulation. However, as seen in studies of yeast gene expression [36, 37], extrinsic noise, equivalent to fluctuations in the parameters themselves, often dominates over intrinsic fluctuations of the processes described here with fixed parameters. Hence, to study this properly, we will need to add a free parameter each characterizing the slow noise in the control parameters (such as concentrations) for modeling the effect of cell to cell variation of Sir proteins and acetylases. Instead, we will discuss the effect of these fluctuations at the limit when the extrinsic part of noise is much slower and stronger than the intrinsic part. At this regime, one can average over the intrinsic noise and use the mean-field approximation. At the end, we propose experiments designed to test the ideas discussed in this dissertation.

Chapter 2

Bifurcation Analysis of a Model for Silencing

2.1 Dynamical Equations

The purpose of this section is to formulate a quantitative version of the conventional biological model of step-wise formation of silenced chromatin (Fig. 1.3), which was discussed in section 1.2. The main parameters involved in final equations are A , *the local degree of acetylation of histones* and S , *the local probability of occupation by Sir complex (Sir2, Sir3 and Sir4)*, both of which could depend on time, as well as on their position on DNA. DNA is represented as a one-dimensional lattice, where each site on the lattice represents either one or more nucleosomes. So in other words, $S_i(t)$ on this lattice, is a number between 0 and 1, representing fractional number of Sir complexes at site i at time t . Fractional degree of acetylation, $A_i(t)$, is defined in the same way. Writing chemical equations, in the mean-field treatment of the system, we get,

$$\frac{dS_i(t)}{dt} = \rho_i(t)(1 - S_i(t))f(1 - A_i(t)) - \eta S_i(t) \quad (2.1)$$

$$\frac{dA_i(t)}{dt} = \alpha(1 - A_i(t))(1 - S_i(t)) - (\lambda + \sum_j \gamma_{ij}S_j(t))A_i(t) \quad (2.2)$$

Note that all the parameters in the above equations are non-negative numbers. In equation (2.1), on the right hand side, the first term is the creation term and the next one is degradation term. The 3-D concentration of ambient Sir complex at site i is denoted by $\rho_i(t)$, which may change as the system evolves. Since free Sir proteins in the environment do not form Sir complexes by themselves, this

quantity actually represents a function of concentrations of all components (Sir2, Sir3 and Sir4) that are ready to make a Sir complex on the site. For example, in the simplest case, when each protein is in low abundance, this function would be proportional to the product of the three concentrations. However, throughout this dissertation we will never need to go into these kind of details. The function $f(x)$ dictates the cooperativity in Sir complex binding and should be a monotonically increasing function of x , $0 \leq x \leq 1$. As the simplest case, we use $f(x) = x^n$, where n is the degree of cooperativity between de-acetylated histones in recruiting Sir proteins. At last, η is the degradation rate of bound Sir complexes. In equation (2.2), the same as equation (2.1), on *RHS*, the first term advocates creation and next one degradation. α represents the constant acetylation rate¹. In the second term, the summation accounts for the contribution of adjacent Sir complexes in de-acetylation of site i . Since Sir complex is only capable of de-acetylation of sites in its neighborhood, γ_{ij} is assumed to drop significantly as $|i - j|$ gets large. In addition, γ_{ij} is assumed to be *symmetric* with respect to i and j , i.e. $\gamma_{ij} = \gamma_{ji}$. Finally, λ is the rate of de-acetylation from the rest of de-acetylase proteins. This rate is assumed to be a constant both in time and position.

In a more general model, all rates can be position dependent. We neglected this effect for η , α and λ by assuming homogeneous concentration of participating enzymes and no drastic conformational changes in DNA that can affect these chemical rates.

2.2 Uniform Solutions

One could analyze the uniform static solutions of these equations and study the stability. Dropping all i indices and replacing the non-local term $\sum_j \gamma_{ij} S_j$ with

¹To be more general, the acetylation term could be $\alpha(1 - A_i)(1 + \sigma - S_i)$ allowing acetylation of histones in silencing complex bound nucleosomes. However, as will be discussed in appendix A, adding this process does not make much of a qualitative difference.

γS , we can rewrite equations as:

$$\frac{dS(t)}{dt} = \rho(t)(1 - S(t))f(1 - A(t)) - \eta S(t) \quad (2.3)$$

$$\frac{dA(t)}{dt} = \alpha(1 - A(t))(1 - S(t)) - (\lambda + \gamma S(t))A(t) \quad (2.4)$$

The stationary states are obtained by solving the algebraic equations produced by setting time derivatives to zero. We analyze first the case where *available* SIR concentrations are kept at a constant level (for example, if creation and degradation rates of SIR proteins are high), meaning $\rho(t)$ in equation (2.3) is assumed to be a given time independent number, ρ . So in terms of scaled parameters $\bar{\rho} = \rho/\eta$, $\bar{\alpha} = \alpha/\lambda$ and $\bar{\gamma} = \gamma/\lambda$, we have:

$$\bar{\rho}(1 - S)f(1 - A) - S = 0 \quad (2.5)$$

$$\bar{\alpha}(1 - A)(1 - S) - (1 + \bar{\gamma}S)A = 0 \quad (2.6)$$

For $f(x) = x^n$, we plot the graph of above nullcline equations for different values of n and chemical parameters. The intersections of two curves represent the fixed points. We find that for $n > 1$ depending on values of chemical parameters we can get either one or three fixed points (Fig.2.1). There is a possibility of having two fixed point when at one of the points two curves are tangent to each other. We will not discuss this case since it is not a physical possibility and will only treat it as a transient state between two physical cases. For $n = 1$, no matter how one chooses chemical parameters, there can be only one fixed point (Fig.2.2). In the three fixed point regime, as it is shown in the Fig.2.3, always the middle one is an unstable saddle point. The other two are stable, or in other words, we are at a bi-stable regime as it could be seen by local analysis (appendix A). One of the two stable states has a lower acetylation and a higher chance of repression, which represents a silenced state (heterochromatin), while the other one with a

higher degree of acetylation and higher chance of de-repression represents an un-silenced state (euchromatin) (Fig.2.1). A more complete mathematical discussion of fixed points and their stability is given in appendix A.

From here on, in our discussion we assume $f(x) = x^n$, $n > 1$ to guarantee possibility of bi-stability in the chemical parameters space. Now, in the bi-stable parameter regime, we have two stable fixed points, silenced(heterochromatin) and un-silenced(euchromatin). If $\bar{\alpha}$, the rate of acetylation, is kept constant, as the value of $\bar{\rho}$, the rate of Sir recruitment, increases the un-silenced (euchromatin) point and middle transient state move towards each other and at some limit value of $\bar{\rho}$, they become one single point where two curve are tangent to each other. Pushing $\bar{\rho}$ to even higher values, leads to a single silenced (heterochromatin) regime i.e. at some upper limit the systems falls from bi-stability to mono-stability (Fig.2.1). In contrast, if $\bar{\rho}$ decreases, the heterochromatin point and middle transient point approach each other and after dropping below a limit the systems becomes mono-stable, the euchromatin state (Fig.2.1). The same behavior can be seen for different values of $\bar{\alpha}$ when $\bar{\rho}$ is kept constant.

Now, suppose that we have decreased $\bar{\alpha}$ to a value where there is only one fixed point, which is a silenced state. Now, if we decrease the value of $\bar{\rho}$ too, we may again fall back into a three fixed point (bi-stable) regime. However, this will not happen when $\bar{\alpha}$ is very small. For small values of $\bar{\alpha}$, if we scan over different values of $\bar{\rho}$, the system will keep showing only one fixed point. In other words, the system transforms from a heterochromatin fixed point to a euchromatin fixed point continuously without crossing the three fixed point regime (Fig.2.4). This behavior, as will be discussed later, can provide us with an explanation to results of SAS-I elimination experiment in section 1.3.2.

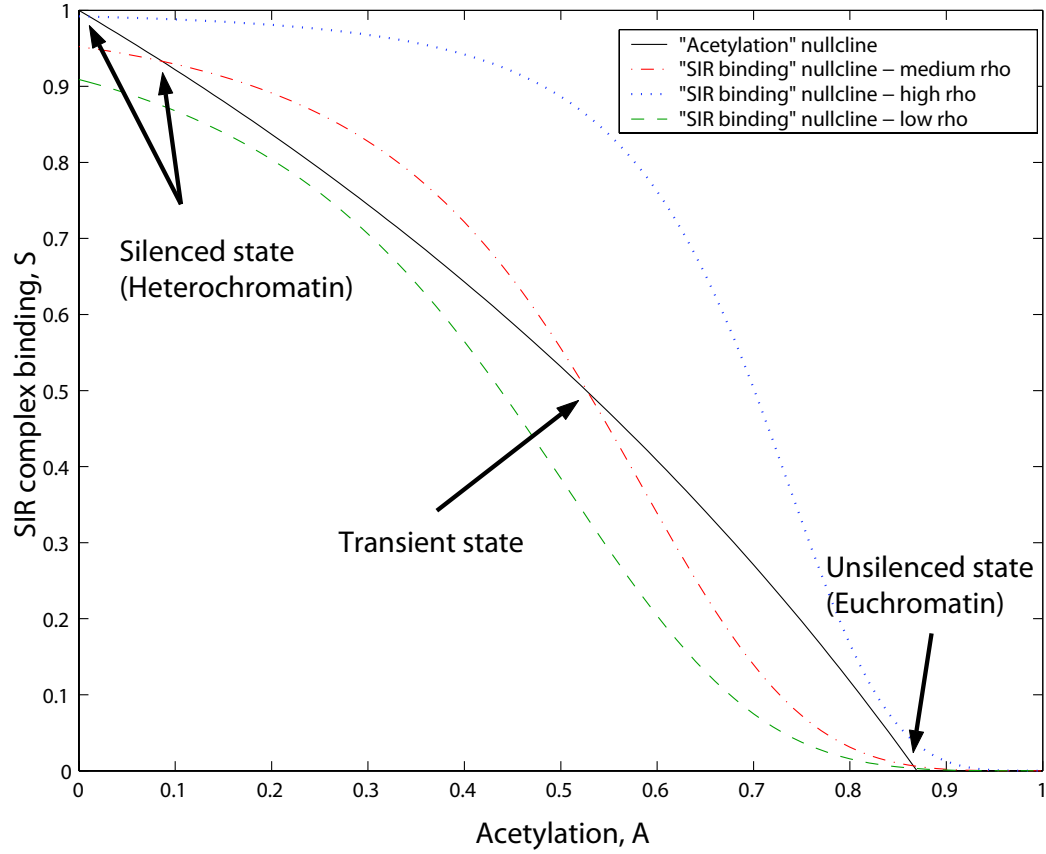


Figure 2.1: The intersections of nullcline curves, three representing static “SIR binding” (dashed lines) and the other one static “Acetylation” (solid line), show fixed points of the system. All graphs are plotted with $f(x) = x^4$, $\bar{\gamma} = 4$ and $\bar{\alpha} = 6.67$. Ambient Sir complex concentration acts as a switch for the bi-stable system. This graph shows how low concentration of Sir complex pushes the system towards euchromatin and high concentration of Sir complex pushes it towards heterochromatin solution. For high, intermediate, low concentrations of Sir complex respectively $\bar{\rho} = 125, 20$ and 10 .

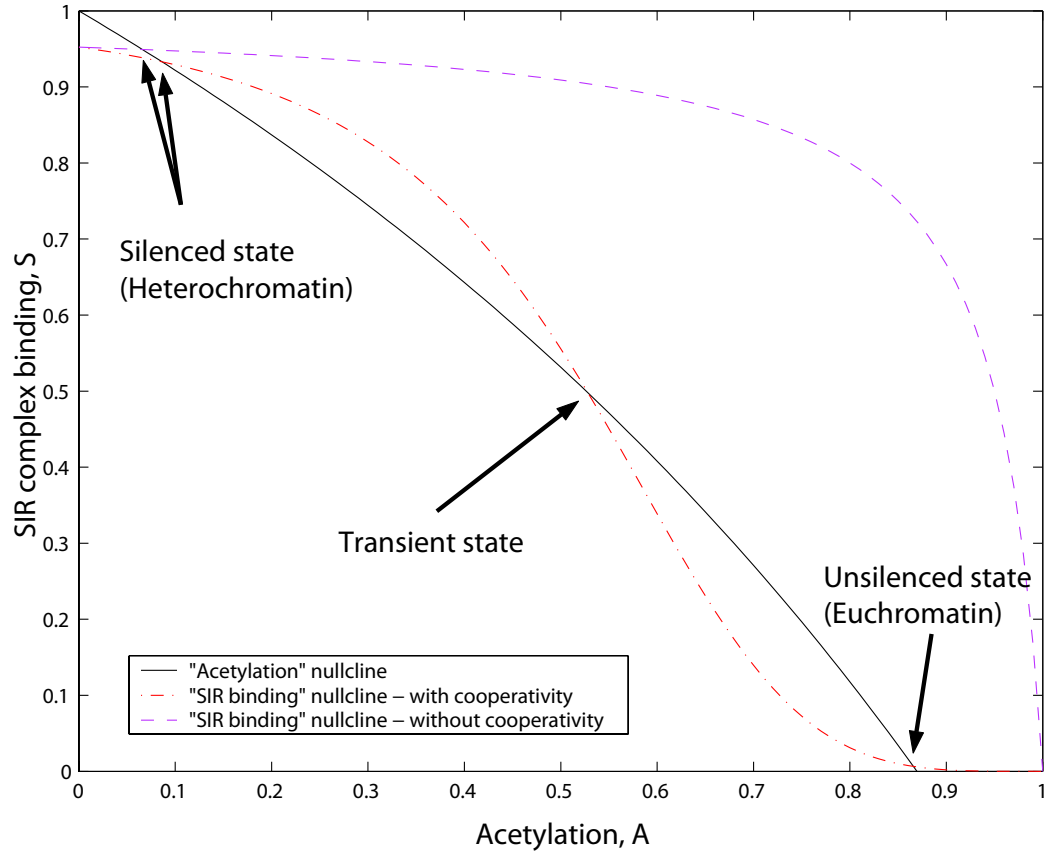


Figure 2.2: The intersections of nullcline curves, two representing static “SIR binding” (dashed lines) and the other one static “Acetylation” (solid line), show fixed points of the system. All graphs are plotted with $\bar{\gamma} = 4$, $\bar{\alpha} = 6.67$ and $\bar{\rho} = 20$. When there is no de-acetylation cooperativity in Sir complex binding, $f(x)$ is linear and there is no bi-stability (only one fixed point solution). But for $f(x) = x^4$ curve, as a result of cooperativity there can be bi-stability (three fixed point solutions).

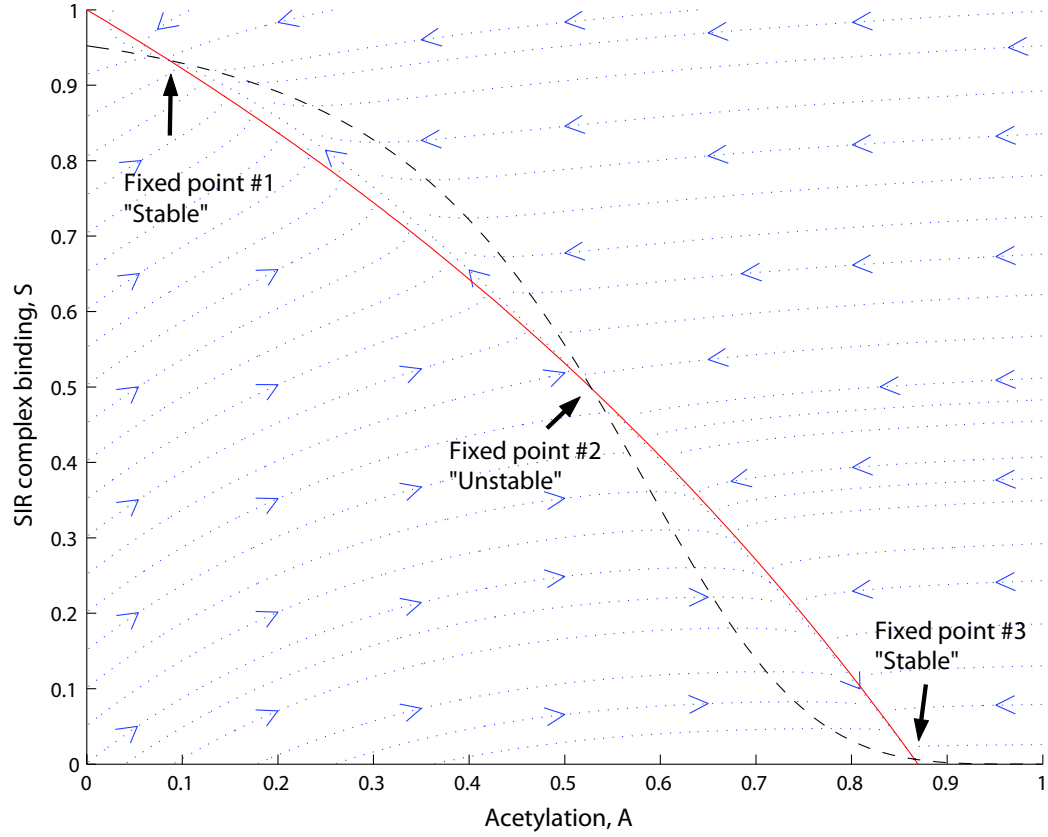


Figure 2.3: The intersections of nullcline curves, one representing static “SIR binding” (dashed line) and the other one static “Acetylation” (solid line), show fixed points of the system. The graphs are plotted with $f(x) = x^4$, $\bar{\gamma} = 4$, $\bar{\alpha} = 6.67$ and $\bar{\rho} = 20$. The flow diagram, demonstrates clearly why the system is bi-stable when there are three fixed point solutions. The middle fixed point is always an unstable saddle point while the other two are stable fixed points.

