Non-Equilibrium Dynamics: Diffusion in Small Numbers and Ribosomal Self-Assembly

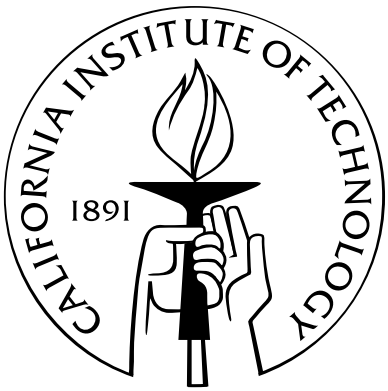
Thesis by

Effrosyni Seitaridou

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2008

(Defended May 8, 2008)

Στους γονείς μου

To my parents

Acknowledgements

I consider myself very lucky to have been at Caltech these past six years. My experience here has been a very rewarding learning process, more so than I could ever hope for. There are many people who are responsible for this, people whose guidance, help, and company were offered so abundantly. I can only begin to thank them here, since whatever I say is definitely not enough.

First and foremost, I would like to thank my advisor, professor Rob Phillips. It has been such a pleasure to interact with him and such an honor to be considered one of his students. He has created an extremely enjoyable and intellectual environment in his lab and, as a graduate student, I could not have asked for more. His guidance and encouragement were always offered, his enthusiasm conveyed what a joy it is to be doing scientific work. It has been so delightful to have known him and worked with him. I can only hope that I have been half as good of a student as he has been an advisor.

I am immensely grateful to Rob for introducing me to such great collaborators as Ken Dill, Kings Gosh, Mandar Inamdar, William Ridgeway, and Jamie Williamson. They are exceptional scientists and excellent teachers and, as such, I have learned a lot from them. It has always been a joy to talk about science with them.

During this learning process, I also had the chance to interact with my labmates, who made being in the lab so much fun. I had to enlist their help at some point or another and their friendship has meant a lot to me. I would especially like to thank Lin Han, Paul Grayson, and Mandar Inamdar. They are admirable and very helpful and kind individuals. I appreciate everything they have done for me.

My appreciation also goes to the members of my committee for taking time away from their research in order to be in my committee. I can only say that I will do my best to meet their high standards.

Last, but not least, I would like to thank my family and David for their continuous love, patience, and support. I love them with all my heart.

Abstract

Biological systems are encountered in states that are far from equilibrium. A change in the cell's condition triggers the flow of energy and matter that causes the cell's transition from that non-equilibrium state to a different state. Our interest is on non-equilibrium systems and the way these relate to the cell's "small numbers" limit as well as to the mechanisms of self-assembly.

Cells contain proteins and nucleotides in numbers smaller than Avogadro's. In addition, advances in single-molecule experiments, which are, by definition, a case of the "small numbers" problem, have emphasized the importance of fluctuations. Does the result we get from a single-molecule measurement agree with what we would get from a bulk measurement? Is it a fluctuation from the mean? It is, thus, of biological interest to see the behavior of non-equilibrium systems at the "small numbers" limit where fluctuations become important. Using microfluidics, we concentrate on the diffusion of a small number of submicron particles in a system that is away from equilibrium. Therefore, we study the "small numbers" limit of Fick's Law, with special reference to the fluctuations that attend diffusive dynamics in order to experimentally test the theoretical predictions obtained via the use of E. T. Jaynes' "principle of maximum caliber."

The process of macromolecular self-assembly is also highly dynamical. The system's components come together, defeating in this way entropic effects, to form the system. In the case of the ribosome, whose importance lies in its ability to synthesize proteins, understanding the mechanism of the highly dimensional process of self-assembly becomes relevant when designing, for example, new antibiotics. The second part of this thesis concentrates on the RNA-protein interactions which, in the case of the ribosome, determine the mechanism of self-assembly. With the use of microfluidic technology and a fluorescence assay we determine the thermodynamics and kinetics of RNA folding and RNA-protein binding for a fragment of the bacterial 30S ribosomal subunit, paving the way for the study of the complete assembly of the 30S subunit.

Contents

List of Figures

List of Tables

Chapter 1

Introduction on Non-Equilibrium Dynamics: Diffusion in Small Numbers and Ribosomal Self-Assembly

1.1 Where is the "Dynamics" in Thermodynamics?

The word "thermodynamics," from the Greek $\theta\acute{\epsilon}\rho\mu\eta$, which means heat, and $\delta\acute{\upsilon}\nu\alpha\mu\iota\varsigma$, which means power, implies that the science of thermodynamics deals with the movement of heat (energy) and how that is changing with time. The four laws of thermodynamics, as we have traditionally learned them, state the following (see, for example [1, 2]):

Zeroth Law: If two thermodynamic systems are each in thermal equilibrium with a third, then they are in thermal equilibrium with each other.

First Law (or the law of conservation of energy): The increase in the internal energy of a system is equal to the amount of energy added by heating the system, minus the amount lost as a result of the work done by the system on its surroundings.

Second Law (or the law of entropy): The entropy of an isolated system not in equilibrium will tend to increase over time, approaching a maximum value at equilibrium.

Third Law: As a system approaches absolute zero, all processes cease and the entropy of the system approaches a minimum value.

From the above statements we see that the laws of thermodynamics concern systems that are in equilibrium or at the process of obtaining an equilibrium state and there is no reference to time [2]. However, just because a system is in equilibrium does not mean that the particles this system consists of are in equilibrium. In fact, at the atomic scales of time and distance, the atomic coordinates vary

in time [1]. When looking at a system microscopically, if we can keep track of all the atomic coordinates, then we can follow a system at any point, even when it is not at equilibrium, and see how the system evolves to reach a different state at a later time. If observation of the coordinates at the atomic scales of distance and time allows us to describe a system regardless of whether it is at equilibrium or not, then why aren't these coordinates included in our study of thermodynamics? There are a couple of different reasons why that is the case. The first reason is that it is hard to keep track of all these coordinates, which are of the order of 10^{23} (Avogadro's number). As a result, it is easier to use the above laws and study systems that are in equilibrium, where the time-varying atomic coordinates become irrelevant, since we are only interested in those properties of the system that are not time-changing. In addition, another reason why the atomic scales of distance and time are not included in the study of thermodynamics is because those scales do not survive the macroscopic measurements. When taking measurements, the instruments normally can resolve down to submicron distances and cannot go down to interatomic distances, which is where the time-changing atomic motion can be observed. Furthermore, the speed by which the measurements are made (of the order of microseconds) is not fast enough to capture the time scale of atomic vibrations which is of the order of femtoseconds [1]. As a result, as Callen puts it [1]: "By definition, suggested by the nature of macroscopic observations, thermodynamics describes only static states of macroscopic systems."

