Chapter 1

Introduction

As the basis of all life phenomena, the interaction of biomolecules has been under the scrutiny of scientists and cataloged meticulously [2]. The recent advent of systems biology represents a natural development from this collection of knowledge: it is more reliable to predict the property of a biological system by taking appropriate account of all the underlying components and their interaction. Unfortunately, the knowledge, which systems biology relies on as the starting point, was usually accumulated by in vitro studies of only a small subset of the system of interest under fundamentally different conditions from its in vivo counterpart. In particular, molecules in a living cell normally find themselves in an extremely crowded environment with a small copy number and significant spatial localization.

Molecular crowding is caused by the existence of a high concentration of macromolecules. On average, each protein occupies a volume of about 10^3 nm^3 in *E. coli* [106]. This is far smaller than that in the dilute solution one would have in a typical in vitro assay. In another word, the free space associated with each macromolecule is comparable to the physical size of the molecule itself. This picture is artistically illustrated in figure 1.1. The diffusion coefficients of biomolecules in such a crowded environment are expected to be smaller than their values in a dilute solution, but the existence of concentrated solute effectively exerts an attracting potential between binding partners [122]. These two effects oppose each other in their influence on bimolecular reaction kinetics, making it difficult to predict the overall consequence of crowding.

It is also very common to find only a handful of each type of molecule in a cell [110]. For example,

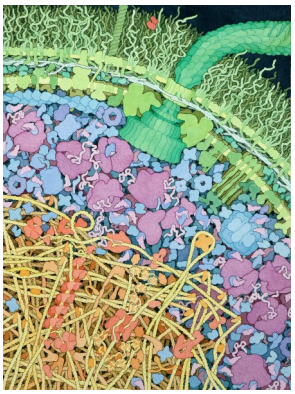

Figure 1.1. The crowded environment in an *E. coli* cell depicted by professor Goodsell (`http://mgl.scripps.edu/people/goodsell`).

there are at most two copies of genetic material and the numbers of regulators are also limited. As a result, the kinetics of chemical reactions is under the influence of stochastic noise and normally different quantitatively or even qualitatively from the prediction of law of mass action [117]. A notable example is the lysis–lysogeny switch of *E. coli*. The seemingly random choice between the two possible fates can be explained by taking into account the stochasticity of chemical reactions when the copy numbers are small (reference [6] and figure 1.2). Finally, real cells are highly compartmentalized. Chemical reactions take place with a well-defined boundary rather than in an infinite space. The boundary effect is more important when system dimension is decreased because surface-to-volume ratio inversely scales as its size. Reaction in a test tube is hardly affected by the presence of boundary, whereas spatial confinement may have important implications in the reaction kinetics in highly localized cellular environments.

Recognizing the important features of in vivo biochemical reactions, it is imperative to study the interaction among biomolecules in a cell-mimetic environment. This thesis summarizes some results of both theoretical and experimental investigations. In chapter 2, the effect of low copy number is exemplified by analyzing the dynamics of a classic chemical oscillator (section 2.1.1) and a well-known biological oscillator (section 2.1.2). In both cases, stochastic noise in chemical reactions adds significant perturbation to the system dynamics, and law of mass action is shown to be insufficient in some regime. Gillespie algorithm is used to simulate the dynamics, and the strength

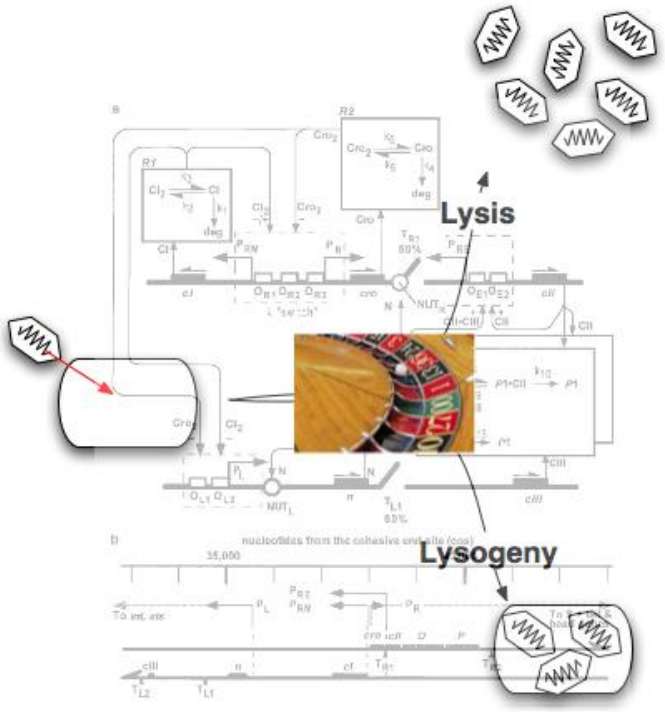

Figure 1.2. *E. coli* cells infected by phage-λ have two possible fates: lysogeny or lysis. The choice is randomly made by the inherent stochasticity of the underlying biochemical reaction network. The reaction scheme in the background is taken from reference [6].

of noise is tuned by changing system volume. In section 2.1.1, a variant of Belousov–Zhabotinsky system is shown to exhibit persistent oscillation close to the bifurcation point where deterministic simulation predicts steady-state behavior. The regularity of the oscillation, as a function of system volume, may have a maximum, indicating the existence of stochastic resonance. In section 2.1.2, the dynamical behavior of peroxidase–oxidase oscillator is investigated as a function of initial conditions. Around the bifurcation point, the system chooses its fate randomly in a way conceptually similar to the lysogeny–lysis decision process (figure 1.2). The effect of confinement is then the topic of section 2.2, which relies on scaling analysis and stochastic simulation. The spatial confinement leads to the synchronization of a group of enzymes coupled by rapid diffusion of small molecules. A more fundamental discussion of molecular diffusion in confined space follows in section 2.3 with the help

of both analytical solution and numerical simulation. Here special attention is paid to the statistical property of molecular collisions during diffusion and the clustering of these events in space and time. Methods of spatial stochastic simulation are also briefly discussed.