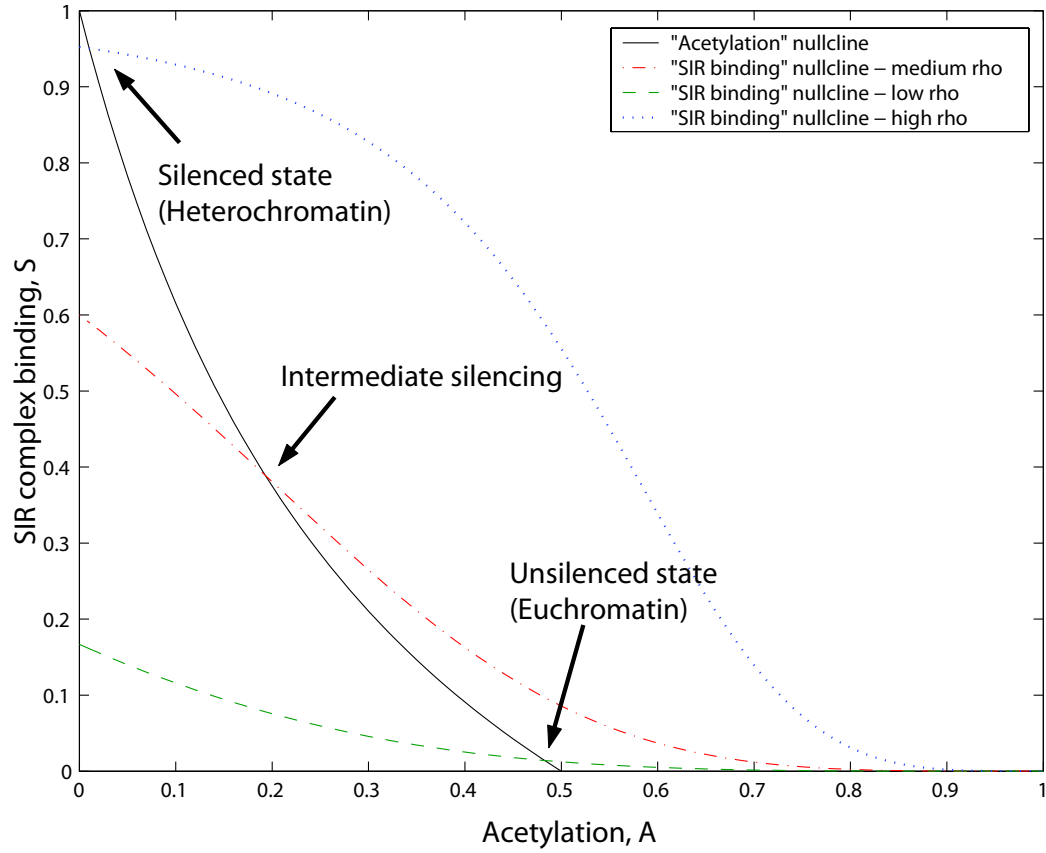


Figure 2.4: The intersections of nullcline curves, three representing static “SIR binding” (dashed lines) and the other one static “Acetylation” (solid line), show fixed points of the system. The graphs are plotted with $f(x) = x^4$, $\bar{\gamma} = 4$, $\bar{\alpha} = 1$ and $\bar{\rho} = 10, 20$ and 125 . At low values of $\bar{\alpha}$ the system loses its bi-stability feature and exhibits a continuous change from euchromatin to heterochromatin state as $\bar{\rho}$ value increases.

2.3 The Hysteresis Behavior

One might ask which of the two stable fixed points is actually chosen by system in the bi-stable regime. In order to answer this question and also to understand the behavior of the system in this regime, we need to employ a dynamical simulation of the system. Going back to equations (2.3) and (2.4) and keeping $\bar{\alpha}$ at some high constant value, we start from some initial value of $\bar{\rho}$. By gradually increasing and decreasing $\bar{\rho}$, at each point we check the final equilibrium state chosen by system. Not very surprisingly, the system demonstrates a *hysteresis* behavior. In other words, the behavior of the system depends on its history i.e. it always keeps the same fixed point from which it entered the bi-stable region. For instance, if we start from a small value of $\bar{\rho}$ at which there is only one un-silenced (euchromatin) fixed point and start increasing $\bar{\rho}$, when the system moves into the bi-stable region it still keeps staying in the un-silenced state. This will be the case no matter how $\bar{\rho}$ changes and even works for discontinuous jumps. However, if $\bar{\rho}$ gets so large that we pass the bi-stable region, the only possible fixed point becomes a silenced (heterochromatin) state and the system abruptly transforms to a silenced state. This discontinuous switch of states is in direct contrast to the continuous change of states that we mentioned for very low values of $\bar{\alpha}$. Now, if we reverse the direction of movement and start decreasing $\bar{\rho}$, as we move back into the bi-stable region, the system keeps staying in silenced state until we exit from the other end and enter the un-silenced region where we suddenly observe a sharp change in the state of system.

To clarify this behavior, one can make a plot of equilibrium value of S versus $\bar{\rho}$ for constant values of $\bar{\alpha}$. Using equation (2.6), we can solve $1 - A$ in terms of the other variables as follows:

$$1 - A = \frac{1 + \bar{\gamma}S}{\bar{\alpha} + 1 + (\bar{\gamma} - \bar{\alpha})S} \quad (2.7)$$

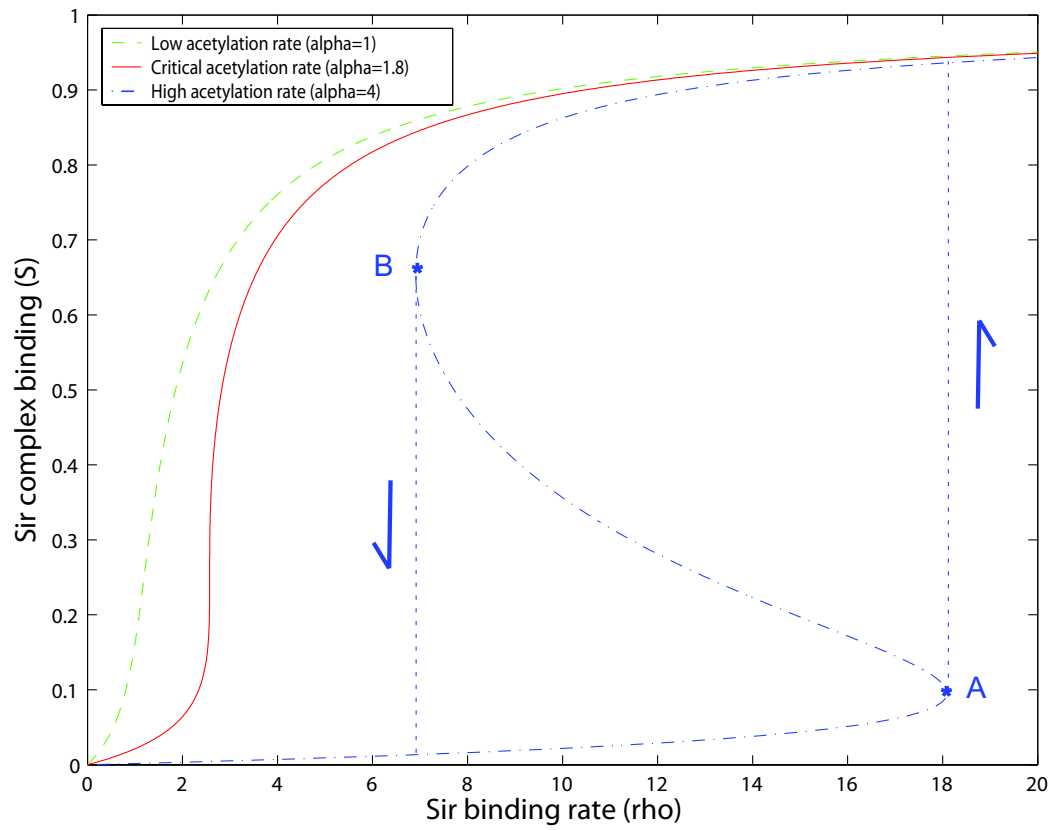


Figure 2.5: Sir occupancy, S as a function of Sir availability, ρ . The S shaped curve indicates multiple solutions as is common with a bifurcation diagram in phase space, the hysteresis behavior.

Now plugging this value into equation (2.5), we will have the following formula for $\bar{\rho}$ in terms of S and $\bar{\alpha}$:

$$\bar{\rho} = \frac{S[\bar{\alpha} + 1 + (\bar{\gamma} - \bar{\alpha})S]^n}{(1 - S)(1 + \bar{\gamma}S)^n} \quad (2.8)$$

Using equation 2.8, we will make a graph of S versus $\bar{\rho}$ at different values of $\bar{\alpha}$ (Fig.2.5). For small values of $\bar{\alpha}$ the graph is monotonically increasing and therefore there is a one to one correspondence between S and $\bar{\rho}$ values, dictating that the transition from un-silenced euchromatin state (low values of S) to silenced heterochromatin state (high values of S) happens continuously along the monotonic curve (Fig.2.5, solid/red and dashed/green curves). However for large values of $\bar{\alpha}$ the graph looks like a sharp "*S-curve*" where at some values of $\bar{\rho}$ there are two corresponding values of high and low S i.e. a bi-stable regime (Fig.2.5, dash-dotted/blue curve). If one increases the value of $\bar{\rho}$ from zero and follows S on this curve, S increases continuously until point (A), where the curve does not allow further continuous increase. Therefore, at this point the value of S jumps to its value at the higher section of the curve (sharp transition from un-silenced to silenced state). For further increase in $\bar{\rho}$, there will be again a one to one correspondence between $\bar{\alpha}$ and S and the system continuously follows the curve. Now if we reverse the direction of the change and start decreasing $\bar{\rho}$, S continuously decreases until point (B), where it jumps to the lower portion of the curve and we get a sharp transition from silenced heterochromatin back to un-silenced euchromatin.

2.4 The Bifurcation Diagram

The interesting behavior of system, motivates the depiction of a phase diagram, indicating regions in the parameter space of $\bar{\rho}$ and $\bar{\alpha}$ leading to mono-stability and to bi-stability. In order to do so, we seek the boundary between the three

fixed-point (bi-stable) region and single fixed-point (mono-stable) region inside the $\bar{\rho}$ and $\bar{\alpha}$ space. We have already learned from previous discussions, that when system moves from bi-stable region to mono-stable region, at this limit two nullcline curves (2.5) and (2.6) in Fig.2.1 are tangent to each other. We use this condition to find the boundary between the two regions in phase space. The values of S and A at fixed points are given by the solution of (2.5) and (2.6), so we take the derivative of both equations with respect to A :

$$n\bar{\rho}(1-S)(1-A)^{n-1} = -\frac{dS}{dA}(1+\bar{\rho}(1-A)^n) \quad (2.9)$$

$$\bar{\alpha}(1-S) + (1+\bar{\gamma}S) = -\frac{dS}{dA}(\bar{\gamma}A + \bar{\alpha}(1-A)) \quad (2.10)$$

We should then equate dS/dA in above equations. By using equations (2.5) and (2.6), we can finally solve A in terms of S and get the simplified equation below:

$$A(S) = \frac{1 + \bar{\gamma}S}{nS(1 + \bar{\gamma})} \quad (2.11)$$

Combining the above equation with equations (2.5) and (2.6) and writing $\bar{\alpha}$ and $\bar{\rho}$ in terms of parameter S ; we have:

$$\bar{\alpha}(S) = \frac{(1 + \bar{\gamma}S)^2}{(1-S)[(\bar{\gamma}+1)(n-1)S - (1-S)]} \quad (2.12)$$

$$\bar{\rho}(S) = \frac{S}{(1-S)} \left(\frac{n(\bar{\gamma}+1)S}{(\bar{\gamma}+1)(n-1)S - (1-S)} \right)^n \quad (2.13)$$

Finally, using above equations, we will be able to graph the pair of parameters $(\bar{\rho}, \bar{\alpha})$ at the boundary between bi-stable and mono-stable region (Fig.2.6). The diagram shows that the bi-stable region in parameter space is a wedge-like region with a pointed *critical point*. Note that, in our graph the horizontal axis is $\bar{\rho}$ and the vertical axis is $\bar{\alpha}$. The mono-stable silenced (heterochromatin) state exists

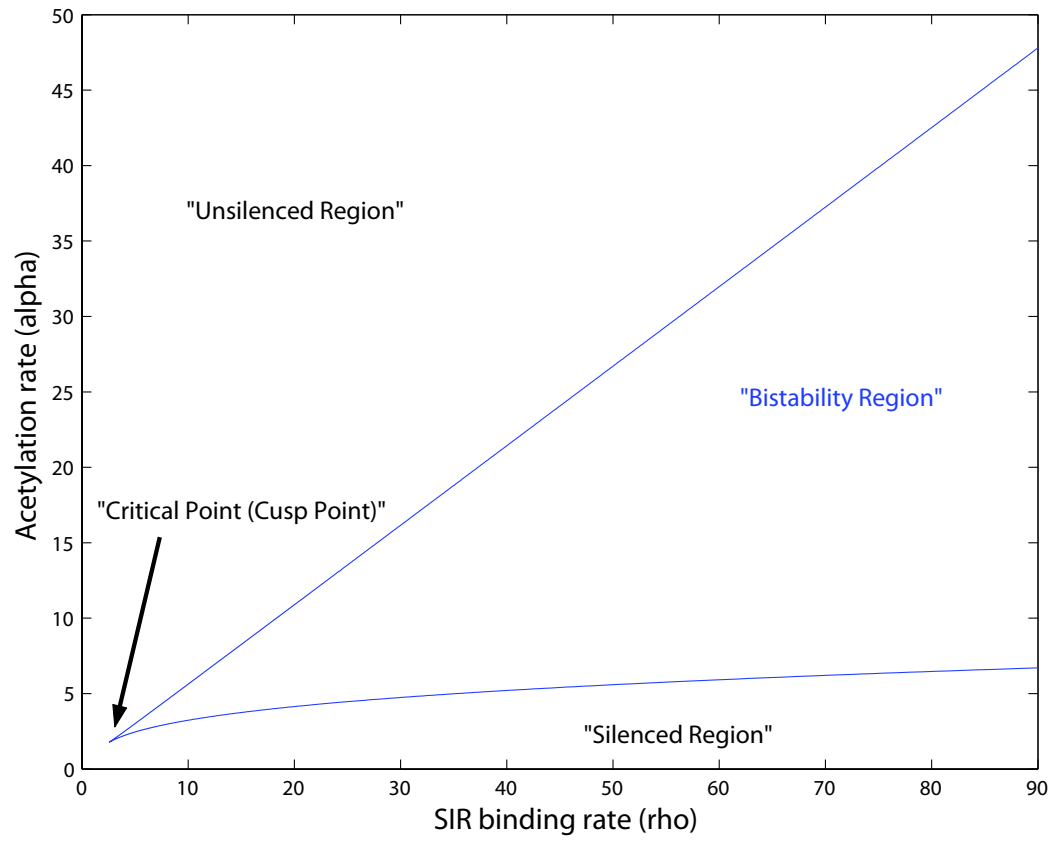


Figure 2.6: The bifurcation diagram in control parameters space, $(\bar{\alpha}, \bar{\rho})$. The degree of cooperativity n was set to 4 and $\bar{\gamma} = 4$.

below the bi-stable region, whereas mono-stable un-silenced (euchromatin) state exists above this region. Because of its special shape, we also call this diagram the *bifurcation* diagram.

2.4.1 The Critical Point

The critical point (the cusp point in bifurcation diagram) corresponds to the minimum values of $\bar{\rho} = \bar{\rho}_c$ and $\bar{\alpha} = \bar{\alpha}_c$ in bi-stable region. In other words, if either $\bar{\rho} < \bar{\rho}_c$ or $\bar{\alpha} < \bar{\alpha}_c$ there will be impossible to have two opposite un-silenced and silenced states existing at the same time. From equations (2.12) and (2.13), we have $\bar{\rho}(S)$ and $\bar{\alpha}(S)$ both as functions of S , then taking their derivatives with respect to S and equating them to zero will give us the value of S at the critical point, S_c :

$$S_c = \frac{n+1}{2n + \bar{\gamma}(n-1)} \quad (2.14)$$

Note that both $\frac{d\bar{\rho}}{dS}(S_c) = 0$ and $\frac{d\bar{\alpha}}{dS}(S_c) = 0$ give us the same value for S_c , as they should. If we call the corresponding value of $A(S_c)$, A_c ; from equation (2.11) we will get:

$$A_c = \frac{2}{n+1} \quad (2.15)$$

Using these values in equations (2.5) and (2.6) will result in critical values of $\bar{\rho}$ and $\bar{\alpha}$:

$$\bar{\rho}_c = \frac{(n+1)^{n+1}}{(n-1)^{n+1}(1+\bar{\gamma})} \quad (2.16)$$

$$\bar{\alpha}_c = \frac{4n}{(n-1)^2} \quad (2.17)$$

Note that, as it can easily be seen from above equations, $n \neq 1$ is a necessary condition for the existence of the critical point. It is also worth to restate that

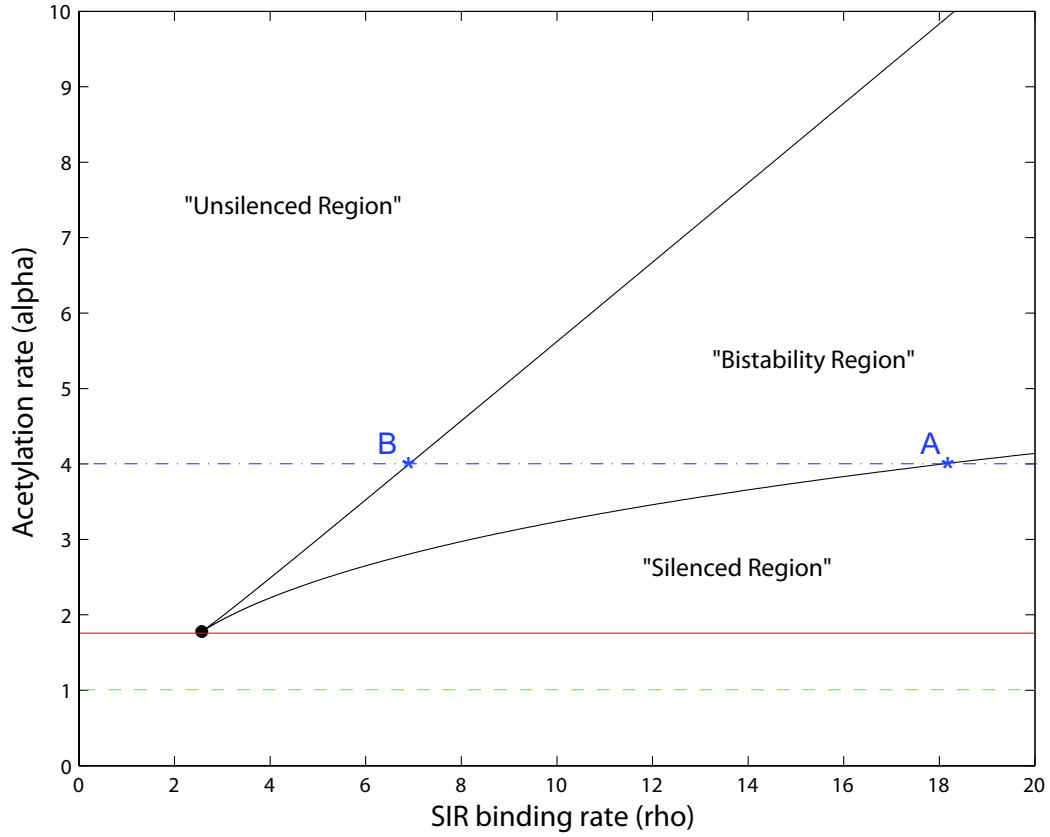


Figure 2.7: The corresponding behavior of system in bifurcation diagram for change of Sir occupancy, S as a function of Sir availability, ρ in Fig.2.5.

the critical point is at low availability silencing factors, $\bar{\rho}$ coupled with low rate of acetylation, $\bar{\alpha}$. This feature will have an important implication when we will later consider the result of experiments involving mutants lacking particular acetylases (discussed in section 1.3.2).