We, therefore, reach an oxymoron: thermodynamics is the study of static states. Thus, it is somewhat deceiving that all these laws, which depend on macroscopic measurements and describe systems in equilibrium, are under the umbrella of "thermodynamics." Where exactly is the "dynamics" in thermodynamics? It seems that what we have traditionally learned as thermodynamics should more appropriately be called "thermostatics."

In addition to the above laws, there have been several advancements in trying to understand equilibrium systems from the point of view of statistical mechanics. Efforts have also been made in understanding systems that are not in equilibrium, even though they are still very close to reaching the equilibrium state. Even more importantly though, there have been approaches for understanding systems that are arbitrarily far from equilibrium. One of these approaches was introduced by E. T. Jaynes and is based on Information Theory and on the use of entropy. This approach is commonly known as the "Principle of Maximum Caliber" and part of this thesis aims to use it in order to obtain predictions about dynamical transport laws and also experimentally test some of the predictions it has made. In the sections that follow we present an overview of the advances made in traditional thermodynamics for equilibrium systems as well as work that was done in order to describe systems that are not in equilibrium. We then present the Maximum Entropy and Maximum Caliber approaches for studying equilibrium and non-equilibrium systems, respectively. Some of the results obtained via the use of this simple formalism and their experimental verification

will be explored in detail in Chapters 2 and 3.

1.1.1 Thermostatics and Equilibrium Systems

As has been mentioned, thermostatics, or traditional thermodynamics, deals with equilibrium systems. In addition, traditional thermodynamics does not reference the atomic spatial and temporal coordinates whose coarse averaging results in the macroscopic thermostatic observables. In order to connect the atomic temporal and spatial coordinates with the laws of traditional thermodynamics stated at the beginning of this chapter we use equilibrium statistical mechanics. In this section we describe the major highlights on the work that has been done in the area of traditional thermodynamics (or thermostatics) and equilibrium statistical mechanics. With this description, the goal is to present some important concepts, like that of entropy and partition function, that describe equilibrium systems that can later be carried across to the study of non-equilibrium systems. We will see that the Principle of Maximum Entropy (and Maximum Caliber, when there is time evolution of the quantities of interest) is based on concepts that are well-characterized in the case of equilibrium.

Perhaps the most fundamental law of equilibrium statistical mechanics is Boltzmann's equation that relates the entropy S of an isolated system to the total number Ω of microstates (i.e., specific detailed microscopic configurations of a system) at a given value of energy U, volume V, and number of moles N of the chemical component of the system. This constitutes the microcanonical ensemble and the Boltzmann equation is:

$$S(U, V, N) = k_B \ln \Omega(U, V, N), \tag{1.1}$$

where k_B is the Boltzmann constant. This equation is considered the historical foundation of statistical mechanics exactly because it makes this connection between the microscopic and the macroscopic worlds: it connects the entropy S, a macroscopic quantity, to the multiplicity Ω of the microscopic degrees of freedom [3]. This equation is based on the assumptions that (i) each microstate is equally likely for an isolated system with energy U, (ii) the systems will go through all the different microstates if given enough time (ergodic hypothesis). These two assumptions together state that the system will stay at each microstate for the same amount of time. The ergodic hypothesis has been very difficult to prove but treatment of the entropy from an information theory approach, as will be explained later in this chapter, seems to provide a powerful alternative view. Finally, (iii) a quantity's time average equals its ensemble average, where the ensemble average of a quantity A is:

$$\langle A \rangle = \sum_{\text{microstates} \equiv \text{i}} A_i p_i, \tag{1.2}$$

where i represents the particular microstate and p_i is the probability that the system is in microstate

i.

If the system is not isolated but is in contact with a heat bath at a temperature T, then the whole system (system of interest plus the heat bath) is in isolation (canonical ensemble). In this case the probability p_i for being in a microstate labeled i is given by the equation:

$$p_i = \frac{\exp(-\beta\varepsilon_i)}{Z(\beta)}, \tag{1.3}$$

where $\beta = 1/k_B T$, ε_i is the energy at level i and $Z(\beta)$ is the partition function given by:

$$Z(\beta) = \sum_i \exp(-\beta\varepsilon_i). \tag{1.4}$$

We can use the Helmholtz free energy $F(T, V, N)$ of the system instead of the entropy $S(U, V, N)$ based on the equation

$$F(T, V, N) = U(S, V, N) - TS(U, V, N) \tag{1.5}$$

and the Helmholtz free energy can be written in terms of the partition function as follows:

$$F = -k_B T \ln Z(\beta). \tag{1.6}$$

Derivatives of the Helmholtz free energy result in the quantities of interest, like the pressure P, entropy, etc.:

$$S = -(\frac{\partial F}{\partial T})_{V,N}, P = (-\frac{\partial F}{\partial V})_{T,N}, ... \tag{1.7}$$

The above formalism can be extended to account not only for energy exchange but also for particle exchange (grand canonical ensemble).

Using equilibrium statistical mechanics together with the laws of traditional thermodynamics can also give us the thermodynamic fluctuations of equilibrium systems. For example, if the energy E of a system is a fluctuating quantity, then deviations from its average value $< E >\equiv U$ are given by $(E - U)$, which is also a fluctuating quantity with $< E - U >= 0$. The mean square deviation (second central moment) is $< (E - U)^2 >$ and is given by:

$$< (E - U)^2 >= -k_B(\frac{\partial U}{\partial(1/T)})_{P/T,N_1,N_2,...} \tag{1.8}$$

In this overview we have presented some of the important concepts that connect thermodynamics to equilibrium statistical mechanics: entropy, via the Boltzmann equation, makes the key connection between macroscopic observables and the microscopic world of atomic coordinates. In addition, we have briefly stated the concepts of partition function, ensemble averages, fluctuations, etc., in the case of equilibrium system ensembles (for a detailed treatment see, for example, [4]). We have

presented these quantities here because we will see later in Section 1.2.2 and in Chapter 2 that we can also obtain such quantities using the Maximum Entropy approach. However, in the next section we move from the study of these different ensembles and the equilibrium case to the work that has been done to describe systems that are not in equilibrium but very close to it. With this next section we aim to give an overview of the effort that has been done in this area.