The development of microfluidic platforms for the investigation of biochemical reactions with configurable constraints is the subject of chapters 3 and 4. Both platforms provide convenient methods to initiate enzymatic reactions with tunable crowding and spatial confinement, a crucial requirement for simulating intracellular environment. The design of the devices also allows for experimental observation over a wide range of timescale from hundreds of milliseconds to dozens of minutes after the reaction starts. As a result, both transient and steady-state behavior are accessible using the same setup. As a proof of concept, chapter 3 reports the application in the observation of the catalytic activity of single β-Galactosidase (β-Gal) restricted in femtoliter-sized microchamber arrays. Repetitive observation of the chemical reactions in the array can be carried out in the same device to get statistically significant results. The dimension of each chamber is less than 10 μm and can be further reduced to increase throughput. Chapter 4 is based on the generation and manipulation of water-in-oil droplets with well-defined and tunable diameter smaller than 10 μm, comparable in size with cellular compartments. Details of the experimental setup to control the microfluidic device are included in this chapter. As the most important finding of this work, the history of droplet formation has a significant effect on the interfacial property of the droplets and leads to a decrease of apparent enzyme activity with decreasing droplet size. This observation is expected to motivate more discussion in microfluidic literature as devices with ever-smaller characteristic size are to be utilized.

The effort devoted to the design and optimization of these microfluidic networks also has practical significance in detecting biomolecules. Many popular detection schemes are based on the generation of fluorescent products catalyzed by the molecule to be detected [7, 74, 91, 113]. For qualitative analysis, the accumulated fluorescence signal crossing the detection threshold is used as evidence for the existence of a certain molecule. Although reducing the volume of the reactor, one of the advantages offered by microfluidics, does not change the sample concentration, the following analysis will show that under the condition of extreme dilution, the detection time may be reduced by using

microreactors. In bulk assay, the fluorescent product concentration c_{P} is expected to be proportional to the concentration c_{E} of the enzyme to be detected, or $c_{\mathrm{P}} = c_{\mathrm{E}}t/\tau_{\mathrm{E}}$ where τ_{E} is the turnover time. If a detection threshold is set to be c_0, one has to wait for at least $t_{\mathrm{bulk}} \sim c_0\tau_{\mathrm{E}}/c_{\mathrm{E}}$ to get the measurement result. In order to reduce t_{bulk} for a certain enzyme, one has to increase c_{E}. In practice, however, it is not always feasible to concentrate the sample as much as one wishes due to practical or economical reasons. This expression of t_{bulk} also seems to suggest that reducing reactor size does not accelerate the detection process. This conclusion, however, is not true when the average number of the enzyme in a reactor is fewer than one due to the reduction of volume, concentration, or both. In fact, for a reactor of volume V, when $c_{\mathrm{E}}V \ll 1$, there is at most one enzyme present and the product concentration can be written as $c_{\mathrm{P}} = t/(V\tau_{\mathrm{E}})$. The time it takes to cross the threshold c_0 is thus $t_{\mathrm{micro}} \sim c_0V\tau_{\mathrm{E}}$. It is obvious to see that reducing system volume will reduce the time consumption linearly. Of course, when the enzyme solution is diluted, one has to carry out a large number N of repeated measurements to get reliable result. Normally, one would require $Nc_{\mathrm{E}}V > 1$. If these repetitive experiments were done sequentially, the total time consumption $Nt_{\mathrm{micro}} \gg t_{\mathrm{bulk}}$, meaning that there is no benefit of system-size reduction. Fortunately microfluidic techniques provide an opportunity of large-scale parallelization [159], which makes it possible to make ensemble measurement with little added time consumption. For example, a large array of microreactors may be observed simultaneously (section 3 and references [113] and [116]). If M reactors can be fabricated in an array, the time required would be reduced by M-fold, or $t^{\mathrm{P}}_{\mathrm{micro}} \sim Nt_{\mathrm{micro}}/M$. In order for $t^{\mathrm{P}}_{\mathrm{micro}} < t_{\mathrm{bulk}}$, one has to have a large enough M such that $M > Nc_{\mathrm{E}}V$. As another example, if water-in-oil droplets are generated as in chapter 4, the enzymatic reaction is initiated once the droplet is formed, and one makes observation at a downstream location to allow for sufficient product generation if any. Because the droplets are flowing in a continuous fashion [134], the detection time is thus asymptotically proportional to the product of the number of droplets to be monitored and their generation time as $t^{\mathrm{P}}_{\mathrm{micro}} \sim Nt_{\mathrm{drop}}$ where t_{drop} is the droplet-formation time. To gain an advantage over traditional method, one just needs $t_{\mathrm{drop}} < t_{\mathrm{bulk}}/N$. Additionally, by simply counting the number of reactors showing high enough fluorescence signal, one can easily calculate the concentration of the enzyme for quantiative purpose [113]. The above analysis

establishes the practical movitation of contriving a microfluidic platform with reduced reactor size.

Finally, the thesis concludes with a summary of the lessons learned from these studies and an outline of future directions.