2.4.2 The Hysteresis Behavior in Phase Space

We can now return to the previous discussion on the behavior of the system at different values of $\bar{\rho}$ and $\bar{\alpha}$. Particularly, let us look at the plot of S versus $\bar{\rho}$ for three constant values of $\bar{\alpha}$ and the hysteresis behavior (Fig.2.5) one more time. The corresponding values of $\bar{\alpha}$ will give us three horizontal lines on the phase

diagram, (Fig.2.7). For low value of $\bar{\alpha}$ the horizontal line (dashed/green line) is under the critical point and as the value of $\bar{\rho}$ increases the system smoothly goes from un-silenced state to silenced state. However, for large values of $\bar{\alpha}$ the horizontal line (dash-dotted/blue line) is above the critical point and to go from un-silenced to silenced region the system has to pass the bi-stability region. If the system enters the bi-stability region from left, i.e. un-silenced region, it keeps its un-silenced state until value of $\bar{\rho}$ increases to point (A) at the boundary. Further increase in $\bar{\rho}$ means entering the silenced region and therefore forcing the system to sharply transform from un-silenced to silenced state. On the other hand, if we enter bi-stability region from right, the system keeps its silenced state until $\bar{\rho}$ decreases to the value at point (B) where we will again have a sharp transition of the system's state, this time from silenced to un-silenced. Finally, the middle horizontal line passes through the critical point and therefore it is at the boundary between two behaviors (The solid/red line in Fig.2.5 and 2.7) .

2.4.3 The Role of Non-local Interaction Factor, γ and Degree of Cooperativity, n

One might ask how the bifurcation diagram discussed above depends on parameter γ in equation (2.4). γ is the parameter that controls the average rate of de-acetylation by Sir complex. In other words, the larger the γ the more chance of de-acetylation, hence more silencing. Therefore, we expect that when γ is large, only large values of α give rise to hyper-acetylated/un-silenced states. Moreover, even moderate values of ρ can cause silencing in this regime. So as it is depicted in Fig. 2.8, the un-silenced euchromatin region of the parameter space will be squeezed and the critical point shifts both towards left for higher γ values. As a result, the bifurcation diagram becomes wider.

The same question rises for, n , the degree of cooperation of de-acetylation in

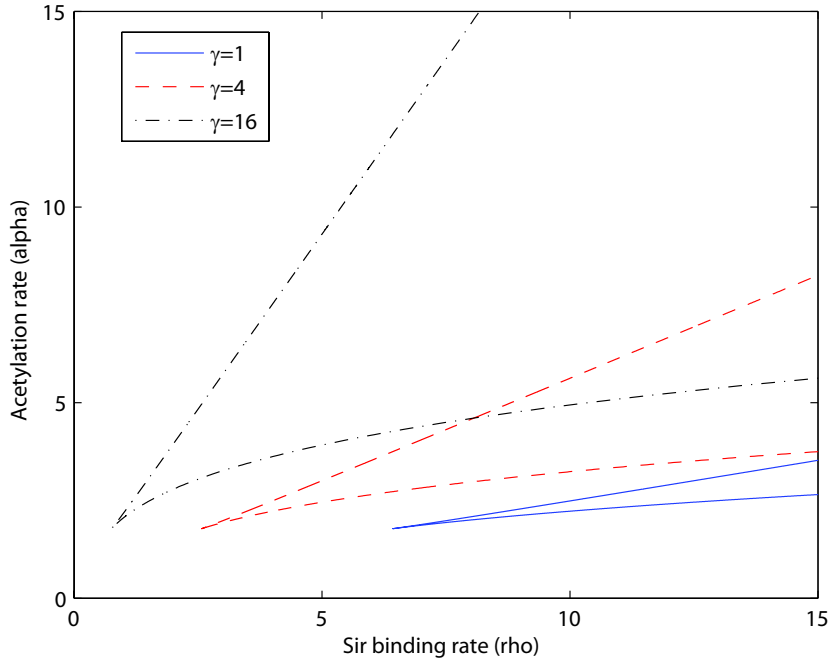


Figure 2.8: The effect of γ on the shape of bifurcation diagram.

recruiting silencing proteins, equation (2.3). As the parameter n increases, the role of cooperation of de-acetylation becomes more accentuated i.e. to get the same silencing effect the system requires more cooperating de-acetylated DNA sites or higher probability of de-acetylation (lower degrees of acetylation) at the corresponding sites. In other words, it will be more difficult to get silenced heterochromatin states than before. Consequently, the un-silenced region in parameter space prevails, squeezing the bifurcation diagram down (Fig. 2.9). The role of non-linear cooperation in dynamical equations of the system will be discussed in more details in chapter 5.

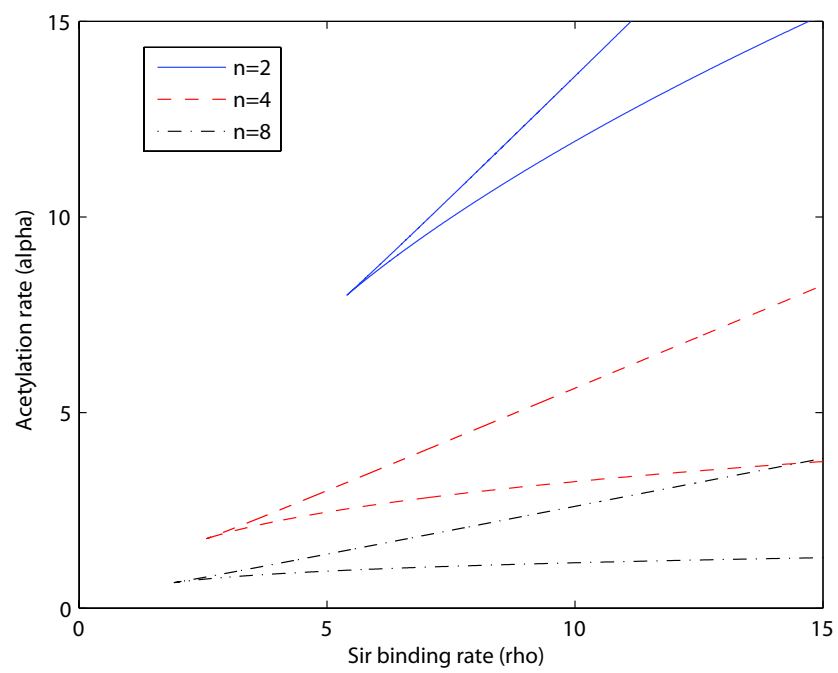


Figure 2.9: The effect of n on the shape of bifurcation diagram.

Chapter 3

Non-uniform Solutions and Front Propagation

In this chapter, we go beyond analyzing the uniform solutions and consider stable non-uniform spatial solutions. In the region of parameter space where the system is bi-stable, it is possible to study how fronts between a silenced state and an un-silenced state form and move on the lattice. In a system with a well defined free energy function, the average motion of a front or interface is determined by the difference of free energies of the two states across the front (Fig. 3.1). As it is shown in Fig. (3.1.A.), the lower free energy state S_2 (usually called the stable state) invades into the meta-stable state with higher free energy S_1 , i.e. as the front moves to the left the sites on the left of the boundary transform from S_1 to S_2 . The opposite happens in Fig. (3.1.C.) when free energy of S_1 is lower and causes the front to move in reverse direction. At the points where the two free energies are the same, the average front velocity is zero (Fig. 3.1.B.). Although, in non-equilibrium systems like the one at hand, there is no useful free energy to be defined, one might still explore the region of parameter space where silenced state invades into the un-silenced ones and vice versa (and the boundary in between where the front becomes stationary).

We study the motion of boundary between the two stable phases in the bi-stable parameter region both in the current discrete model and the continuum version of the model where the lattice is replaced by a continuous 1-D system.

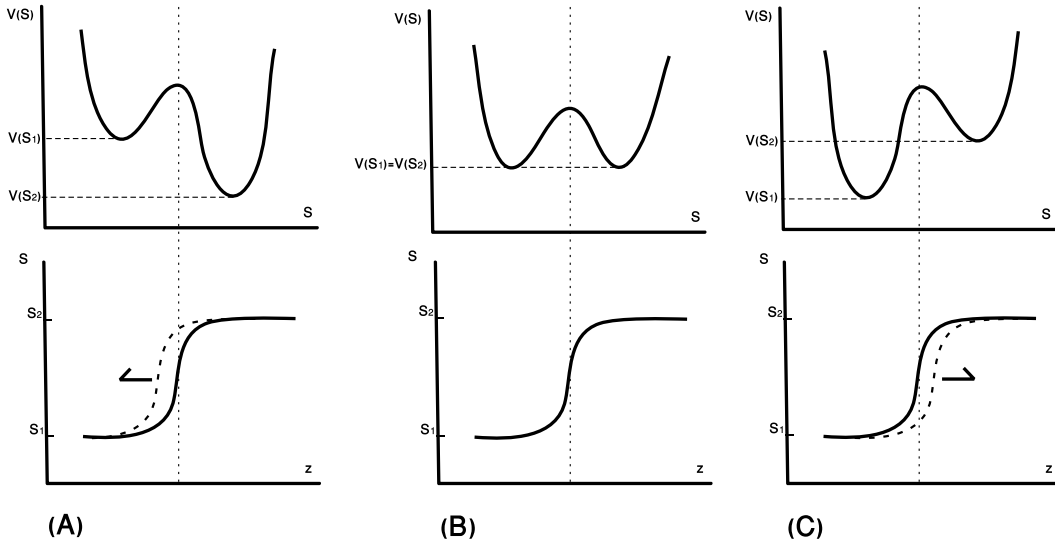


Figure 3.1: Front propagation in a system with a well defined free energy

3.1 The Continuum Limit

The continuum versions of equations (2.1) and (2.2) are given by discrete parameter i and j replaced by continuous parameters x and y , and the sum replaced by an integral; as follows:

$$\frac{\partial S(x, t)}{\partial t} = \rho(1 - S(x, t))f(1 - A(x, t)) - \eta S(x, t) \quad (3.1)$$

$$\frac{\partial A(x, t)}{\partial t} = \alpha(1 - A(x, t))(1 - S(x, t)) - (\lambda + \int \gamma(x - y)S(y, t)dy)A(x, t) \quad (3.2)$$

Similar to the discrete case, function $\gamma(x - y)$ is an *even* function that sharply falls to zero as $|x - y|$ gets large compared to a length scale ξ , say, making the integrand negligible in this region $|x - y| \gg \xi$. Since we are only interested in qualitative nature of the bifurcation diagram, we can make a further approximation restricted to the situation where $S(x, t)$ changes slowly over the distance of the order of ξ . Thus we Taylor expand $S(y, t) = S(x + (y - x), t) = S(x + u, t)$

around x , keep up to the second order in $|u| = |y - x|$ and disregard higher orders. Equation 3.2 reduces to:

$$\frac{\partial A(x, t)}{\partial t} = \alpha(1 - A(x, t))(1 - S(x, t)) - \left(\lambda + \Gamma_0 S(x, t) + \Gamma_2 \frac{\partial^2 S(x, t)}{\partial x^2} \right) A(x, t) \quad (3.3)$$

where $\Gamma_0 = \int \gamma(u) du$ and $\Gamma_2 = 1/2 \int \gamma(u) u^2 du$. Note that we set $\Gamma_1 = \int \gamma(u) u du$ equal to zero since $\gamma(u)$ is an even function. Equations (3.1) and (3.3) are then the continuous forms of set of principal equations (2.1) and (2.2).

3.1.1 Wave Solutions and Zero Velocity Line

The analysis of the continuum system follows the standard route [17, 18, 19]. Seeking for front moving solutions, we assume wave solutions $A(x, t) = A(x - ct)$ and $S(x, t) = S(x - ct)$, and if $z = x - ct$:

$$0 = c \frac{dS(z)}{dz} + \rho(1 - S(z))f(1 - A(z)) - \eta S(z) \quad (3.4)$$

$$0 = c \frac{dA(z)}{dz} + \alpha(1 - A(z))(1 - S(z)) - \left(\lambda + \Gamma_0 S(z) + \Gamma_2 \frac{d^2 S(z)}{dz^2} \right) A(z) \quad (3.5)$$

For each set of parameters, there is a front velocity, c , for which there is only “one” (or none) continuous solution that represents a transition between the stationary states representing un-silenced euchromatin and silenced heterochromatin.

For the purpose of this dissertation, we would focus on the part of the parameter space where the front velocity is zero. This analysis of problem for $c = 0$ is considerably simpler, above equations will become:

$$0 = \bar{\rho}(1 - S(z))f(1 - A(z)) - S(z) \quad (3.6)$$

$$\bar{\Gamma}_2 \frac{d^2 S(z)}{dz^2} = \bar{\alpha} \frac{(1 - A(z))(1 - S(z))}{A(z)} - 1 - \bar{\Gamma}_0 S(z) \quad (3.7)$$

where $\bar{\Gamma}_0 = \Gamma_0/\lambda$ and $\bar{\Gamma}_2 = \Gamma_2/\lambda$. Now, since equation (3.6) is an algebraic equation, allowing A to be expressed in terms of S , we can define a potential $V(S)$ as follows:

$$V(S) = -S - 1/2 \bar{\Gamma}_0 S^2 + \bar{\alpha} \int_0^S dS \frac{(1 - A(S))(1 - S)}{A(S)} \quad (3.8)$$

So that other equation, namely equation (3.7), could be written as

$$\bar{\Gamma}_2 \frac{d^2 S}{dz^2} = \frac{dV(S)}{dS} \quad (3.9)$$

The values of parameters, for which the potential $V(S)$ has two local minima with equal potential values, correspond to existence of a zero velocity front (Fig. 3.1.B.). Note that, we were able to use this potential *only* to describe zero velocity fronts, and not for the general traveling solution. This is because only at zero velocity, the system is at equilibrium and therefore free energy can be well defined. To find out the zero velocity line in parameters space, one needs to find the relationship between $\bar{\rho}$ and $\bar{\alpha}$ resulting in having two local minima with equal values in function $V(S)$. At these minima $dV/dS = 0$ and therefore $d^2 S/dz^2 = 0$ from equation (3.9). Applying this condition to equation (3.7) reduces the pair of equations (3.6) and (3.7) to nullcline equations (2.5) and (2.6) with $\bar{\gamma} \approx \bar{\Gamma}_0$. This simply means that as one would have expected the local minima in continuous equations coincide with two uniform solutions found from nullclines in chapter 2. So if S_1 and S_2 correspond to two stable uniform solutions in bi-stable region, we need to have $V(S_1) = V(S_2)$ or:

$$\bar{\alpha} \int_{S_1}^{S_2} dS \frac{(1 - A(S))(1 - S)}{A(S)} = S_2 - S_1 + 1/2 \bar{\Gamma}_0 (S_2^2 - S_1^2) \quad (3.10)$$

Since $A(S)$ is given by algebraic equation (3.6), it is a function of $\bar{\rho}$. Therefore, $\bar{\alpha}$ can be easily found as a function of $\bar{\rho}$ by above equation. For the case of

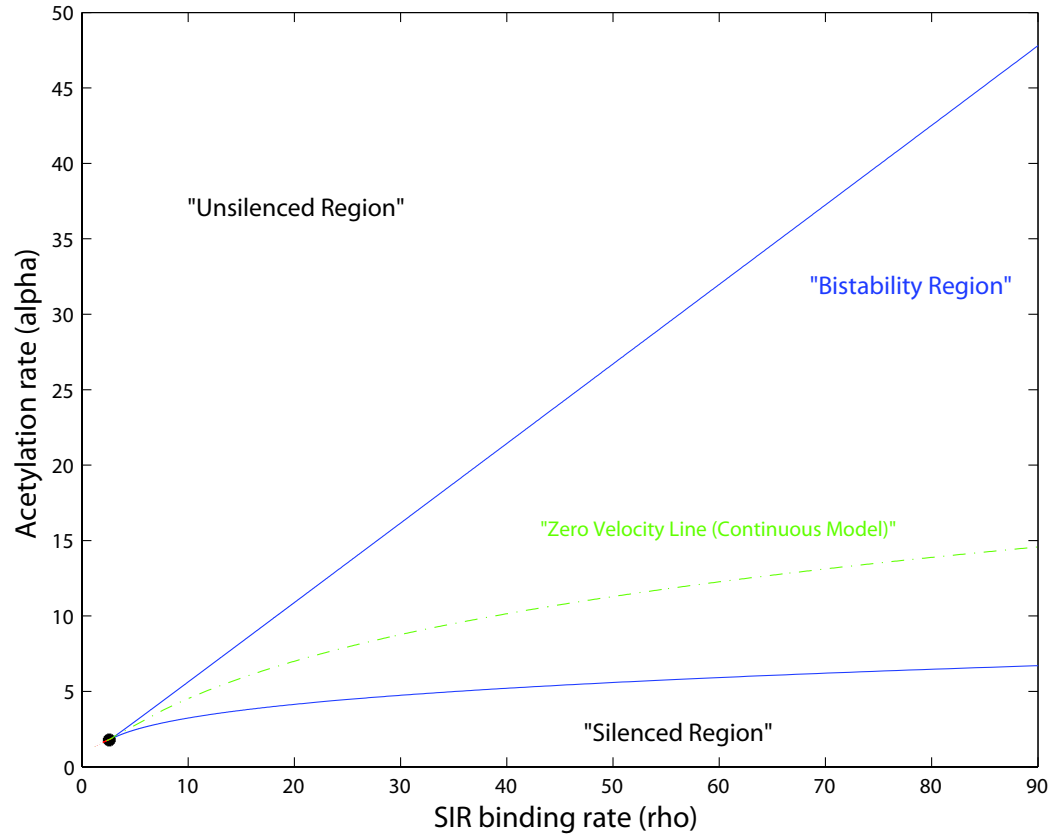


Figure 3.2: The zero velocity line, the line of stable boundary formation between euchromatin and heterochromatin states in continuum regime, inside bi-stable region of phase diagram. The degree of cooperativity n was set to 4 and $\bar{\gamma} = 4$.