1.1.2 When Systems are Close to Equilibrium

Even though thermostatics and equilibrium statistical mechanics have given us insight into the various ensembles, it is still the dynamic processes that are particularly interesting. This is because dynamic processes occur and can describe the behavior of biological systems, economic systems, the weather, planetary motion, etc. When we extend traditional thermodynamics to include rates of processes, then we talk about "irreversible thermodynamics." This area of thermodynamics is still based on the laws stated at the beginning of this chapter but also includes the law of "time reversal symmetry." Based on this law, all laws of physics remain unchanged if the time t is replaced by its negative $-t$ and if, at the same time, the magnetic field B is replaced by its negative $-B$ [1]. Irreversible thermodynamics is based on the Onsager Reciprocity Theorem developed by Onsager in 1931. In this section we will briefly discuss the Onsager relations as well as other statistical theorems such as the fluctuation-dissipation theorem, the Green-Kubo relations and the linear response theory. We present this work here because, as will be stated along the way, the Information Theory/Maximum Entropy method has also been used to obtain the same results, while it has also been suggested that this formalism provides a more general treatment which, therefore, can be extended to describe systems arbitrarily far from equilibrium [5].

In order to talk about irreversible processes we need to include two parameters: the force (or affinity) that drives the process and the response to this process. Let's denote each of the affinities as $\mathcal{F}_i$. One of the extensive parameters of the system, X_k, will change as a response to the affinity. The flux, or the rate of change of the extensive parameter, is given by:

$$J_k \equiv \frac{dX_k}{dt}. \tag{1.9}$$

Furthermore, for systems that are purely resistive, that is, the fluxes at a given instant depend only on the values of the affinities at that instant (i.e., there is no memory in the system), we can define the kinetic coefficients L_{jk} as follows:

$$L_{jk} = (\frac{\partial J_k}{\partial \mathcal{F}_j})_0. \tag{1.10}$$

For the kinetic coefficients we adopt a notation that exhibits their functional dependence on an

externally applied magnetic field, **B**, suppressing the dependence on other intensive parameters: $L_{jk} = L_{jk}(\mathbf{B})$ and the Onsager theorem states that [6]

$$L_{jk}(\mathbf{B}) = L_{kj}(\text{-}\mathbf{B}). \tag{1.11}$$

This equation says that there is a symmetry between the linear effect of the jth affinity on the kth flux and the linear effect of the kth affinity on the jth flux, when these effects are measured in opposite magnetic fields. An example of this kind of interference between irreversible transport processes that take place simultaneously is in the case of heat conduction and electrical conduction: an electric current in a circuit of different metallic conductors will result in a heating of their junction. Similarly, if the junctions are at different temperatures, an electromotive force will appear in the circuit. The above equation describes the symmetry between these effects [6]. If the affinities are really small, then the "linear purely resistive process" holds, which is described mathematically by:

$$J_k = \sum_j L_{jk}\mathcal{F}_j. \tag{1.12}$$

It is important to note at this point that the Onsager reciprocal relations have been derived by E. T. Jaynes using the Maximum Entropy approach [7, 8], suggesting that this formalism can be used for systems that are not in equilibrium. A simple derivation of these Onsager relations by applying the same Maximum Entropy formalism to a simple model was described in [9]. This formalism will be discussed later in this chapter.

Let's assume that we perturb a system from equilibrium by disturbing one of its extensive parameters $\mathcal{A}_i$ by $\delta\mathcal{A}_i = \mathcal{A}_i - \langle\mathcal{A}_i\rangle$. The perturbation on the Hamiltonian will be of the form

$$\Delta H = -\sum_j f_j\mathcal{A}_j. \tag{1.13}$$

The f_js are called the external forces. Let $\overline{A_i}$ denote the average of the quantity $\mathcal{A}_i$, using the modified Hamiltonian (i.e., the modified probability distribution function), and $\langle\mathcal{A}_i\rangle$ denote the average using the equilibrium probability distribution function. Then, for small values of forces, it can be shown that [10]

$$\begin{aligned}\overline{\delta\mathcal{A}_j}(t) &= \overline{\mathcal{A}_i}(t) - \langle\mathcal{A}_i\rangle, \\ &= \beta\sum_j f_j\langle\delta\mathcal{A}_i(t)\delta\mathcal{A}_j(0)\rangle.\end{aligned} \tag{1.14}$$

The summand in Eq. 1.14, which is the flux autocorrelation function, is called the Kubo function

and is represented as,

$$C_{ij}(t) = \langle \delta\mathcal{A}_i(t)\delta\mathcal{A}_j(0)\rangle. \tag{1.15}$$

The most general version of the Kubo formula is:

$$\xi_{ij}(t) = -\beta\dot{C_{ij}}(t), \tag{1.16}$$

and connects the equilibrium fluctuations to the dynamic susceptibility, $\xi_{ij}(t)$. The dynamic susceptibility is defined by writing the most general formula for the dynamical linear response to an external time-dependent perturbation $\sum_j f_j(t)\mathcal{A}_j$:

$$\overline{\delta\mathcal{A}_i}(t) = \sum_j \int_{-\infty}^{t} dt'\xi_{ij}(t-t')f_j(t'). \tag{1.17}$$

This general Kubo formula can be used to give relationships between transport coefficients and the equilibrium correlation functions. These relationships are called Green-Kubo formulas and they have been used to describe, for example, electric conductivity and magnetic susceptibility. In the case of Ohm's law we know that the current I is proportional to the applied voltage V and σ is the proportionality constant (electric conductivity): $I = \sigma V$. However, this linear relationship is true only for small applied voltages and breaks down when V increases.

Based on causality, which requires that the effect must follow cause and, thus, $\xi(t) = 0$ if $t < 0$, we can get the fluctuation-dissipation theorem:

$$\xi''(\omega) = \frac{1}{2}\beta\omega C(\omega), \tag{1.18}$$

where $\xi''(t) = \frac{i}{2}\beta\dot{C}(t)$. In this equation, on the right-hand side we have the fluctuations while on the left the dissipation. This theorem is a powerful tool for predicting the behavior of a system that is close to equilibrium from its reversible fluctuations in thermal equilibrium. It is based on the assumption that the response of a system in thermodynamic equilibrium to a small applied force is the same as its response to a spontaneous fluctuation. Therefore, there is a direct relation between the fluctuation properties of the thermodynamic system and its linear response properties. Examples of the applications of the fluctuation-dissipation theorem include Brownian motion, which we briefly describe in the next section, and the thermal noise in a resistor. In the former case, the idea is that the random forces (fluctuations) which cause the erratic motion of a particle in Brownian motion would also cause drag (dissipation) if the particle were pulled through the fluid. In the case of the thermal noise in a resistor, the fluctuation-dissipation theorem is used to describe the electronic noise generated by the thermal agitation of the electrons inside an electrical conductor

at equilibrium, which happens regardless of any applied voltage.