$f(x) = x^n$, we solved the integral numerically using *MATLAB* and plot zero velocity line inside the bi-stability region in phase diagram (represented by the dashed/green line in Fig.3.2).

According to the diagram, zero velocity line starts from the critical point and divides the bi-stable region into two sections. The upper section, which is next to mono-stable un-silenced regime corresponds to invasion of un-silenced state into the silenced areas. The opposite happens in the lower section. Therefore, each section of the bi-stable region just mentioned, points to a different direction of the movement of front, similar to what happens in Fig.3.1.A and 3.1.C.

3.2 The Discrete Model

We also need to study the formation of boundary in the discrete model i.e. solve equations (2.1) and (2.2) directly, using numerical methods. For this means, an explicit finite difference method was employed (a more detailed description of the method is given in appendix B).

By scanning different values of parameters $\bar{\rho}$ and $\bar{\alpha}$, inside the bi-stability region one finds a band-like region (as opposed to the zero velocity *line* in continuum case) where a stable zero velocity boundary between the silenced and un-silenced states can form (in the rest of the phase space either the silenced region shrinks to zero or it covers all the sites). This result actually has been very well studied for front propagation failure in lattice models as in [20]. The boundaries of the band are represented by the dashed/red lines in Fig.3.3. This band shrinks to the zero velocity line as one takes the continuum limit. We will call this region in phase space the *zero velocity band*.

Now, One might ask which of these descriptions, the continuous or the discrete case, are closer to reality. If we count each nucleosome as a unit and expect one silencing complex per nucleosome, then that provides us with a natural lattice

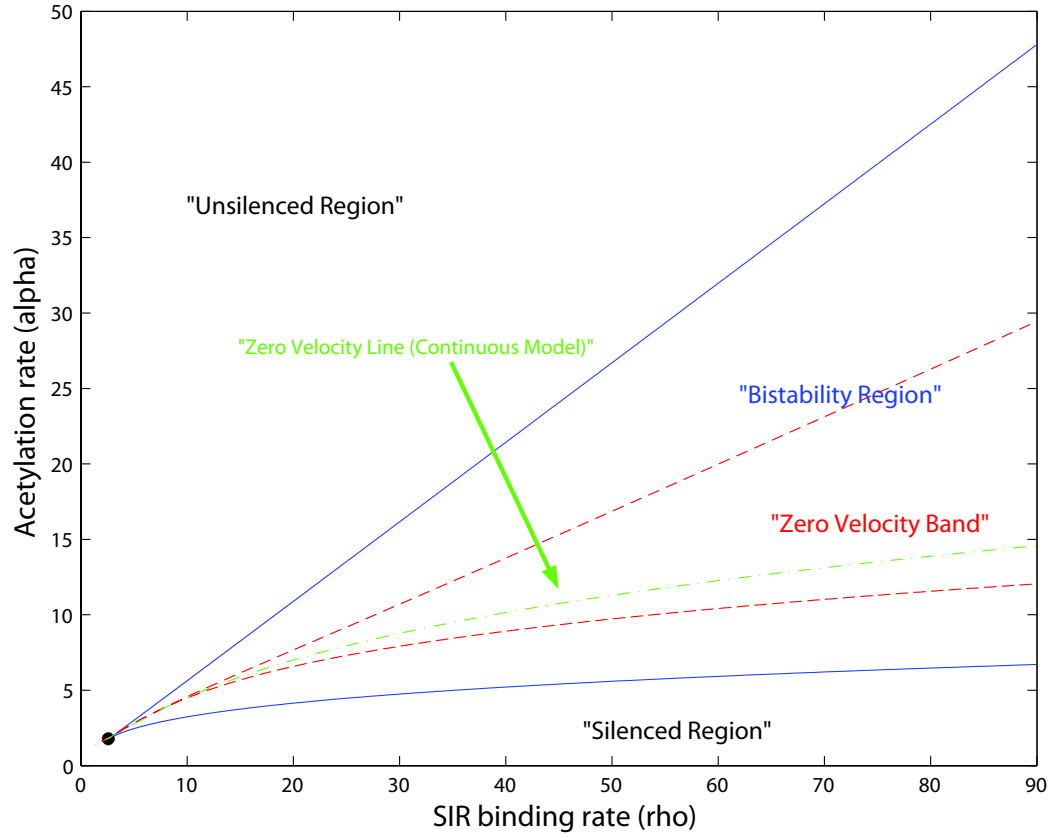


Figure 3.3: The zero velocity band, the region of stable boundary formation between euchromatin and heterochromatin states in discrete regime, inside bi-stable region of phase diagram. The degree of cooperativity n was set to 4 and $\bar{\gamma} = 4$.

spacing. However, the nucleosomes are not quite static. They could move around or disappear (if the histone-octamer falls off DNA). If the time scale of nucleosome dynamics is much slower than that of the silencing process, then we are justified in taking the nucleosome array as a lattice to operate upon. If the time scales are the other way around, we might average out the nucleosome fluctuations and get an effective continuum description. The truth probably is somewhere in between, leading to a fuzzy region of low front mobility crossing over to high front mobility regions, above and below zero velocity line in the bifurcation diagram.

3.3 Front Dynamics

As it was discussed in section 1.3.2 lowering acetylation rate, α as in the *sas2* mutant case [25, 26], may result in the counter-intuitive behavior of reducing silencing rather than helping it. This is a question of dynamics, so to find an answer we need to study the movement of the front by starting from some initial configuration of silenced and un-silenced areas at equilibrium, then change the chemical parameters, and let the system evolve until it reaches equilibrium again. Since in the bi-stability parameter region there is no unique stable configuration, there is no trivial answer to above problem if the final parameters sit in this region.

In particular, we are interested to see if the initial nucleation necessarily results in spreading of the silencing. Suppose we are in the bi-stable region of the parameter space and start from the locally stable uniform un-silenced solution, i.e. all sites of our 1-D lattice are hyper-acetylated/un-silenced. We want to know what happens if we nucleate a small region of silencing, say, by tethering a protein that recruits the silencing factors locally. So we force nucleation at a site in the middle of lattice, i.e. we keep a site silenced all the time. With initial parameters inside the zero velocity band and by utilizing numerical methods

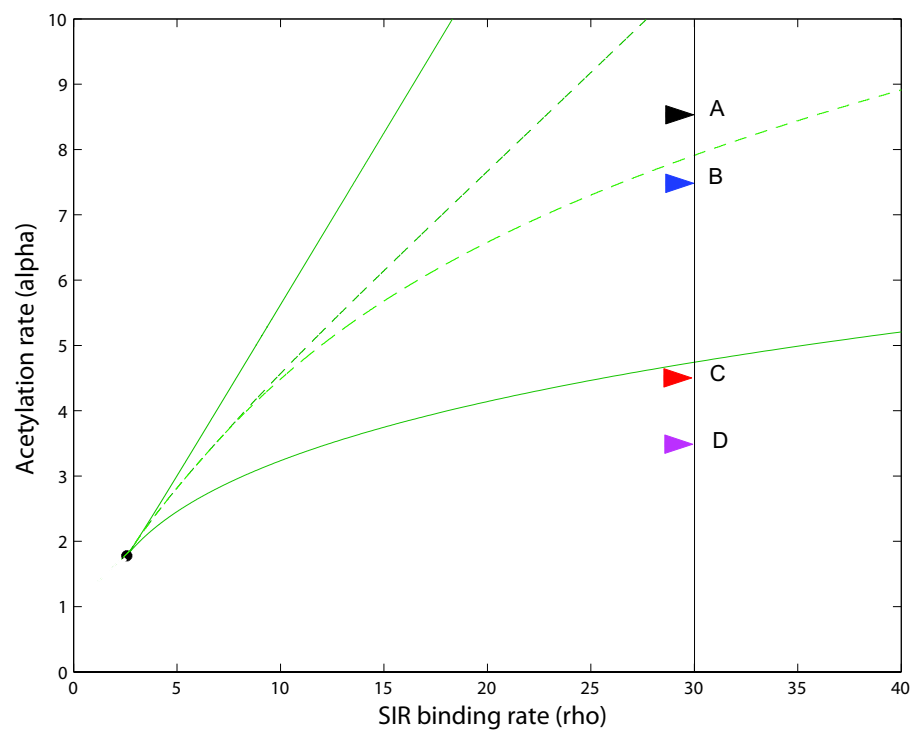
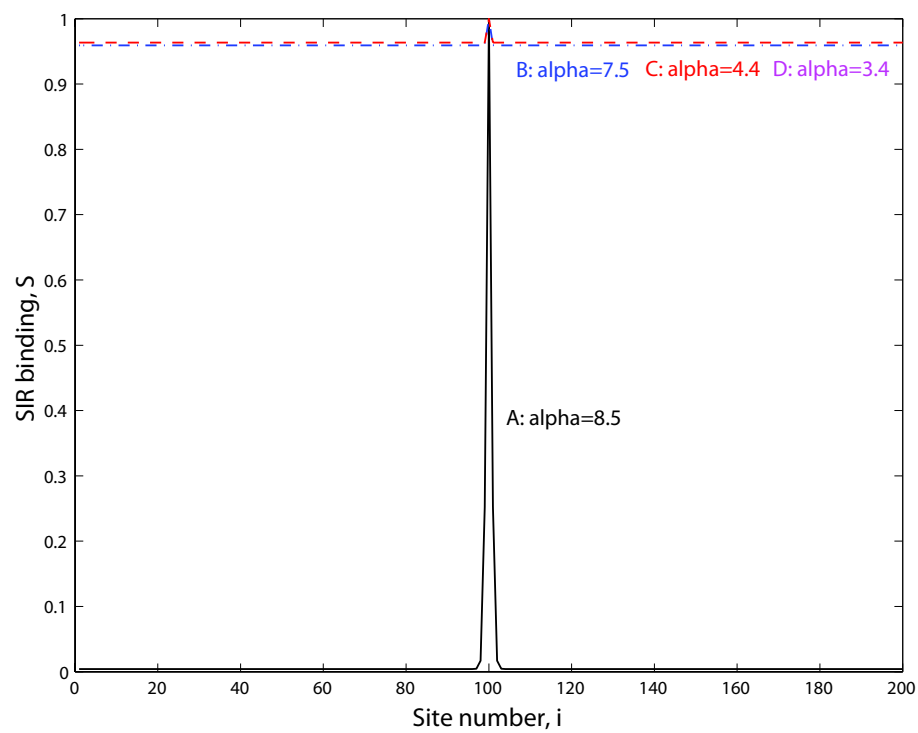


Figure 3.4: Front dynamics at infinite supply of Sir proteins, ρ constant ($\bar{\rho} = 30$).

in appendix B, we let the system evolve until it reaches the equilibrium. Not surprisingly, there will be no spreading and the final configuration will not be very different. The spreading will be limited only to the neighborhood of the nucleation site. The system needs only to conform to the desired shape of the boundary to reach equilibrium and this does not take more than a few sites at vicinity (case A: solid/black curve, Fig. 3.4). Then if one decreases α , say, by knocking off an acetylase protein; nothing changes unless we move out of the zero velocity band. For a α under the zero velocity band yet inside bi-stability area, since there will be no stable boundary between two states the silencing spreads and covers all the space (case B: Dash-dotted/blue curve, Fig. 3.4). It is obvious that, the same thing would happen independent of the initial configuration, if final value of α was under the bi-stability region inside the silenced state territory (case C: Dashed/red curve and case D: dotted/magenta curve, Fig. 3.4). Note that in particular, lowering the value of α in no way reduces the silencing effect. In search of an answer to that paradoxical phenomenon, in the next section we will study the effect of finite supply of Sir proteins.

3.4 The Role of Finite Supply of Sir Proteins

So far, it was assumed that the available ambient concentrations of Sir proteins is fixed, reflected in ρ (or $\bar{\rho}$) being held constant. We could use our insights, into the bifurcation diagram, to infer what would happen if the total number of Sir proteins (the sum of those in solution and those bound to DNA) were fixed. This is particularly interesting in the bi-stable region.

Our treatment is very similar to studying phase equilibrium with a fixed number of particles. For example, consider a liquid gas mixture at a constant temperature in a fixed volume with fixed number of particles, and imagine that there is an interface between the two states. The interface moves, and the fractions

of particles in the different states change till chemical potential of the two states become equal. Under this final condition, the interface does not move anymore, apart from thermal fluctuations around the average position. As we noted before, in our problem, we may not define a free energy, but we could indeed talk about average movement of interface between two states, namely the front, and the condition under which the interface stops moving.

Going back to original equations (2.1) and (2.2), we need to replace the constant $\rho_i(t)$ by a $\rho(t)$ which is given as follows:

$$\rho(t) = \kappa(S_{total} - \sum_k S_k(t)) = \rho_0(1 - \frac{\sum_k S_k(t)}{S_{total}}) \quad (3.11)$$

where S_{total} , the total number of functional Sir complexes in the system, and ρ_0 are both constants.

Depending upon the size of the silenced region and S_{total} , one would get interesting titration effects in this model. So we actually consider two cases; case I, with S_{total} large enough to cover the whole lattice and case II, with not sufficient value of S_{total} (it can only cover a limited area of lattice). Now, suppose we employ the same experiment we began in last section and start with a nucleation center at the middle of entirely hyper-acetylated/un-silenced lattice. If we are above the zero velocity band, no stable boundary can form and high acetylation rate makes it impossible for silenced patch to spread into the un-silenced region. Even inside the zero velocity band, as with the previous case of infinite supply of Sir protein, the spreading will only be limited to the vicinity of nucleation and it stops as soon as stable boundary can form. Therefore, in both cases silencing is going to remain localized around the region of recruitment.

However if the acetylation rate, α , is tuned down further; we go into the region where the silencing can spread into un-silenced area i.e. below the zero velocity band. Naively, one would always expect more silencing as in the infinite S_{total} case. However, since ρ is no longer fixed, as silencing spreads, $\rho(t)$ reduces. At

this point one of the two following things may happen. The front could stop because $\rho(t)$ reduces enough to reach a point on bifurcation diagram where the propagation velocity is zero i.e. inside the zero velocity band. Thus the effect of reducing α would be to effectively reduce ρ as well, so that the system always stays on the zero velocity region. This result can happen in both case I and II (case B, Fig.3.5 and cases B, C & D, Fig.3.6). However, in case I there is a possibility that before reaching the zero velocity band, silencing has covered all the lattice. So in this case the final state can be any point inside the bi-stability region, or inside mono-stable silenced region and not necessarily in the zero velocity band (cases C & D, Fig.3.5).

In conclusion, Fig.3.5 and 3.6 show how the simulation results agree with the predictions above. In all the cases we started with an initial stable nucleation region inside the zero velocity band, then jumped to lower values of α (black triangles) and let the system evolve until it reaches its final stable configuration (colored triangles). As it is shown in Fig.3.6 for limited supply of Sir proteins the system always stops on the zero velocity band, silencing a limited area on the lattice. In Fig.3.5, however, S_{total} is large enough to cover the whole space; so silencing of limited patches of DNA only happens when α is not too small.

The importance of the results from this simulation is not limited to the above discussion on spreading and front propagation. Comparing results of case C and D in Fig.3.6 reveals that lowering α can actually reduce the silencing effect. This effect can be explained intuitively as follows; since with limited supply of Sir proteins and decreasing α system has to stay on the zero velocity band, the value of ρ also decreases, consuming more ambient Sir proteins. In other words, more free Sir proteins sit on the lattice and expand the silenced area. However, the decrease in both parameters α and ρ moves the system closer to the critical point, where the values of S and A for two opposite states converge. In other words, the average probability of Sir binding, S , *decreases* and the average probability

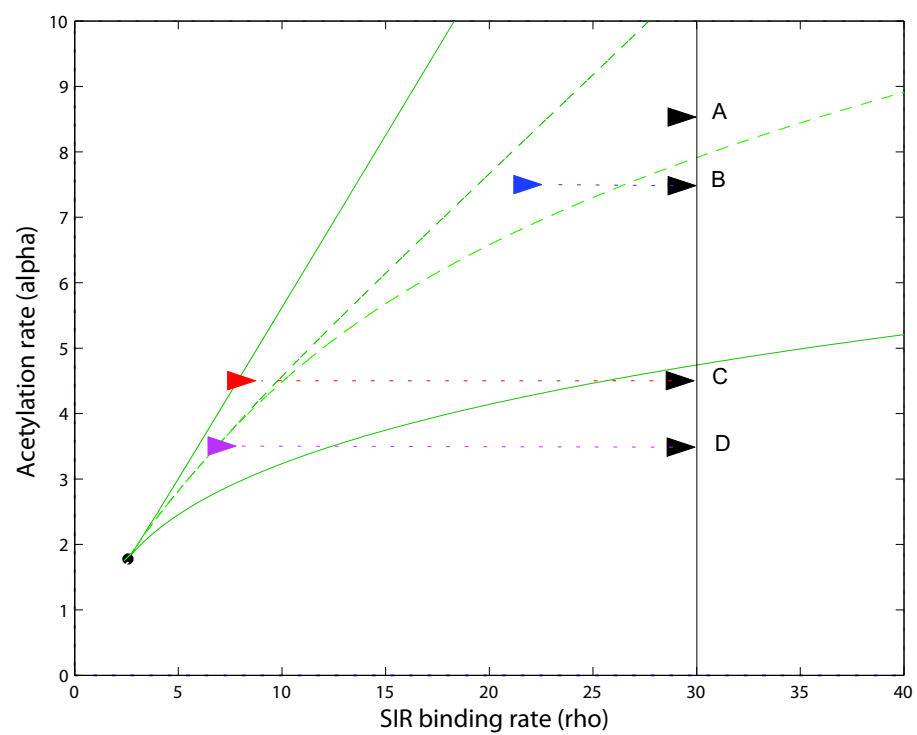
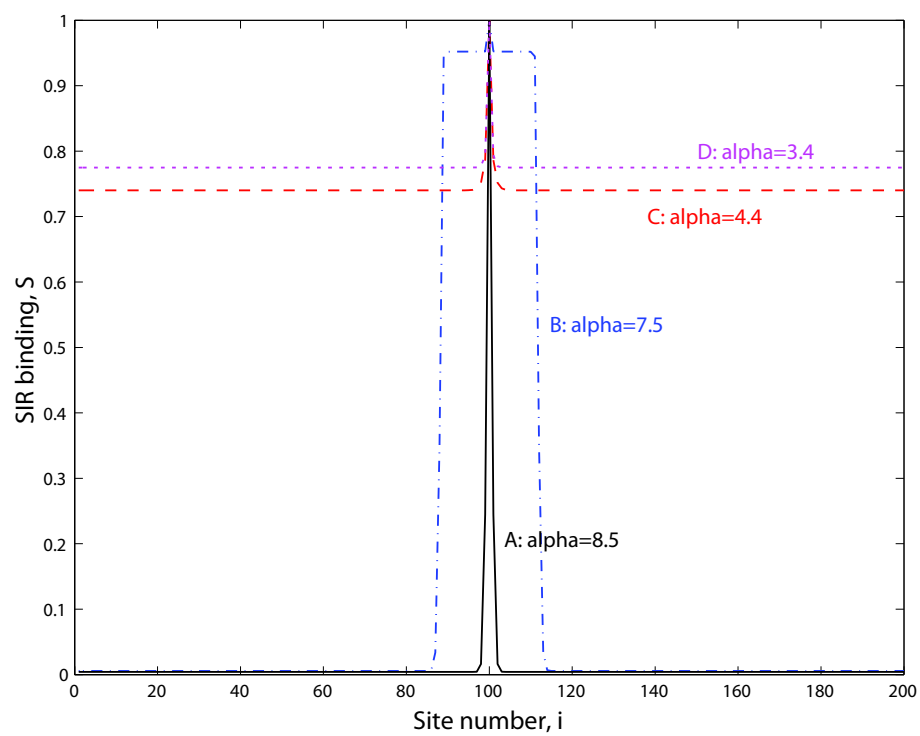


Figure 3.5: Front dynamics at limited supply of Sir proteins, enough to cover the whole lattice.

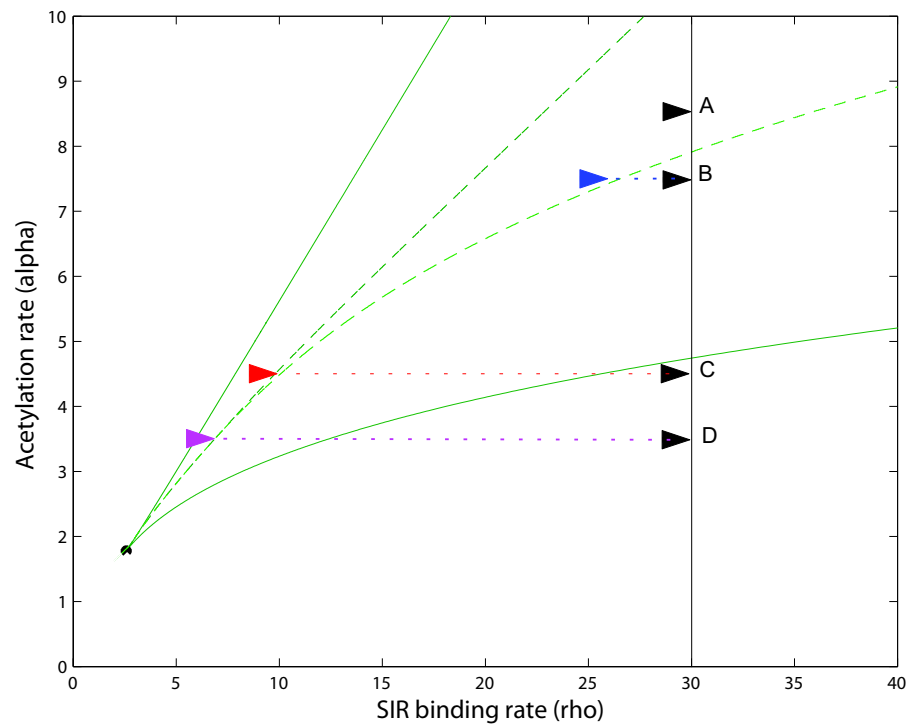
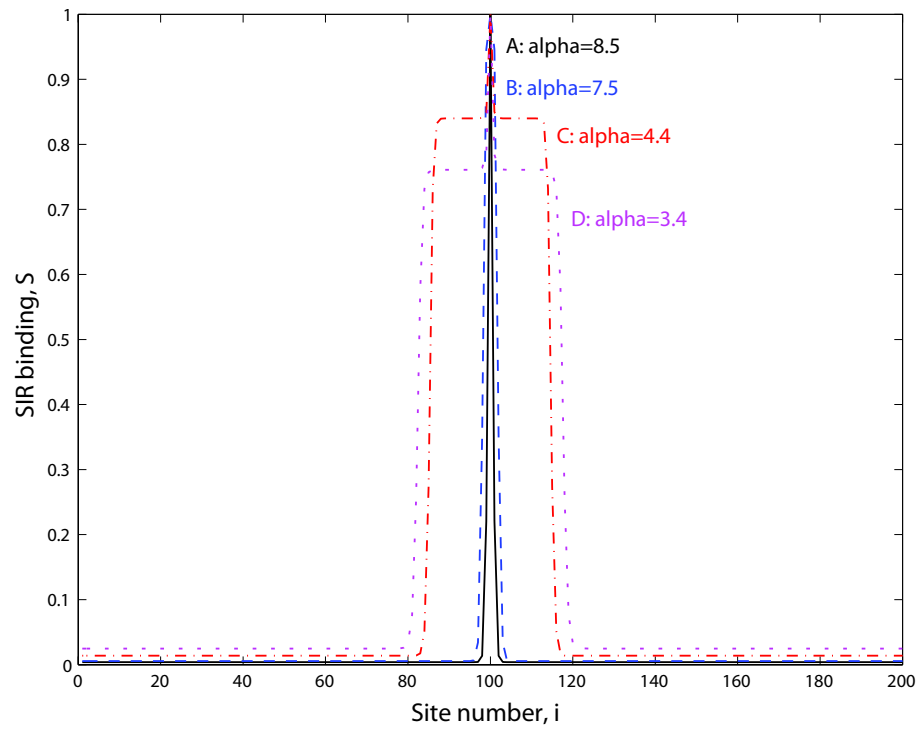


Figure 3.6: Front dynamics at limited supply of Sir proteins, enough to cover a limited patch on the lattice.

of acetylation, A , *increases* and both get closer to their critical values.

In addition, around the critical point, the bi-stability region is very narrow and by small fluctuations the system may cross bi-stable line and switch states. So lowering acetylation rate, α as in the mutants of *sas2* which code for acetylases [25, 26], not only can reduce average silencing but also move the system closer to the critical point and results in very sharp changes in silencing under the small changes in Sir availability. So we believe that the resulting system becomes extremely susceptible to cellular noise and would display a wide distribution of expression. Thus, as opposed to the naive expectation that *SAS2* deletion will just make everything more transcriptionally silent, one should find individual cells that show good expression from the "*silent*" loci.

In addition, the mono-stable intermediate silencing state in *SAS-I* elimination experiment (section 1.3.2) may correspond to the case where α has dropped to a value lower than the critical value, α_c . In this case, the system follows the same behavior and consumes ambient Sir proteins until ρ decreases to a value close to the critical point. However, since α is smaller than α_c , the final state will be in the mono-stable region of phase space at a point between un-silenced and silenced regions, where there is no distinction between two states. This point very well describes the *porous* heterochromatin configuration of this state.

Note that, as discussed in chapter 1, in reality the silencing could also stop at special boundary elements on DNA where some other process stops the spread of silencing [21, 24]. Alternatively, if there are more than one region in DNA where silencing spreads by the same mechanism and if at least one of these regions does not possess a boundary element, then we are led to the same situation, namely ρ reducing enough to stop front movement. We will explore the biological consequence of this observation in the chapter 6.

Chapter 4

The Role of Stochasticity

An important aspect of biochemical interactions inside cell is that they all show some degree of stochasticity. Numerous studies on gene expression in living cells have shown the *intrinsic* stochasticity in transcription and translation [32]. For instance, one can look at the bacteriophage λ switching phenomenon, which is a good example of noise in cellular processes as the switching between the bi-stable lysis and lysogenic states happens randomly [33]. Basically the underlying reason to this intrinsic noise is that, most of the time only a few constituent protein molecules are present to participate in interactions, so the interactions happen like sudden bursts at random times [34, 35]. Thus, in practice, cellular interactions are very susceptible to thermodynamic fluctuations. As a result, the mean-field approaches like what we took with equations (2.1) and (2.2) in previous chapters are generally far from reality and just demonstrate the average behavior of cell in these circumstances. However, as it was discussed at the end of chapter 1, studies show that *extrinsic noise* (the fluctuations in the chemical parameters themselves) is often slower and prevails intrinsic fluctuations described above [36, 37]. So one might still be able to learn a lot about the effect of extrinsic fluctuations at the mean-field approximation of the system.

In contrast, we also mentioned many observations on epigenetic inheritance that show a very robust inheritance of chromatin configuration to the decedent cells. One might ask, despite all the noise involved, how cells can behave with such robustness. Unfortunately, an exact answer to that question requires a detailed knowledge of chromatin configuration and molecular processes during cell division

which, because of the physical constraints in probing them, it has not yet been available to us.

Our model can actually provide us with some insights regarding the robustness of the average chromatin configuration to extrinsic fluctuations, sudden changes in chemical constants. We will discuss the results of our simulation to explain this average configurational robustness in the first section of this chapter. In the next section, however, we will actually consider the intrinsic noise in a stochastic treatment of our model and will see that the model can still show the same form of bifurcation diagram under thermodynamic fluctuations and thus the similar bi-stable/mono-stable behavior persist.

4.1 Extrinsic Fluctuations and Front Robustness

Going back to discrete equations (2.1) and (2.2); so far the main chemical parameters of the model α and ρ (or ρ_0 , in the case with limited supply of Sir), have been treated as constants, ignoring any fluctuations in their value. We will show that, in the mean-field approach, depending on the intensity of α and ρ variation, the system can still display the same average silencing configuration and maintain the same position of the barrier. Assume that we are in the middle of the zero velocity band and some contiguous part of DNA lattice is silenced (Solid/black curve in Fig. 4.1.A & Fig. 4.2.A). The borders between two states are very well defined and the system is at equilibrium. Now keeping ρ fixed (The supply of Sir proteins is un-limited), we shift the value of α to either of its value at the boundaries of zero velocity band and using the numerical method in appendix B, we let the system evolve (Dashed/green curve, Fig. 4.1). The result as it is shown in the picture, exhibits no shift in the position of the boundary. The only thing is the degree of silencing that slightly changes inside the silenced patch. Clearly, if α moves out of the zero velocity band, the front will move until either of the regions

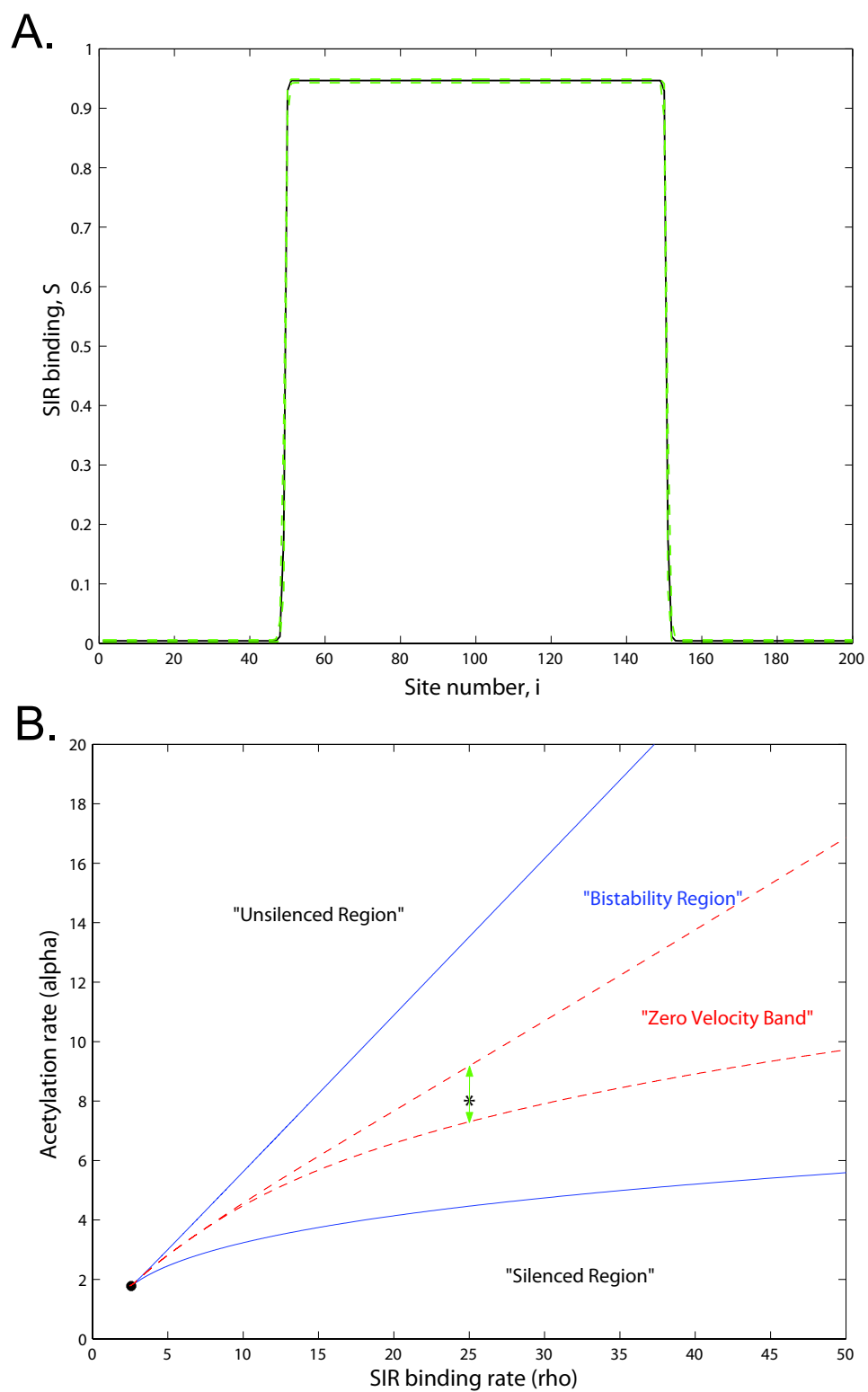


Figure 4.1: Average front robustness due to fluctuations in chemical parameter α at un-limited supply of Sir proteins

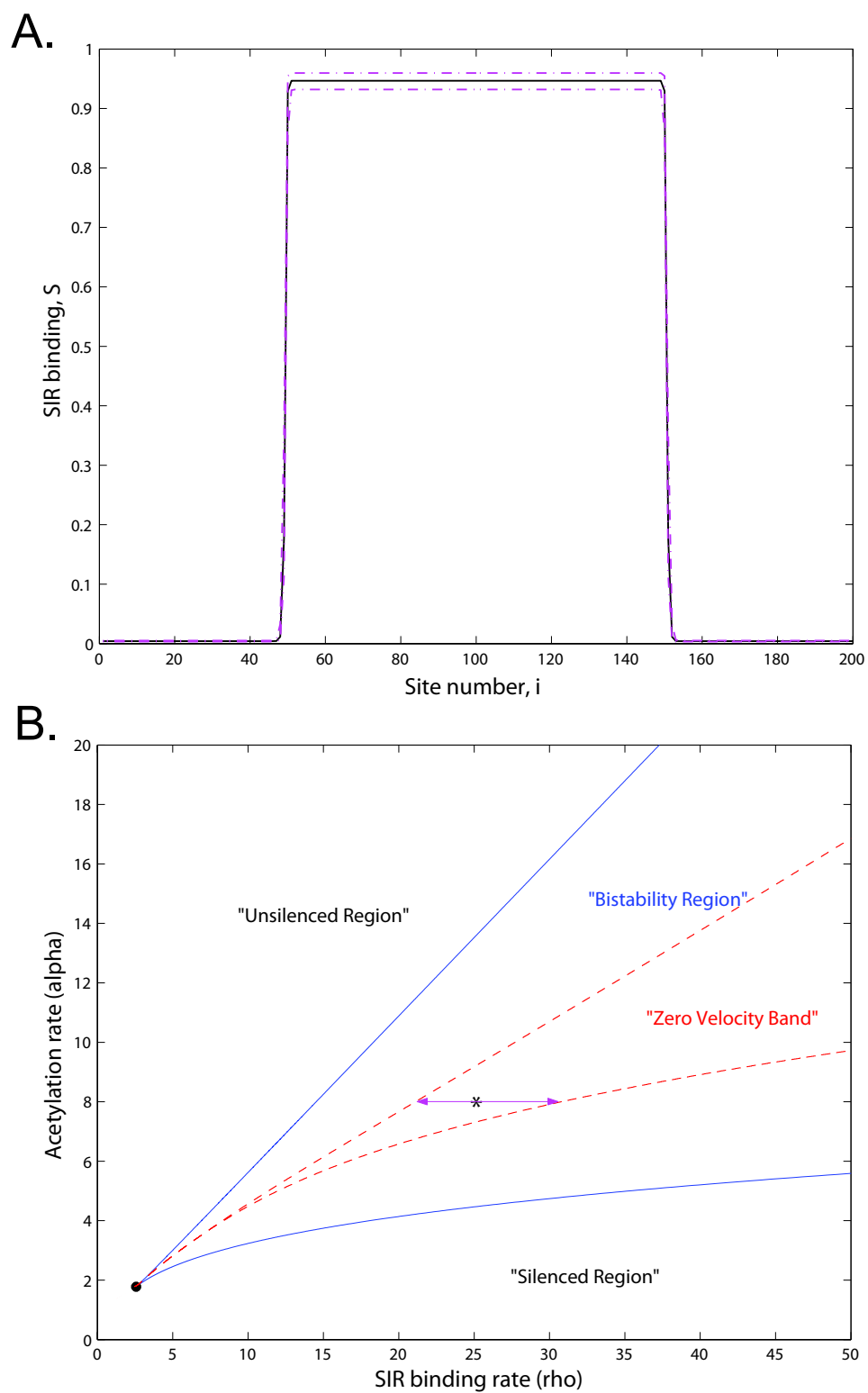


Figure 4.2: Average front robustness due to fluctuations in chemical parameter ρ at un-limited supply of Sir proteins