In summary, the linear response theory and the fluctuation-dissipation theorem provide the microscopic basis of the Onsager relations, while the Green-Kubo relations give a method for obtaining the transport coefficients for non-equilibrium processes by observing the equilibrium fluctuations in the system. There are two main reasons why we have briefly presented this work: the first is in order to give an overview of the theoretical advancements that have been done for describing systems that are not in equilibrium, even though they are still very close to it. This work can explain the behavior of the measured quantities of an experimental system, like the current as a response to an applied voltage, the heating of a conductor, etc. The second reason for this presentation is because E. T. Jaynes has claimed to have reproduced these results via the use of the Maximum Entropy formalism, which can be extended to systems that are arbitrarily far from equilibrium [7, 8]. Before we proceed to the description of this formalism, we first briefly present Brownian motion and then an overview of the work on explaining systems that are arbitrarily far from equilibrium.

1.1.3 Brownian Motion

Since Chapter 3 deals with the diffusion of submicron particles, it is important at this point to briefly discuss Brownian motion. The name of the motion came from Robert Brown who observed the random motion of a pollen of grain that was suspended in water. The cause of this random motion was found by Einstein, who related the random jumps of the grain to the collisions it underwent with the water molecules, as a specific example of the fluctuation-dissipation theorem described above. He went on to explain the motion by introducing the theory of random walk. The Langevin equation gave an alternative description which led to the Fokker-Planck equation [11]:

$$\frac{\partial p(x,t)}{\partial t} = \frac{\partial^2 D p(x,t)}{\partial x^2} + \frac{\partial}{\partial x}\left(\frac{1}{\zeta}\frac{\partial U}{\partial x}\right), \tag{1.19}$$

where $p(x,t)$ is the probability distribution as a function of position x and time t, ζ is the friction coefficient of the fluid that the particle of mass m is suspended subject to a potential field $U(x)$. $D = k_B T/\zeta$ is the diffusion constant (T is the temperature of the fluid and k_B is the Boltzmann constant). This last relation for D is the Stokes-Einstein formula. What Eq. 1.19 describes is the time evolution of the non-equilibrium probability density function of position and velocity of a particle. The experiment we present in Chapter 3 is meant to observe the free expansion of a colloidal suspension of particles and obtain this probability density function. This is why in that chapter we will solve Eq. 1.19 for the case of $U(x) = 0$ and for the initial and boundary conditions of our experimental realization.

1.2 Putting the Dynamics back in Thermodynamics

So far we have discussed systems that are either in equilibrium or have been moderately perturbed from it. In the latter case, the systems are still very close to equilibrium. In this section we will discuss approaches for understanding processes that are arbitrarily far from equilibrium. We start with Jarzynski's equality (the non-equilibrium work theorem), a modification of which we will see in Chapter 3, and the fluctuation theorem and then we present the Information Theory and Maximum Entropy/Caliber method. As was discussed above, the goal is to give an overview of the work that has been done to describe non-equilibrium systems and compare that to the simplicity of the Maximum Entropy/Caliber formalism.

1.2.1 Jarzynski's Equality and the Fluctuation Theorem

In this section we present the work that was done in order to describe systems regardless of how close they are to equilibrium. Jarzynski's equality and the fluctuation theorem are two such results and they both deal with the Second Law of Thermodynamics. Jarzynski's equality relates free energy differences between two equilibrium states and non-equilibrium processes and is a strong statement about equilibrium thermodynamical quantities being exactly derivable from irreversible (non-equilibrium) trajectories. The fluctuation theorem gives the relative probability that the entropy of a system which is currently away from thermodynamic equilibrium (maximum entropy) will increase or decrease over a given amount of time. In other words, it provides us with a mathematical expression for the probability (due to fluctuations) that entropy will flow in a direction opposite to that dictated by the Second Law of Thermodynamics.

The setting for the Jarzynski equality is a system that is in equilibrium with a heat bath at temperature T. Let λ be the extensive parameter of the system. $\lambda = \lambda_1$, when the system is in equilibrium in state 1. If we vary λ as a function of time, from its initial value $\lambda(0) = \lambda_1$ to its final value at $t = t_f$, $\lambda(t_f) = \lambda_2$ with a prescribed time-history $\lambda(t)$ and repeat the same procedure N times, then the work obtained during the ith trial is W_i. The Jarzynski's equality [12, 13, 14] states that,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \exp\left(\frac{-W_i}{k_B T}\right) = \left\langle \exp\left(-\frac{W_i}{k_B T}\right)\right\rangle = \exp\left(-\frac{F_2 - F_1}{k_B T}\right). \tag{1.20}$$

F_1 and F_2 are the free energies of the system, when the values of λ are λ_1 and λ_2, respectively. This equality is true no matter what the time history of loading $\lambda(t)$ is and this is where the importance of Jarzynski's equality lies. $\lambda(t)$ can be rapid or slow and, therefore, does not limit the implementation of experiments and computations to processes that are slow enough to allow for quasi-static equilibria. Whereas traditional thermodynamics tells us that the work done to drive

the system from one state to another is always greater than or equal to the difference in free energy between the two states ($\langle W \rangle \geq \Delta F$), where the equality holds only when the system is driven reversibly, Jarzynski's equality says that we can get the free energy difference between two states by averaging a large number of irreversible experiments, and that our free energy estimates will not be an upper bound, they will be the actual free energy difference. The lack of limitations in $\lambda(t)$ is why Jarzynski's equality is often used to obtain the free energy landscapes of biological systems [15].