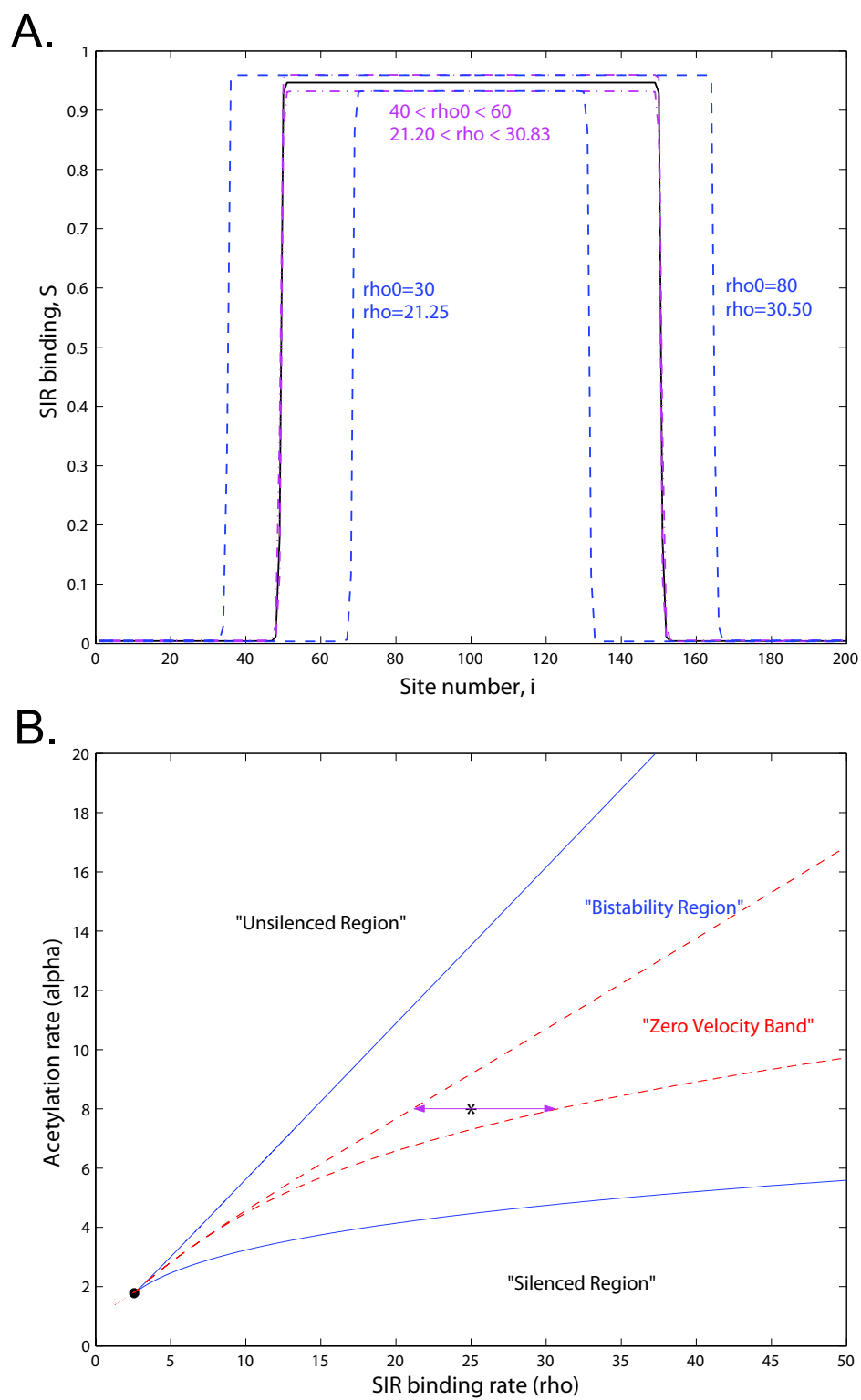


Figure 4.3: Average front robustness due to fluctuations in chemical parameter ρ_0 at limited supply of Sir proteins

shrinks to zero. Now if one does the same simulation with ρ changing instead of α , the result will still demonstrate persistence in the position of the boundary between two states (Dash-dotted/magenta curve, Fig. (4.2)). For the case with limited supply of ambient Sir proteins, one would still expect the same behavior, although now the sudden shift is assumed to be in ρ_0 (Fig. 4.3). However, in this case, when the value of ρ_0 crosses the upper/lower bounds, the average shift in the position of boundary will be *limited* since the value of ρ always returns inside the zero velocity band to maintain a stable boundary (Dashed/blue curve, upper diagram, Fig. 4.3). In other word, when the supply of Sir proteins is limited, the position of boundary is even more robust to changes in chemical parameters.

In conclusion, depending on the location of our system inside the parameter space or more precisely inside the zero velocity band, there is some degree of persistency in the average position of the boundary after it is settled. This will provide us with some insights on why chromatin which is subject to different extrinsic fluctuations during the cell life cycle can still maintain the similar configuration.

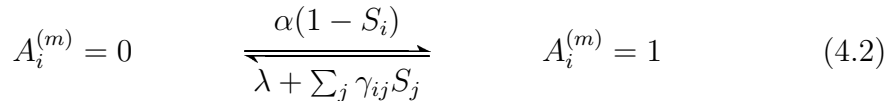
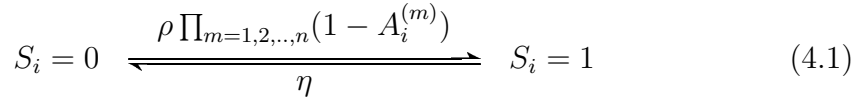
4.2 The Intrinsic Stochastic Treatment of the System

So far, we assumed a mean-field approximation of the system by introducing chemical equations (2.1) and (2.2) as the principal equations governing the silencing phenomenon. As it was discussed at the beginning of this chapter, stochastic nature of cellular interactions demands a stochastic approach to withhold a more realistic perspective of the system's behavior. The intrinsic stochastic nature of the interactions can be treated in equations (2.1) and (2.2) when the parameters A_i and S_i are not considered average values anymore and only acquire a few discrete values on a random basis, e.g. S_i can only take 1 for Sir complex sitting at site i or 0 for Sir complex not sitting there.

The mean-field treatment of the system can only be equal to the average of its stochastic version when all terms in equations (2.1) and (2.2) are linear or when there is no cooperativity. More precisely, the average value of any non-linear function f of the stochastic variable x is not in general equal to function f of the average value of stochastic variable x ; or $\langle f(x) \rangle \neq f(\langle x \rangle)$. So we need to check if the average stochastic treatment of the system actually agrees with our qualitative results from the mean-field approach or not.

One of the big drawbacks to the mean-field approach is that when one considers a stable fixed point as silenced to un-silenced, it only represents the average behavior of the system. In contrast, when one studies single cell data, there are no stable fixed points because of stochastic deviations. Therefore to be more concrete, one is required to check if the stochastic deviation from the average value is small enough that they can actually be considered a *stable point* in stochastic regime. So the stochastic simulation of the system is necessary, not just to examine the average qualitative results from mean-field approach but also to see if the fixed points in the mean-field model can represent any stable points in the real stochastic behavior of chromatin.

For the principal equations (2.1) and (2.2), with $f(x) = x^n$, we assumed a digital variable S_i and n digital variables $A_i^{(m)}$, $m = 1, 2, \dots, n$ at each site of the 1-D lattice. Therefore, we have the following rates between opposite digital states:



where all the chemical parameters are kept constant. Then we employed a Monte Carlo simulation on a lattice of length $L = 10$ with periodic boundary conditions. Two opposite initial conditions was chosen for each set of parameters:

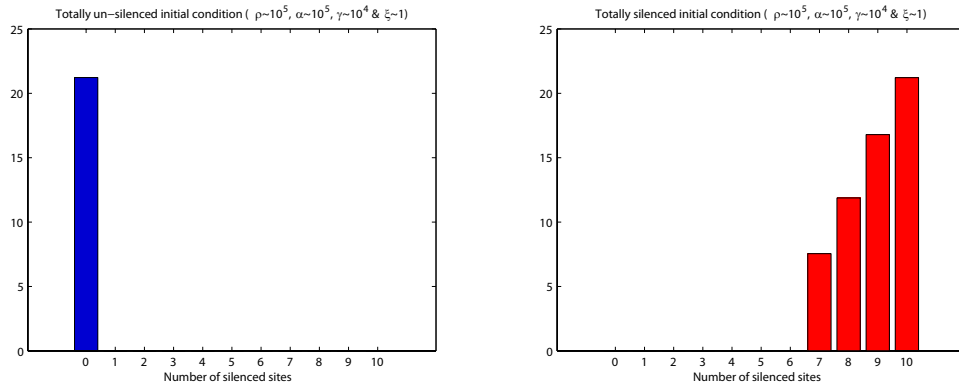


Figure 4.4: Histogram of the natural logarithm of number of silenced sites in a $L=10$ site lattice, starting from two opposite initial conditions. The corresponding values of chemical parameters are indicated at the top of each graph. Note that, the degree of cooperativity n equals 2 and total number of samplings is more than 10^9 . The diagrams exhibit the possibility of two sharp stable fixed points in the stochastic model.

an entirely hypo-acetylated/silenced lattice and an entirely hyper-acetylated/un-silenced lattice. We also assumed the parameter γ_{ij} to fall off exponentially with a length scale ξ . For a more detailed description of the method please refer to appendix C. By scanning on different values of ρ and α , we could distinguish different points of the phase space according to the average behavior of the system and the time scale of convergence towards the average state. As a matter of fact, the qualitative shape of the bifurcation diagram will be similar to the bifurcation diagram of the mean-field approach, verifying that the mean-field approach does actually provide a good qualitative description of the average behavior of the stochastic system. Moreover, results from the histograms clearly show that in some regions of phase space the stable points can be very robust to noise and the system stays in silenced and un-silenced states for a *comparatively* long time (Fig. 4.4).

However, as it is seen in Fig.4.5, at some points of the parameter space one might get intermediate solutions that do not correspond to either of the known

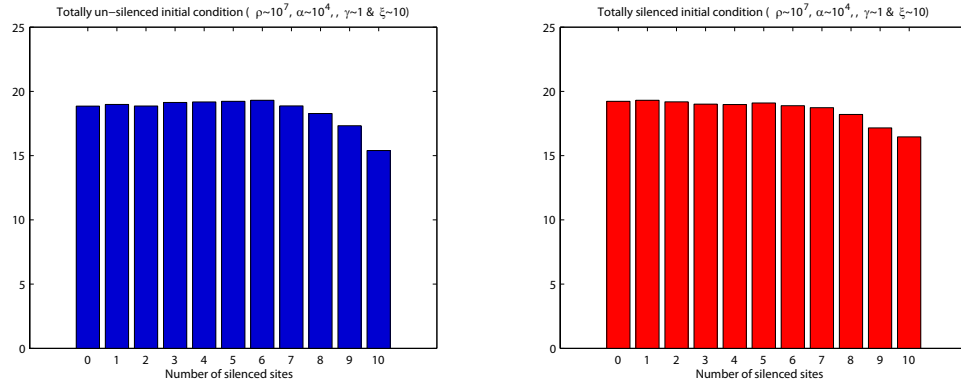


Figure 4.5: Histogram of the natural logarithm of number of silenced sites in a $L=10$ site lattice, starting from two opposite initial conditions. The corresponding values of chemical parameters are indicated at the top of each graph. Note that, the degree of cooperativity n equals 2 and total number of samplings is more than 10^9 . The diagrams exhibit the possibility of an intermediate state in stochastic regime.

stable solutions, i.e. entirely silenced or un-silenced states (Fig.4.4). These kind of solutions can represent either a uniform intermediate solution between entirely silenced and un-silenced lattice states (a point outside the bi-stable parameter regime) or they can correspond to a bi-stable system switching back and forth between two stable states (a point inside the bi-stable parameter regime). These two type of solutions have the same average behavior. However, if we keep one of the parameters constant and change the other one, we may be able to see the difference in the behavior of two cases (Fig.4.6 and Fig.4.7). As it is seen in Fig.4.7, in the mon-stable regime, convergence to a final average state could always be observed during our screening time. Where as, for the bi-stable regime (Fig.4.6), convergence to a final average state is much longer and may not happen during the screening time. One might say that, both systems may eventually converge to the same average behavior in a very long screening time. It may be true; but at least we can be sure that in time periods not too *long* the cell exhibits totally different regimes: bi-stable or mono-stable.

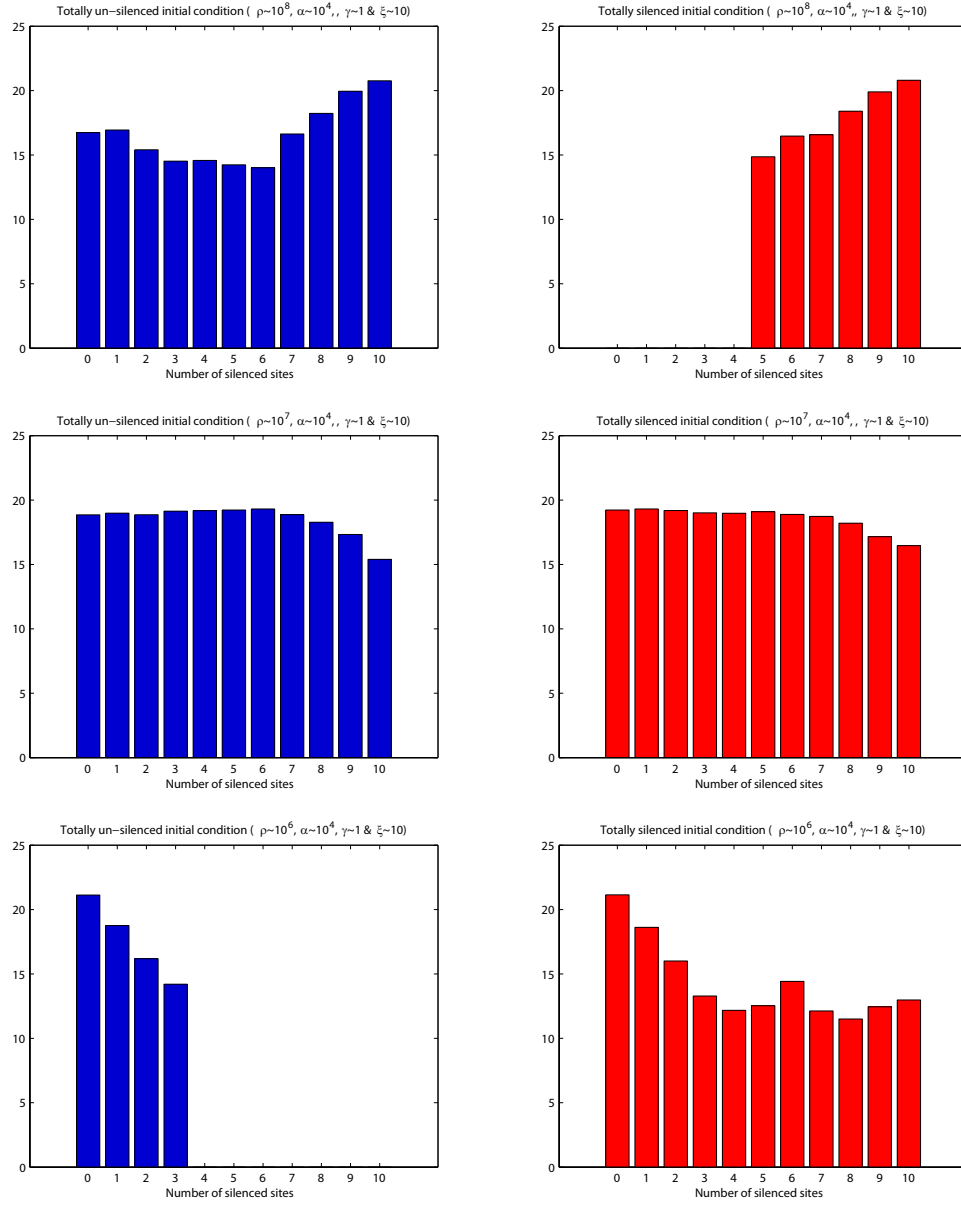


Figure 4.6: Histogram of the natural logarithm of number of silenced sites in a $L=10$ site lattice, starting from two opposite initial conditions. The corresponding values of chemical parameters are indicated at the top of each graph. Note that, the degree of cooperativity n equals 2 and total number of samplings is more than 10^9 . The transition from the average intermediate state, inside bi-stable regime of parameter space, to one of the average stable states takes a longer time comparing to a mono-stable system, Fig.4.7.

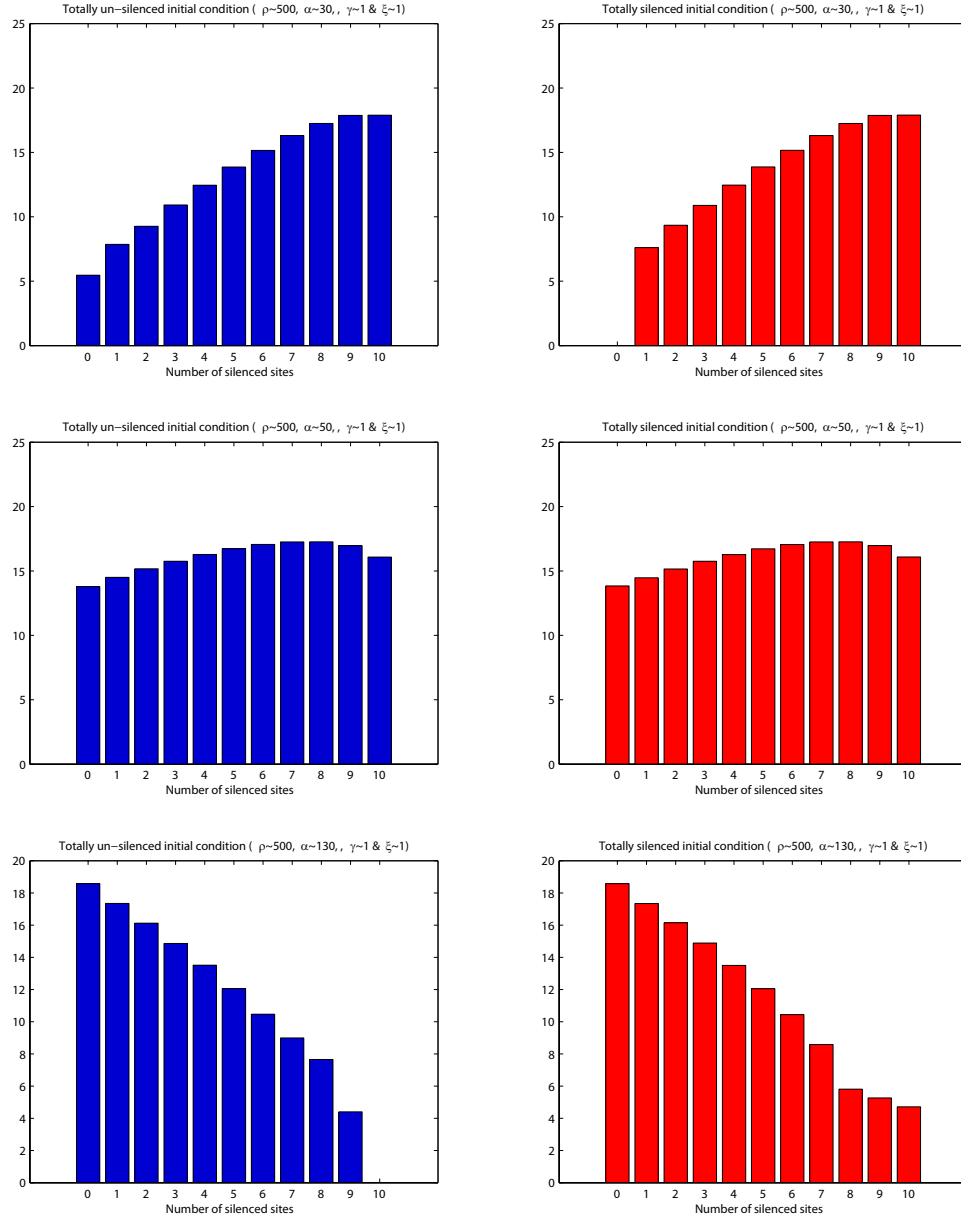


Figure 4.7: Histogram of the natural logarithm of number of silenced sites in a $L=10$ site lattice, starting from two opposite initial conditions. The corresponding values of chemical parameters are indicated at the top of each graph. Note that, the degree of cooperativity n equals 2 and total number of samplings is more than 10^8 . The transition from average intermediate state, inside mono-stable region of parameter space, to a different average state trough changing parameters, happens faster comparing to a bi-stable system, Fig.4.6

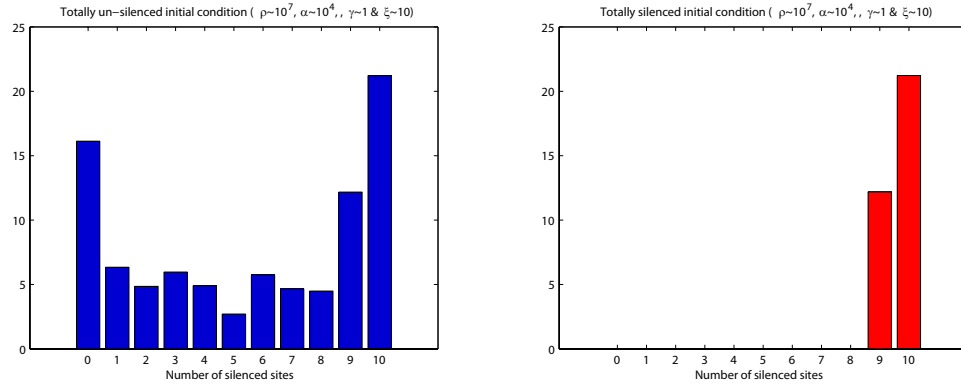


Figure 4.8: Histogram of the natural logarithm of number of silenced sites in a $L=10$ site lattice, starting from two opposite initial conditions. The corresponding values of chemical parameters are indicated at the top of each graph. Note that, the degree of cooperativity n equals 2 and total number of samplings is more than 10^8 . The diagrams show that, in shorter time scales the system in Fig.4.5 actually exhibits bi-stable states.

Note that, we can also check the difference between two cases above by assuming a shorter screening time. As a result, one observes that, in a shorter time scale the bi-stable system actually falls into either of the stable states and it stays there for some time until it switches to the other state (Fig. 4.8). In contrast, the mono-stable system converges to the same final state very fast. In conclusion, as it was said earlier, by differentiating between the systems according to the time of their convergence, one will be able to find qualitatively the same shape of bifurcation diagram as we discovered in mean-field approach.

Chapter 5

Alternative Possible Non-linearity in the System

In the model discussed so far, we assumed $f(x) = x^n$ in the main equations (2.1) and (2.2) and bi-stability demands that that $n > 1$, meaning that we need a certain degree of cooperativity in how de-acetylated histones recruit the silencing complex. This might very well be the case. However, the cooperativity in that particular interaction is not absolutely essential when we have other non-linear effects in play.

One rather plausible effect is as follows. Transcription of a gene is often associated with a higher rate of acetylation of histones. It is believed to be one of the reasons why highly transcribed genes are hard to silence. For example, a tRNA gene, usually producing a large amount of RNA, has been found to have an important role in a silencing boundary [24]. One might therefore imagine that silencing, which affects local transcription rates, indirectly affects the local acetylation rate. One way to model this is to introduce an additional function $g(1 - S_i)$ in the local acetylation rate making it $\alpha(1 - A_i(t))(1 - S_i(t))g(1 - S_i(t))$. If there is no such feedback from silencing, we could have $g(y) = 1$. We will consider $g(y) = y^{m-1}$, $m = 1$ being the case of no feedback, where as the simplest models of feedback would lead to $m = 2$. For a general value of m (and n) our model now would be given by the following equations.

$$\frac{dS_i(t)}{dt} = \rho_i(t)(1 - S_i(t))(1 - A_i(t))^n - \eta S_i(t) \quad (5.1)$$

$$\frac{dA_i(t)}{dt} = \alpha(1 - A_i(t))(1 - S_i(t))^m - (\lambda + \sum_j \gamma_{ij} S_j(t))A_i(t) \quad (5.2)$$

Thus, the nature of non-linearity in these models is characterized by a number doublet (m, n) . We discussed $(1, n)$ models in the previous chapters and found that we need n to be greater than one for these subclass of models. So to study the general (m, n) models, we can follow the same routine. Thus, to find a bifurcation diagram, generalized nullcline equations are as follows:

$$\bar{\rho}(1 - S)(1 - A)^n - S = 0 \quad (5.3)$$

$$\bar{\alpha}(1 - A)(1 - S)^m - (1 + \bar{\gamma}S)A = 0 \quad (5.4)$$

We also need to find the general form of equation (2.11) to hold at the boundary of the bi-stability region:

$$A = \frac{1 + \bar{\gamma}S}{nS[(m - 1)\bar{\gamma}S + m + \bar{\gamma}]} \quad (5.5)$$

Using above equations, we can use S as a parameter and plot $(\bar{\alpha}(S), \bar{\rho}(S))$ to get the phase diagram for (m, n) models (Fig.5.1). As a result we found that for $(m, 1)$, $m > 1$ the shape of the bi-stability region is qualitatively the same wedge-like band. In general, for any $m \geq 1$ and $n \geq 1$ excluding $m = n = 1$ the (m, n) model will result in a bifurcation diagram with essentially the same structure (Fig.5.1). As a consequence, many of the qualitative results we provided for $(1, n)$ models is also true for the general case (m, n) models, provided that m and n are never both less than or equal to 1.

In other words, both kinds of models, those with Sir binding depending strongly non-linearly on the degree of de-acetylation as well as those where the effect of silencing on local transcription feeds back on the acetylation rate, show qualitatively similar behavior needed to describe silencing in nucleus. Hence we will continue using the results of the $(1, n)$ models, fully keeping in mind that there is a broader class of models leading to the same qualitative predictions.

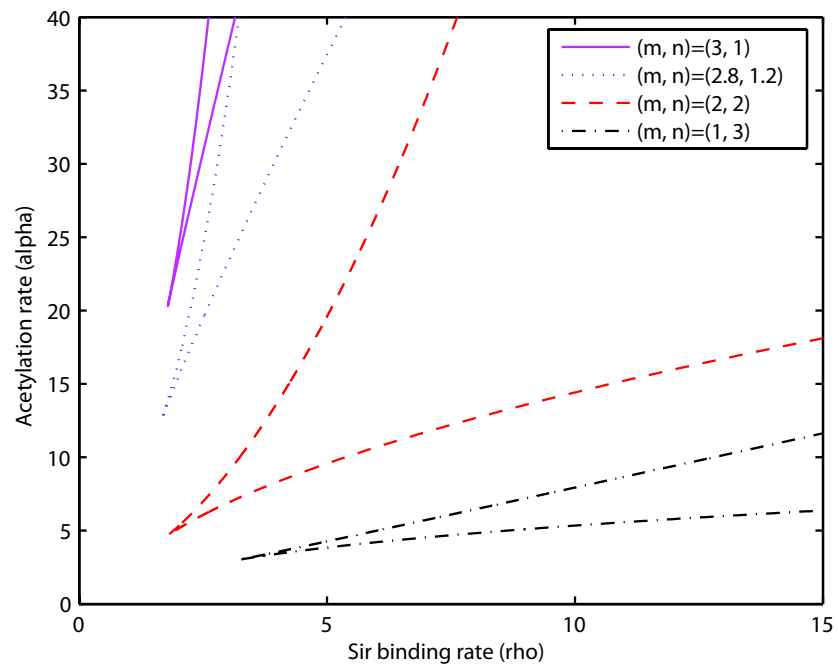


Figure 5.1: The qualitative configuration of bifurcation diagram is essentially the same for (m, n) models, with $m \geq 1$ and $n \geq 1$ excluding $m = n = 1$.

Chapter 6

Biological Consequences of the Model

The bifurcation diagram presents a classification of qualitatively different kinds of dynamics possible within the model. It provides us with a more precise vocabulary for discussing qualitative consequences of alternative models. Combining this with experimental facts, we should be able to place the wild type yeast and various mutants in this diagram.

Outside the bi-stable region, in the un-silenced region, the dynamics decides a self-consistent level of silencing. Recruitment of silencing complex at one place only affects a small region around it, with the effects dying off exponentially with distance from the nucleation center. The upper part of the bi-stable region, with higher values of α (Region I in Fig.6.1), is not qualitatively very different in that regard. The only difference comes in, when one considers stochastic dynamics, which allows for occasional formation of silencing in the whole region.

In the lower half of the region, Region II in Fig.6.1, and also in the silenced region of phase space, nucleation leads to spreading. This is the region where the naive expectation from the popular biological model matches the results of mathematical analysis. We have argued, that under some conditions, the dynamics of Sir depletion would lead the systems starting in this region into the border of the two regions (zero front velocity curve, Fig.6.1). A locus of DNA, described by parameters of region II, and the silenced region could possibly see non-specific silencing induced by stochastic nucleation of silencing.

But in this bifurcation diagram, where is the point corresponding to silencing

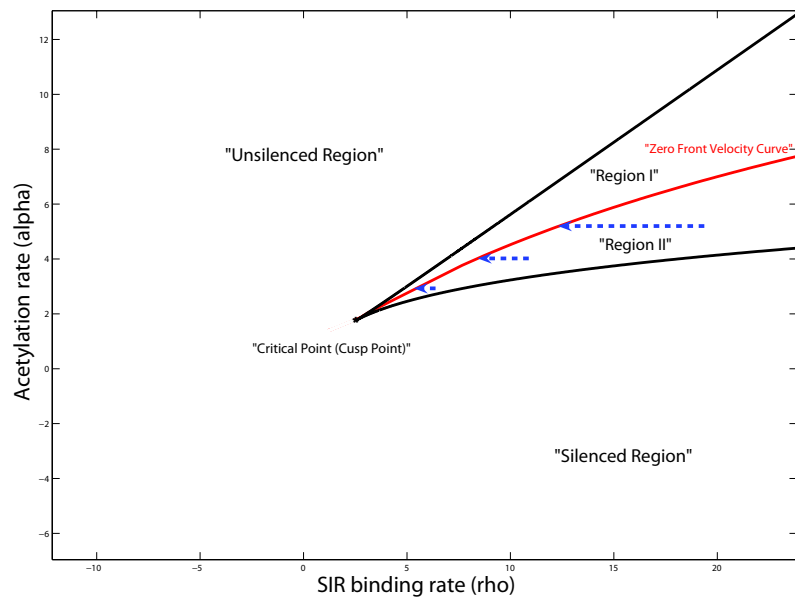


Figure 6.1: Changes in the rate of silencing, ρ in response to decreasing of the rate of acetylation, α , when the total supply of Sir complexes are limited. The system approaches the zero velocity line close to the critical point.

dynamics in silent mating loci in wild type yeast? The fact that the silent loci in the *sir1* mutants could be in either state, suggests that one is in the region allowing bi-stability. A tougher question to answer is which part of the bi-stable region it is in. In Region II, there is quite some chance of getting undesirable non-specific silencing. On the other hand, in Region I, typically, the silencing spreads very little from the nucleation center, unless the system is very close to the cusp point (or critical point, Fig.6.1). In fact, if one defines a length scale by how far the effect of silencing local nucleation spreads, that length scale diverges exactly at the cusp point. The system could operate at a point where this length scale is large.

The dynamics represented in the popular cartoon model of silencing, reviewed in [3], corresponds the behavior in Region II. Such models come with explicit requirement of boundary elements to stop the spreading. On top of that, there should be an argument why the chance of nucleation in somewhere else in the genome does not cause spontaneous non-specific silencing. Alternatively, one could possibly argue why the probability of accidental nucleation is low. However, if the system is in Region I, then one could observe a reduction in silencing with increasing the distance from the nucleation center, namely the silencer. Such a claim has been made by some researchers [22]. Although one could argue for both options, we would side with the second option, namely, that the operating point of wild type silencing loci is in Region I, but not too far from the cusp point to have a large spreading length scale.

We could now discuss the consequences of lowering the acetylation rate as it happens in, say, the *sas2* mutant [25, 26]. We argued that if there are fronts of silencing that are not pinned down by boundary elements somewhere in the genome, then our argument about ρ (Sir binding rate) reducing and moving the system back to the zero velocity line/region, applies. This is indeed a possibility in yeast. Although the silent mating loci have well defined boundary elements,

the same may not be true of all the telomeric regions. This result might explain certain counterintuitive features of mutants of certain genes like *sas2* which code for acetylases. If the reduced acetylation rate in *sas2* mutant is close to a certain level, the system will flow back to close to the cusp point at tip of the bi-stable region. Near the cusp point, the degree of silencing changes very sharply with the changes of Sir availability. We believe that the resulting system becomes extremely susceptible to cellular noise and would display a wide distribution of expression. Thus, as opposed to the naive expectation that *SAS2* deletion will just make everything more transcriptionally silent, one should find individual cells that show good expression from the “silent” loci. We speculate, whether this is the reason why the *SAS* genes may have been picked up in an assay looking for defects in silencing. Recent single cell measurements observation for GFP expression from *sir1sas2* gene mutant cells show a wide but uni-modal distribution of expression in a cell population, where as *sir1* cell population show bimodal distribution, characteristic of epigenetic states [39].

Another simple consequence the bifurcation diagram that one could say qualitative things about the epigenetic switching rate in different parts of the bifurcation diagram. For example, we expect the switching rate to get faster near the cusp point. We expect as the level of Sas2 is lowered continuously (decreasing α), we will see a rise in switching rate, as the system would move toward the cusp point.

Chapter 7

Discussion

We have formulated a mathematical version of the model of silencing and computed the bifurcation diagram of the system. This diagram is consistent with several observations about mutants. It is, in principle, possible to explore the whole two dimensional control parameter space experimentally. For example, one could study single cell GFP fluorescence from a reporter in *HMR* while modifying ρ by regulating Sir proteins, and modulating α via changing the level of Sas2.

In addition to the *sas2* mutant, which we discussed extensively, one of the mutants that we want to understand is *dot1*. Part of the reason to study this mathematical model is the apparent paradox: if the Sir2,3,4 system itself can propagate further from region with stochastic nucleation of silencing, why many other regions, not contiguous to silencing at nucleation sites, do not show occasional heritable silencing in manner . In fact, a screen high copy disruptors of telomeric silence [29], produced, among others, a gene called *DOT1* whose deletion cause nonspecific silencing. Understanding how Dot1 affects silencing requires us to consider additional states like methylation of histones [30]. Based on our preliminary study of a full model of the system with additional states it seems that the simpler model studied in this paper, with some change of parameters, could effectively capture the effect of Dot1. This is one future direction that we are pursuing.