If we now assume the same system that is coupled to a heat bath of constant temperature is forced out of its equilibrium state by a time-dependent work process, then we can use the fluctuation theorems [16, 17] to describe the entropy production. The general form of the fluctuation theorems is [16],

$$\frac{P(+\sigma)}{P(-\sigma)} \simeq \exp(\tau\sigma), \tag{1.21}$$

where $P(+\sigma)$ is the probability of observing an entropy production rate, σ, measured over a trajectory of duration τ. Thus, using this theorem, we can find the probability of a fluctuation that will result in an entropy consumption (i.e., decrease). As an extensive quantity, entropy production depends on the size of the system. If the system is large, so is σ, and, therefore, the probability of an entropy consumption is negligible. Similarly for large observation times τ, as can be seen from Eq. 1.21. The Second Law of Thermodynamics is, therefore, valid for large systems and observation times and the fluctuation theorem can better describe small systems, such as biological systems, observed for short times. Thus, one implication is that small machines (such as nanomachines or even mitochondria in a cell) will spend part of their time actually running in a way opposite to what would be expected based on the Second Law of Thermodynamics where entropy increases. When the fluctuation theorem is combined with the central limit theorem, the Green-Kubo relations for linear transport coefficients described above can also be obtained when close to equilibrium. However, the fluctuation theorem applies to fluctuations far from equilibrium and is, therefore, more general than the Green-Kubo relations.

1.2.2 Information Theory and Maximum Caliber

Now that we have presented the work in the area of statistical mechanics for describing systems that are in equilibrium, close to it, or arbitrarily far from it, we move on to presenting an alternative formalism that can be used to describe all such systems. This formalism is based on an Information Theory/Maximum Entropy approach that was first developed by E. T. Jaynes [18]. A big part of this thesis is based on the use of this approach in order to describe dynamical laws of transport and on the realization of an experiment that aims to verify or falsify some of its predictions. In this section, we discuss this formalism in detail. In Chapters 2 and 3 we will use this method in detail

to derive those dynamical laws of transport and describe the experiment, respectively.

From the Information Theory perspective, we know that the entropy, S, of an equilibrium system involves the summation $S = -\sum_\nu p_\nu \ln p_\nu$ over all possible microstates ν (p_ν is the probability that the system is in microstate ν). With the Maximum Entropy method, we can look at entropy from this Information Theoretic approach and make the best inference about the probability distribution of the system given the information we have about the measurable, macroscopic quantities [5, 18]. Maximum Entropy is used for equilibrium systems but, as will be seen, it can also be extended to non-equilibrium systems. Thus, we can write the entropy for systems no matter how far they are from equilibrium.

The key idea is the following [18]: for systems far from equilibrium, instead of writing the entropy over the system's *microstates* we can write it over *microtrajectories*. This method is called the Caliber [7] and with the "principle of Maximum Caliber" [7] the probabilities of the microtrajectories are obtained by maximizing the caliber, subjected to the physical constraints [7, 8, 18].

Maximum Entropy methods for equilibrium systems are well understood [7, 8, 18, 19, 20, 21] and there are also some properties derived for non-equilibrium systems [22, 23, 24]. In the next section we discuss Maximum Entropy for equilibrium systems [5]. In the next chapter we will show, however, specific examples of how the Principle of Maximum Caliber can be used to develop models for the dynamical laws of transport (diffusion, heat flow, chemical kinetics, etc.).

1.2.2.1 Maximum Entropy For Systems in Equilibrium

As was discussed in Section 1.1.1, the study of equilibrium systems has thus far made use of the ergodic hypothesis. With Maximum Entropy, this hypothesis is not used and, as a result, the method can be extended to non-equilibrium systems.

We mentioned at the beginning of the chapter that the experimental macroscopic measurements, which give us the macrostate of the system (volume, temperature, etc.), cannot give us information about the microscopic atomic coordinates of the particles that make up the system. We also discussed that equilibrium statistical mechanics does exactly that: it connects the atomic temporal and spatial coordinates with the macroscopic observables and the laws of traditional thermodynamics. We also mentioned that the experimentally determined macrostate is a spatial and time average of the atomic coordinates. Therefore, a given macrostate can correspond to many microstates that give out the same average. Important questions now arise: Given the macrostate, what is the probability that the system will be in a specific microstate? How can we assign these probabilities to the microstates? Obtaining those probabilities will allow us to find the fluctuations of the macroscopic observables.

In order to assign these probabilities using Maximum Entropy, we start by assuming that some quantity x can take on the values $(x_1, x_2, ..., x_n)$, where n can be finite or infinite. Given the average values of functions $f_1(x), f_2(x), ..., f_m(x)$, where $m < n$, we would like to find the probability

assignment $p_i = p(x_i)$ ($p_i \geq 0$, since the individual probabilities cannot be negative), which satisfies the given data (i.e., constraints):

$$\sum_{i=1}^{n} p_i = 1 \tag{1.22}$$

$$\sum_{i=1}^{n} p_i f_k(x_i) = \langle f_k(x) \rangle = F_k \quad k = 1, 2, ..., m. \tag{1.23}$$

The first constraint is the normalization constraint where, since the system is in some microstate, all the probabilities have to sum up to 1. The second constraint tells us that the measured quantity is essentially the probability-weighted average of all the functions. From the above we see that the number of equations m is less than the total number of variables n. Thus, we are missing information and, therefore, there is no unique solution. To obtain the least biased solution given the constraints (Eqs. 1.22 and 1.23) we maximize the "information measure" $S_I = \sum_i p_i \ln p_i$ (Shannon entropy). This quantity is known to have the properties of consistency and uniqueness and, therefore, it is the correct measure of the "amount of uncertainty" in a probability distribution. Thus, in order to obtain the distribution p_i we maximize the information theory entropy

$$S_I = -\sum_{i=1}^{n} p_i \ln p_i. \tag{1.24}$$

subject to the constraints Eqs. 1.22 and 1.23. We use the method of Lagrange multipliers. The entropy subject to the constraints is:

$$S_I = -\sum_{i=1}^{n} p_i \ln p_i + \lambda_0 \sum_{i=1}^{n} p_i + \sum_{k=1}^{m} \lambda_k \sum_{i=1}^{n} f_k(x_i) p_i. \tag{1.25}$$

In order to maximize this quantity we differentiate and set the result equal to 0, i.e.,

$$\delta S_I = \sum_i (-1 - \ln p_i + \lambda_0 + \sum_{k=1}^{m} \lambda_k f_k(x_i)) \delta p_i = 0. \tag{1.26}$$

Thus the probability p_i is given by,

$$p_i = \exp(-\lambda_0 + 1) \exp(\sum_{k=1}^{m} \lambda_k f_k(x_i)). \tag{1.27}$$

The normalization condition described above implies that,

$$\exp(-1 + \lambda_0) = \frac{1}{\sum_{i=1}^{n} \exp(\sum_{k=1}^{m} \lambda_k f_k(x_i))}. \tag{1.28}$$

The denominator on the right hand side is the partition function which is denoted by $Z(\lambda_1, ..., \lambda_m)$.