We finally mention two issues not dealt at all within this dissertation that needs further attention. One is that our model of DNA, as a one dimensional

system, may be called into question if the heterochromatin formation happens very fast (compared to the speed with which silencing spreads), making the DNA fold up into higher order organization quickly. The other interesting issue is inheritance of silencing. Could we have our model capture inheritance in a coarse grained manner, or do we stand to gain something by modeling the probable silencing of duplicated DNA explicitly? Of course, for any biological model, there are many ways of making it more realistic. However, not many of these ‘improvements’ change the qualitative properties of the bifurcation diagram. We believe our model includes enough features of the biological phenomena to be a good starting point for more refined discussion of the qualitative behavior of this system.

Chapter 8

Appendix A: Uniform Fixed Points and Their Stability

In this appendix, we consider a more general form of the equations (2.3) and (2.4) and discuss the mathematical methods which can be used to find the stable fixed points of the system. Consider the following uniform dynamical equations:

$$\frac{dS(t)}{dt} = \rho(1 - S(t))f(1 - A(t)) - \eta S(t) \quad (8.1)$$

$$\frac{dA(t)}{dt} = \alpha(1 - A(t))g(1 - S(t)) - (\lambda + \gamma S(t))A(t) \quad (8.2)$$

where all the Greek letters are positive constant variables. S and A are real parameters that can only take values between 0 and 1, inclusive. $f(x)$ and $g(y)$ are non-negative, monotonically increasing function in the interval $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Note that, uniform version of both the set of non-linear equations (5.1) and (5.2) and the general form of acetylation term mentioned in foot notes of section 2.1, are special cases of above equations. Assume the change of variables, $X = 1 - A$ and $Y = 1 - S$; rewriting the above equations in X and Y :

$$\frac{dY}{dt} = \eta - Y(\rho f(X) + \eta) \quad (8.3)$$

$$\begin{aligned} \frac{dX}{dt} &= (\lambda + \gamma - \gamma Y) - X [\alpha g(Y) + (\lambda + \gamma - \gamma Y)] \\ &= (\lambda + \gamma - \gamma Y) \left[1 - X \left(\frac{\alpha g(Y)}{\lambda + \gamma - \gamma Y} + 1 \right) \right] \end{aligned} \quad (8.4)$$

Now, define functions $\tilde{f}(X)$ and $\tilde{g}(Y)$ as follows:

$$\tilde{f}(X) = \bar{\rho}f(X) \quad (8.5)$$

$$\tilde{g}(Y) = \frac{\bar{\alpha}g(Y)}{1 + \bar{\gamma} - \bar{\gamma}Y} \quad (8.6)$$

where $\bar{\alpha} = \alpha/\lambda$, $\bar{\rho} = \rho/\eta$ and $\bar{\gamma} = \gamma/\lambda$. Both $\tilde{f}(X)$ and $\tilde{g}(Y)$ follow the same property we assumed for $f(x)$ and $g(y)$, i.e. non-negative, monotonically increasing functions in the interval $0 \leq X \leq 1$ and $0 \leq Y \leq 1$. So, in terms of these new functions we have:

$$\frac{dY}{dt} = \eta [1 - Y(\tilde{f}(X) + 1)] = F(X, Y) \quad (8.7)$$

$$\frac{dX}{dt} = (\lambda + \gamma - \gamma Y) [1 - X(\tilde{g}(Y) + 1)] = G(X, Y) \quad (8.8)$$

Then, the fixed points for above equations satisfy following nullclines:

$$Y_f = \frac{1}{\tilde{f}(X_f) + 1} = \hat{f}(X_f) \quad (8.9)$$

$$X_f = \frac{1}{\tilde{g}(Y_f) + 1} = \hat{g}(Y_f) \quad (8.10)$$

where X_f and Y_f denote the fixed point values. Note that $\hat{f}(X)$ and $\hat{g}(Y)$ are monotonically decreasing functions inside the interval $0 \leq X \leq 1$ and $0 \leq Y \leq 1$, which themselves are confined between 0 and 1. For the special cases being studied in this dissertation, the general shape of \hat{f} and \hat{g} is given in Fig.8.1; i.e. either they decrease with a positive curvature (positive second derivative) or decrease like an *S-curve*. The solution to above equations is then given by the intersection of one such curve with another one reflected at the $x = y$ line. Therefore, it is easy to see that for all cases in this dissertation, there is either one or three answers to above equations.

One can investigate the stability of the fixed points by Taylor expansion of equations (8.7) and (8.8) around the fixed points. So if $X = X_f + \delta X$ and $Y = Y_f + \delta Y$, for small δX and δY , keeping up to the first non-zero term:

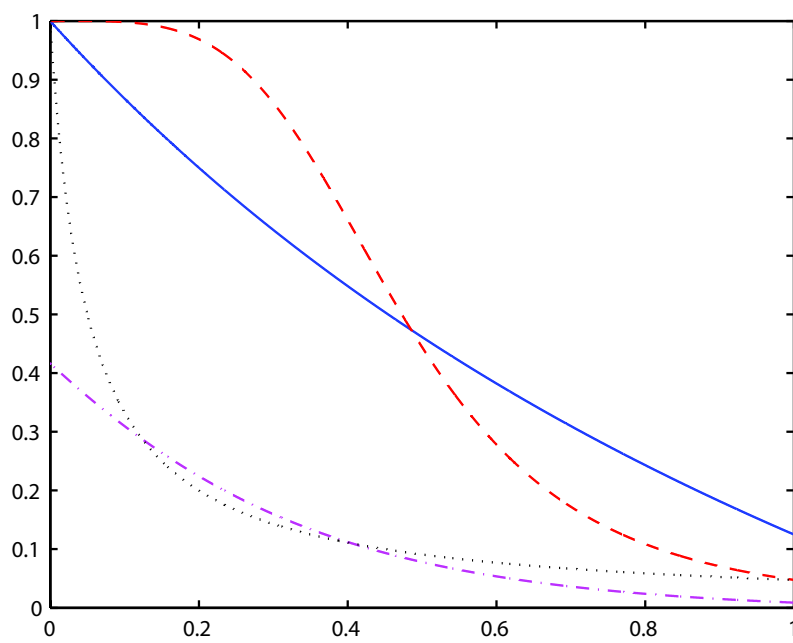


Figure 8.1: The shape of monotonically decreasing functions \hat{f} and \hat{g} in general nullclines.

$$\frac{d}{dt} \begin{pmatrix} \delta Y \\ \delta X \end{pmatrix} = \begin{pmatrix} \frac{\partial F}{\partial Y} & \frac{\partial F}{\partial X} \\ \frac{\partial G}{\partial Y} & \frac{\partial G}{\partial X} \end{pmatrix} \begin{pmatrix} \delta Y \\ \delta X \end{pmatrix} \quad (8.11)$$

In order to have stability at (X_f, Y_f) , the above Jacobian matrix has to have negative eigenvalues. The two eigenvalues satisfy the below quadratic equation:

$$Z^2 - \left(\frac{\partial F}{\partial Y} + \frac{\partial G}{\partial X} \right) Z + \left(\frac{\partial F}{\partial Y} \frac{\partial G}{\partial X} - \frac{\partial G}{\partial Y} \frac{\partial F}{\partial X} \right) = 0 \quad (8.12)$$

So to have negative solutions, we should have:

$$\frac{\partial F}{\partial Y} + \frac{\partial G}{\partial X} < 0 \quad (8.13)$$

$$\frac{\partial F}{\partial Y} \frac{\partial G}{\partial X} - \frac{\partial G}{\partial Y} \frac{\partial F}{\partial X} > 0 \quad (8.14)$$

According to equations (8.7) and (8.8), both $\frac{\partial F}{\partial Y}$ and $\frac{\partial G}{\partial X}$ are negative when $0 \leq X \leq 1$ and $0 \leq Y \leq 1$, hence the condition (8.13) is always satisfied. Since this condition means that the sum of eigenvalues is negative, we can either have two negative eigenvalues or one negative and the other positive. In other words, each fixed point can either be a stable point or an unstable saddle point. In the three fixed point regime then, we can either have two stable points flanking the middle saddle point or two saddle points flanking the middle stable point. However since for small Y , dY/dt is always positive (equation (8.7)) and it only changes sign when Y passes its nullcline value (equation 8.9) and also since the same holds for X ; one can find that the middle fixed point is always the saddle point. Hence the flanking points are actually stable. One can also draw the vector field of the flow $(\frac{dX}{dt}, \frac{dY}{dt})$, using softwares such as *MATLAB*, to investigate the stability of the points (Fig. 2.3).

Note that, now in order to find the numerical values of stable fixed points, we can always start from either $X = 0$ or $Y = 0$ and by recursively plugging X and Y inside equations (8.9) and (8.10) and obtaining new values approach the stable fixed point answers.

Chapter 9

Appendix B: Numerical Methods in Discrete Model Approach

We use an array of variable length $L = 10$ to 200 sites, setting the value of S and A at each boundary equal to either one of the two stable uniform solutions. Two forms of γ_{ij} is considered. First, γ_{ij} is assumed to be non-zero only for nearest neighbors ($|i-j| = 1$) and be zero elsewhere. Second, $\gamma_{ij} = \gamma_0 \exp(-|i-j|/\xi)$ with a positive length scale, ξ . For the initial condition, we assume a configuration with a very sharp transition in the middle of the lattice i.e. half of the sites on one side are set to hypo-acetylated/silenced uniform solution and the rest of the lattice which are on the other side are set to hyper-acetylated/un-silenced uniform solution. Then an *explicit finite difference method* is employed to find the numerical answers of equations 2.1 and 2.2:

$$S_{ij+1} = \Delta t [\rho(1 - S_{ij})f(1 - A_{ij}) - \eta S_{ij}] + S_{ij} \quad (9.1)$$

$$A_{ij+1} = \Delta t \left[\alpha(1 - A_{ij})(1 - S_{ij}) - (\lambda + \sum_k \gamma_{ik} S_{kj}) A_{ij} \right] + A_{ij} \quad (9.2)$$

where j indicates the number of steps in time, and Δt is the time step. Convergence to a stable state was very fast and in most cases with only slight change from the initial configuration.

At the end, to check that our answer has actually converged to the stable solution, we can apply the time independent form of equations (2.1) and (2.2) recursively on the answer:

$$S_i = \frac{\rho f(1 - A_i)}{\rho f(1 - A_i) + \eta} \quad (9.3)$$

$$A_i = \frac{\alpha(1 - S_i)}{\alpha(1 - S_i) + \lambda + \sum_k \gamma_{ik} S_k} \quad (9.4)$$

Chapter 10

Appendix C: Monte Carlo Simulation of the Stochastic Model

In order to employ a stochastic simulation of equations (4.1) and (4.2), one can take a Monte Carlo approach. Assume that we have a lattice of size L and at each site i of this lattice n digital parameters $A_i^{(m)}$, $m = 1, 2, \dots, n$ and one digital parameter S_i . S_i and $A_i^{(m)}$ can take only be 0 or 1. We also choose the parameter γ_{ij} to fall off exponentially with a length scale ξ and apply periodic boundary condition. Then, at each step of the simulation, one picks a random site i and also one of the parameters $\{S_i, A_i^{(m)}\}$ at random; then follows the rules below:

1. for $A_i^{(m)}$;
 - (a) if $A_i^{(m)} = 1$
 - i. if $(\lambda + \sum_j \gamma_{ij} S_j) \geq \alpha(1 - S_i)$, switch to $A_i^{(m)} = 0$
 - ii. else, switch to $A_i^{(m)} = 0$ with the rate $\frac{\lambda + \sum_j \gamma_{ij} S_j}{\alpha(1 - S_i)}$.
 - (b) if $A_i^{(m)} = 0$
 - i. if $(\lambda + \sum_j \gamma_{ij} S_j) \leq \alpha(1 - S_i)$, switch to $A_i^{(m)} = 1$
 - ii. else, switch to $A_i^{(m)} = 1$ with the rate $\frac{\alpha(1 - S_i)}{\lambda + \sum_j \gamma_{ij} S_j}$.
2. for S_i ;
 - (a) if $S_i = 1$
 - i. if $\eta \geq \rho \prod_m (1 - A_i^{(m)})$, switch to $S_i = 0$

- ii. else, switch to $S_i = 0$ with the rate $\frac{\eta}{\rho \prod_m (1 - A_i^{(m)})}$.
- (b) if $S_i = 0$
 - i. if $\eta \leq \rho \prod_m (1 - A_i^{(m)})$, switch to $S_i = 1$
 - ii. else, switch to $S_i = 1$ with the rate $\frac{\rho \prod_m (1 - A_i^{(m)})}{\eta}$.

For each set of control parameters, we run the simulation twice i.e. for opposite initial configurations: uniformly silenced/hypo-acetylated and uniformly un-silenced/hyper-acetylated lattice. As a result, we avoid getting stuck in only one of the two local minima (fixed points) by choosing each opposite fixed point once.

References

- [1] Alberts, B., et al. (2002) *Molecular Biology Of The Cell* (Gerland, New York, NY).
- [2] Lodish, H., et al. (2004) *Molecular Cell Biology* (WH Freeman, New York, NY).
- [3] Stillman, B. & Stewart, D. (Editors) (2005) in *Epigenetics (Cold Spring Harbor Symposia on Quantitative Biology)* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), Vol. 69.
- [4] Gilbert, S. (2003) *Developmental Biology* (Sinauer, Sunderland, MA).
- [5] Grewal, S.I. & Moazed, D. (2003) *Science* **301**, 798-802.
- [6] Moazed, D., Rudner, A.D., Huang, J., Hoppe, G.J. & Tanny, J.C. (2004) in *Reversible protein acetylation (Novartis Foundation Symp. 259)* (Wiley, Chichester), pp. 48-62.
- [7] Rusche, L. N., Kirchmaier, A. L. & Rine, J. (2002) *Mol. Biol. Cell* **13**, 2207-2222.
- [8] Aparicio, O. M., Billington, B. L., Gottschling, D. E. (1991) *Cell* **66**, 1279-1287.
- [9] Pillus, L. & Rine, J. (1989) *Cell* **59**, 637-647.
- [10] Ptashne, M. (1992) *A Genetic Switch* (Cell Press and Blackwell, MA).
- [11] Shea, M.A. & Ackers, G.K. (1985) *J. Mol Biol.* **18**, 211-230.
- [12] Arkin, A., Ross, J. & McAdams, H.H. (1998) *Genetics* **149**, 1633-1648.
- [13] Novick, A. & Weiner, M. (1957) *Proc. Natl. Acad. Sci. USA* **43**, 553-566.
- [14] Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I. & van Oudenaarden, A. (2004) *Nature* **427**, 737-740.
- [15] Vilar, J.M., Guet, C.C. & Leibler, S. (2003) *J. Cell. Biol.* **161**, 471-476.
- [16] Gardner, T.S., Cantor, C.R. & Collins, J.J. (2000) *Nature* **403**, 339-342.
- [17] Aronson, D.G. & Weinberger, H.F. (1975) in *Partial Differential Equations and Related Topics* ed. Goldstein, J.A. (Springer Lecture Notes in Mathematics), Vol. 446, pp. 5-49.

- [18] Cross, M. C. & Hohenberg, P. C. (1993) *Reviews of Modern Physics* **65** (**3**, part II), 851-1112.
- [19] Fall, C.P., Marland, E.S., Wagner, J.M. & Tyson, J.J. (Editors) (2002) *Computational Cell Biology* (Springer, NY).
- [20] Keener, J. S. (1987) *SIAM J. Appl. Math.* **47**, 556-572.
- [21] Bi, X. & Broach, J. R. (1999) *Genes Dev.* **13**(9), 1089-1101.
- [22] Xu, E.Y., Zawadzki, K.A. & Broach, J.R. (2006) *Mol. Cell* **23**, 219-229.
- [23] Bi, X., Braunstein, M., Shei, G. & Broach, J.R. 1999 *Proc. Natl. Acad. Sci. USA* **96**(21), 11934–11939.
- [24] Kamakaka, R. M., et al. (1999) *Genes Dev.* **13**, 698-708.
- [25] Kimura, A., Umehara, T. & Horikoshi, M. (2002) *Nat. Genet.* **32**, 370-377.
- [26] Suka, N., Luo, K. & Grunstein, M. (2002) *Nat. Genet.* **32**, 378-383.
- [27] Schaper, S., Franke, J., Meijnsing, S.H. & Ehrenhofer-Murray, A.E. (2005) *J. Cell Sci.* **118**(7), 1473-1484.
- [28] Ehrenhofer-Murray, A.E., Rivier, D.H. & Rine, J. (1997) *Genetics* **145**(4), 923-934.
- [29] Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M. & Gottschling, D.E. (1998) *Genetics* **150**, 613-632.
- [30] van Leeuwen, F., Gafken, P.R. & Gottschling, D.E. (2002) *Cell* **109**, 745-756.
- [31] Rao, C.V., Wolf, D.M. & Arkin, A.P. (2002) *Nature* **420**, 231-237. Erratum (2003) *Nature* **421**, 190.
- [32] Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D. & van Oudenaarden, A. (2002) *Nat. Genet.* **1**, 69-73.
- [33] Hasty, J., Pradines, J., Dolnik, M. & Collins, J.J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 2075-2080.
- [34] Yarchuk, O., Jacques, N., Guillerez, J. & Dreyfus, M. (1992) *J. Mol. Biol.* **226**, 581-596.
- [35] Chapon, C. (1982) *EMBO J.* **1**, 369-374.
- [36] Blake, W.J., Kærn, M., Cantor, C.R. & Collins, J.J. (2003) *Nature* **422**, 633-637.
- [37] Raser, J. M. & O'Shea, E. K. (2005) *Science* **304**, 1811-1814.

- [38] Grewal, S.I. & Elgin, S.C. (2002) *Curr. Opin. Genet. Dev.* **12**, 178-187.
- [39] Xu, E.Y., Zawadzki, K.A. & Broach, J.R. (2006) *Mol. Cell* **23(2)**, 219-229.
- [40] Lee, T.I., et al. (2006) *Cell* **125(2)**, 301-313.
- [41] Bernstein, B.E., et al. (2006) *Cell* **125(2)**, 315-326.

Vita

Mohammad Sedighi

- 2007** Ph.D. in Physics, Rutgers University
- 2000** B.Sc. in Applied Physics from Sharif University of Technology
- 1996** Graduated from Allameh Helli High School
-
- 2001-2006** Teaching Assistant, Department of Physics and Astronomy, Rutgers University