Thus, the probability of the microstate p_i is

$$p_i = \frac{\exp(\sum_{k=1}^{m} \lambda_k f_k(x_i))}{Z(\{\lambda_k\})}. \tag{1.29}$$

Now that we have the probability of the microstate i, we can obtain the Lagrange multipliers by substituting the above equation into our original equation for the constraint:

$$\begin{aligned} \langle f_k \rangle &= \sum_{i=1}^{n} f_k(x_i) p_i, \\ &= \sum_{i} \underbrace{\frac{f_k(x_i) \exp(\sum_k \lambda_k f_k(x_i))}{\sum_j \exp(\sum_k \lambda_k f_k(x_j))}}_{Z(\lambda_1, ..., \lambda_m)}. \end{aligned} \tag{1.30}$$

The above equation can also be written as,

$$\langle f_k(x_i) \rangle = \frac{\partial}{\partial \lambda_i} \ln Z(\lambda_1, ..., \lambda_m), \tag{1.31}$$

where k goes from 1 to m. Solving these m equations will give us the Lagrange multipliers and the probabilities of the microstates. The variances of quantities are found to be

$$\Delta f_k^2 = \langle f_k^2 \rangle - \langle f_k \rangle^2 = \frac{\partial^2}{\partial \lambda_k^2} \ln Z. \tag{1.32}$$

This general derivation gave us a "least biased" distribution for a problem with missing information. As a specific example of this results, let's say that we are only given the average energy $\langle E \rangle$ of a system. We can obtain the probabilities by using Eq. 1.29. In this example, this equation will turn out to be the Boltzmann distribution, as was seen in Section 1.1.1, if the Lagrange multiplier λ is equal to $1/k_B T$. Thus, using Maximum Entropy we got the well-known results of equilibrium statistical mechanics without using the ergodic properties of the system. Also, at thermodynamic equilibrium, the information entropy S_I is the same as the experimental entropy S_E [5].

With this Information Theory approach the obtained probability p_i is a representation of our belief that the system is in microstate i. If we conduct another measurement in order to determine the macrostate of the system under the same constraints, we expect reproducibility in the macroscopic behavior. However, the underlying microstates are not reproducible. Again, the fact that the macroscopic measurement/behavior is reproducible shows that this measurement is a characteristic of each of the vast number of microstates compatible with those constraints: the microscopic details are irrelevant for the prediction of the macro quantities and are averaged out by the instruments conducting these experiments. Therefore, we can see that if the application of Jaynes procedure fails, that is because there are constraints that had not been taken into account.

Thus, we have seen that the Maximum Entropy approach can recover the known results of equilibrium statistical mechanics. By using the experimental constraints we can obtain the least biased probabilities of the system's microstates. So far, since we are in equilibrium, the constraints do not depend on time. However, if the constraints become time-dependent, then we can find the probabilities of microtrajectories, instead of microstates, that will generalize this method to non-equilibrium systems.

1.2.2.2 Maximum Caliber for Non-Equilibrium Systems

As has been mentioned, the application of the Maximum Caliber approach to simple, non-equilibrium systems will be discussed in Chapter 2. In this section we briefly discuss the Maximum Caliber formalism which aims to describe systems that are far from equilibrium.

By definition, in the equilibrium case the macroscopic observables (i.e., constraints) were time independent. In contrast, in order to study non-equilibrium systems, we need to include the fact that the observables are now functions of time: as the system evolves in order to reach the equilibrium state, our macroscopic measurements will produce results that change with time. Therefore, in order to move from Maximum Entropy to Maximum Caliber, we need to make the following adjustments: our treatment using Maximum Caliber will be the same as with Maximum Entropy with the exceptions that the constraints will now be time-dependent and we will have microtrajectories instead of microstates.

We follow the same notation as in the previous section but now the variable x is a function of time t. If we consider a finite duration of time T, the variable $x(t)$ will have values $x_1(t), ..., x_n(t)$ for t between 0 and T. The experimental constraints $f_1(t), ..., f_m(t)$ are also time-dependent and are given by,

$$\sum_{i=1}^{n} f_k(x_i(t)) = \langle f_k(x(t)) \rangle = f_k(t) \quad k = 1, ..., m \quad t \in [0, T]. \tag{1.33}$$

Eq. 1.33 tells us that the system evolves in the phase space x via individual microtrajectories $x_i(t)$. This microtrajectory $x_i(t)$ of duration T is represented by Γ and the entropy in this non-equilibrium case is written over the microtrajectories, Γ. Simply, Γ represents the time function $x_\Gamma(t)$ for t between 0 and T. The entropy can now be written as

$$S_I = -\sum_{\Gamma} p_\Gamma \ln p_\Gamma. \tag{1.34}$$

Just like before, we subject this entropy to the constraints given in Eq. 1.33. These constraints are for all values of time between 0 and T. Therefore, the corresponding Lagrange multipliers are also dependent on time. We take into account all the constraints and write the entropy over the

trajectories:

$$S_I = -\sum_{\Gamma} p_\Gamma \ln p_\Gamma + \lambda_0 \sum_{\Gamma} p_\Gamma + \sum_{\Gamma} \left(\sum_{k=1}^{m} \int_0^T \lambda_k(t) f_k^{(\Gamma)}(t) dt \right) p_\Gamma. \tag{1.35}$$

Again, just like in the equilibrium case, in order to maximize the above quantity we differentiate it and set it equal to 0:

$$\delta S_I = \sum_{\Gamma} (-1 - \ln p_\Gamma + \lambda_0 + \sum_{k=1}^{m} \int_0^T \lambda_k(t) f_k^{(\Gamma)}(t)) \delta p_\Gamma = 0. \tag{1.36}$$

Consequently, the probability of the trajectory is:

$$p_\Gamma = \frac{1}{Z(\lambda_1(t), ..., \lambda_m(t))} \exp \left(\sum_{k=1}^{m} \int_0^T \lambda_k(t) f_k^{(\Gamma)(t)} \right) \tag{1.37}$$

with

$$Z = \exp(1 - \lambda_0) = \sum_{\Gamma} \exp \left(\sum_{k=1}^{m} \int_0^T \lambda_k(t) f_k^{(\Gamma)}(t) \right). \tag{1.38}$$

which is now a partition functional. We use the expression for p_Γ from Eq. 1.37 to find that the trajectory entropy $-\sum_\Gamma p_\Gamma \ln p_\Gamma$ has a maximum value of [8]

$$S_I^{\max}(f_1(t), ..., f_m(t)) = Z + \lambda.A, \tag{1.39}$$

where we have defined $\lambda.A$ to be a short form for the expression

$$\lambda.A = \sum_{k=1}^{m} \int_0^T \lambda_k(t) f_k(t). \tag{1.40}$$

$S_I^{\max}(f_1(t), ..., f_m(t))$ is the time-dependent caliber of the system.

In his work, E. T. Jaynes presents the Maximum Caliber method by making use of quantum mechanical operators and derives the Kubo formulas for transport coefficients, as well as the Onsager reciprocal relations, but without making use of the assumptions that the system is close to equilibrium [7, 8, 18]. It is perhaps due to this allowance to get rid of all these assumptions that led Jaynes to hold such a strong opinion about the universality of the Principle of Maximum Caliber. The reason why this method is expected to work is the same as why the maximum entropy method works for equilibrium statistical mechanics. As was explained in the previous section, the macroscopic behavior is experimentally reproducible under the applied constraints, while the microscopic path is not, and, therefore, the macroscopic observable is characteristic of each of the vast number of microscopic paths compatible with those constraints. The microscopic details are irrelevant to

the prediction of macroscopic quantities. The method of Maximum Caliber manages to get rid of the irrelevant microscopic information.

1.2.2.3 Interpretation of the Lagrange Multipliers

In the Maximum Entropy and Maximum Caliber approaches we have been using the method of Lagrange multipliers in order to be able to find the probabilities of microstates and microtrajectories. However, what is the physical interpretation of the Lagrange multipliers? In our treatment thus far the macroscopic observables, or constraints on the system, were of the form

$$\langle A \rangle = \sum_{\nu} p_{\nu} A_{\nu} \tag{1.41}$$

and

$$\langle A(t) \rangle = \sum_{\Gamma} p_{\Gamma} A_{\Gamma} \tag{1.42}$$

when the system was in equilibrium and far from equilibrium, respectively. For the experimentalist conducting a measurement, however, very often it is not the constraints that can be controlled but the associated Lagrange multiplier. For example, we can measure the average energy $< E >$ by controlling the temperature T that the system is in (reminder that $\lambda = 1/k_B T$ and, thus, T is essentially the Lagrange multiplier). We cannot measure the average energy directly. In the non-equilibrium case, we do not control the flux but only the driving force. For example, in an electronic circuit, we cannot control the current I but only the applied voltage V. Therefore, we can see that the Lagrange multiplier, which is associated with the macroscopic constraint, is the driving force and the actual experimentally controlled quantity. This is exactly the meaning of the Lagrange multipliers: they represent the experimentally controlled variable that will lead to the observable and, therefore, the system constraint. However, even though the relation between the Lagrange multiplier and the average constraint is relatively easy to see in these examples, it is not always the case, especially when we look at non-equilibrium systems. For example, let's assume that we have charged particles in an externally applied, time-dependent electric field. In this case we could have as a constraint the average particle flux $\langle J(t) \rangle$. We assume that the Lagrange multiplier associated with this constraint is the driving force but it is not evident how this multiplier is associated with the experimentally applied field. As another example, in the case of a colloidal particle hopping between two energy wells A and B, the Lagrange multipliers are interpreted as transition probabilities [25]. The transition probabilities are dependent on the energy landscape of the two wells. This energy landscape is the experimentally controlled quantity, because it can be adjusted by tuning the laser that is responsible for forming the traps, and is also the "driving" force that causes the particle to

hop between wells. However, a general interpretation for the Lagrange multipliers in the case of non-equilibrium systems is yet to be found.

1.3 Biology and Non-Equilibrium Systems

So far we have discussed thermostatics and thermodynamics and the work that has been done in the area of statistical mechanics in order to describe systems that are not in equilibrium. However, the question still remains: Why are non-equilibrium systems interesting? Why do we need models that can explain, describe, and predict the behavior of non-equilibrium systems? Even though non-equilibrium systems provide us with a lot of interesting physics like, for example, in the case of the formation of spiral galaxies, one main application of non-equilibrium dynamics comes from the area of biology. Biological systems are almost by definition non-equilibrium systems. If the biological cell were in equilibrium, it would not be able to respond to the ever-changing environmental conditions and, therefore, it would not be functioning properly. As a result, it is important that we use non-equilibrium models to explain, describe, and predict the behavior of biological systems.

1.3.1 Cells, Proteins, and Ribosomes

The cell is very commonly known as the "fundamental unit of life" [26]. This is because it has the ability to respond to changing environmental conditions in order to sustain itself, grow, and replicate. Nothing smaller than a cell can be considered living. Even viruses need cells in order to be able to propagate their own genetic legacy. Cells of the same or different type also interact with one another in the case of multicellular organisms. In order for the cell to be able to successfully execute all these functions, which can change with time, it cannot be static. It has to be robust and time changing itself. Therefore, the living cell is a non-equilibrium system. An example of a dynamic process is shown in Figure 1.1 which depicts images of fertilized sea urchin eggs as they are dividing. In order for the cell to divide a lot of different dynamic processes have to take place, such as protein synthesis, chromosome replication, etc.

About 30% of the cell weight is due to macromolecules, which includes proteins, polysaccharides and DNA. Approximately half of the macromolecules are proteins and, in a typical eukaryotic cell, it is estimated that proteins make up approximately 20% of the cell weight [27]. Proteins determine not only the structure but also the type (i.e., function) and behavior of the cell, which is why they are very commonly referred to as the building blocks of cells [26]. Since all proteins are formed by the linear arrangement of 20 different types of monomers, called amino acids, it is the sequence of the amino acids that determines the protein and, consequently, its structure and function. Therefore, protein synthesis is a fundamental part of the life of the cell.

In a typical liver cell we find approximately 10^9 proteins [27], a number much smaller than

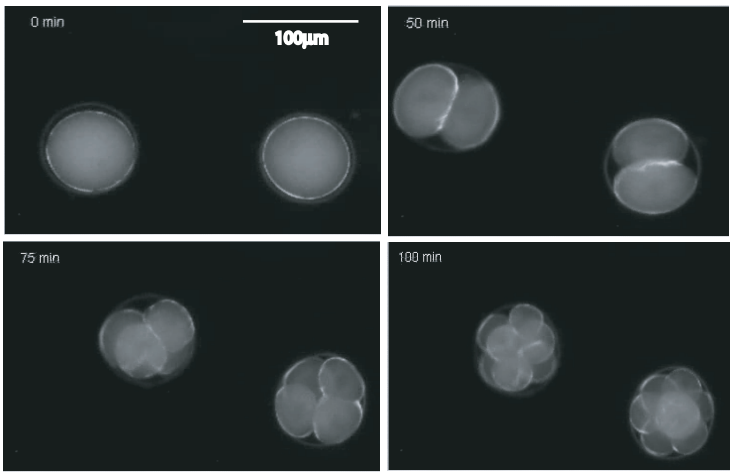


Figure 1.1: Division of two fertilized sea urchin eggs. This panel shows brightfield images during the process of division of two fertilized sea urchin eggs *in vitro*. The time point after the fertilization is shown in the upper left corner of each image.

Avogadro's number of 6×10^{23}, which is what physicists and chemists are most familiar with. In addition, DNA is found inside cells in a single copy. Thus, with respect to Avogadro's number, systems inside cells occur at small numbers. As will be explained in Chapters 2 and 3, this represents an example of a bigger problem, that of the behavior of systems (cells, experimental measurements, etc.) at the "small-numbers" limit.

The process of protein synthesis (i.e., translation) is executed by a structure inside the cell, the ribosome (see Figure 1.2a). When the cell is in need of more proteins, the ribosome self-assembles: the different parts that it consists of (ribosomal RNA and proteins) come together to form it. The process of self-assembly is not unique to ribosomes. Viral capsids (i.e., protein shells that contain and provide protection for the viral genome), for example, are also products of a self-assembly process. The question to ask, therefore, is why do these protein units (in the case of viral capsids) and RNA-protein complexes (in the case of the ribosome) come together, overcoming entropic and electrostatic forces, to form a highly symmetric capsid or a ribosomal subunit? The answer to this question will be examined in detail in Chapter 4. In the following sections of this chapter, however, we describe the function and importance of the ribosome and we give a historical overview of our knowledge surrounding its structure and function.

1.3.2 What is a Ribosome?

Ribosomes are complexes of ribosomal RNA (rRNA) and proteins that are found in all cells, either as free (suspended in the cytosol) or as bound to the rough endoplasmic reticulum or the nuclear envelope (in the case of eukaryotic cells). Both free and bound ribosomes are identical in structure

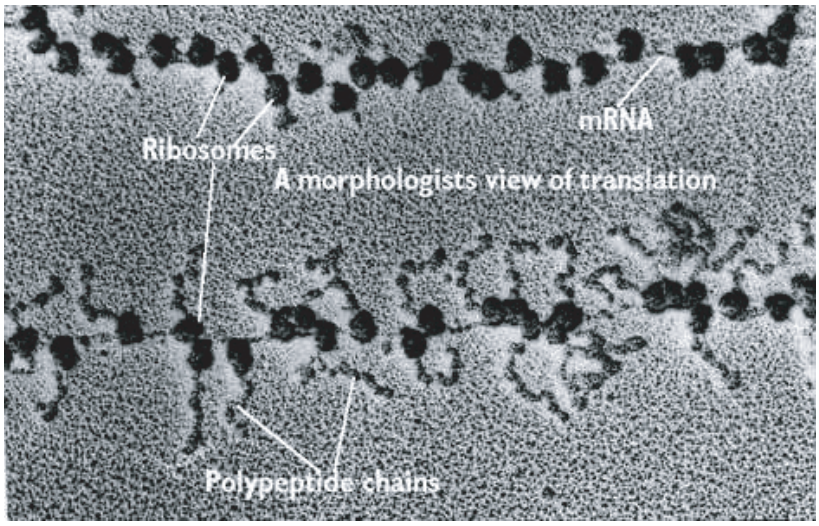
and function. As has already been mentioned, the ribosome is the macromolecular machine responsible for translation. In a bacterial cell there are approximately 20,000 ribosomes [28]. An *E.coli* ribosome is about 20 nm in diameter and has a mass of around 2700 kDa [29]. The way ribosomes synthesize proteins, as will be explained in detail in Chapter 4, is by binding the messenger RNA (mRNA) and, via a series of conformational changes, by moving along the mRNA as it binds the amino-acid-carrying transfer RNAs (tRNA). The ribosome then removes the amino acid from the tRNA and binds it to the growing protein chain (see Figure 1.2b). Ribosomes of both prokaryotes and eukaryotes consist of two subunits, the large and the small. For bacterial and archaeal ribosomes the large and small subunits are called 50S and 30S, respectively [30]. In the case of *E. coli* the 30S subunit contains one large rRNA (16S rRNA) molecule and 21 proteins. The 50S subunit consists of one large (23S rRNA) and one small (5S rRNA) rRNA molecule and 34 different proteins. Together these subunits are responsible for translating the messenger RNA (mRNA) and synthesizing the proteins.

1.3.3 Historical Overview

Ribosomes were first observed in the mid-1950s by George Palade who used an electron microscope [33]. Even though the identification of this cell structure happened quite early, the atomic structure of the ribosome has not been very easy to determine, as it has taken approximately 40 years. The first 2D shapes were provided with the use of transmission electron micrographs in the 1960s [34]. The shapes seen were quite crude and not of high resolution. Nevertheless, in the 1970s, image reconstitution techniques were developed that aimed to create 3D models [35, 36, 37, 38, 39] based on those images. At the same time, 2D gel electrophoresis provided the separation and identification of the individual ribosomal proteins [40] (see Figure 1.3), while immune electron microscopy, i.e., the visualization in electron micrographs of complexes formed between biological structures and specific antibodies, provided the sites of many of the ribosomal proteins (for a sample of the work, see [41, 42, 43, 44, 45, 46, 47]). The work of Masayasu Nomura established how the ribosomal subunits self-assemble via partial reconstitution experiments [48]. These experiments were based on the principle that, if one protein is left out from the mixture of proteins and rRNA then, as a result, certain proteins will not bind to the partially assembled subunit because of their interaction with the missing protein. Nomura's work was the first to provide some insight into the ribosomal self-assembly process, which is the main focus of Chapter 4.

The neutron scattering experiments that were conducted from the mid 1970s till the 1980s provided the distances between the centers of mass of the ribosomal proteins. Given this information about the separations between the various proteins, it was then possible to extract the 3D structure of the ribosome [49, 50, 51]. The development of cryoelectron microscopy was the key for determining the molecular structure of the prokaryotic ribosome. This took place in the 1990s mainly due to the

(a)



(b)

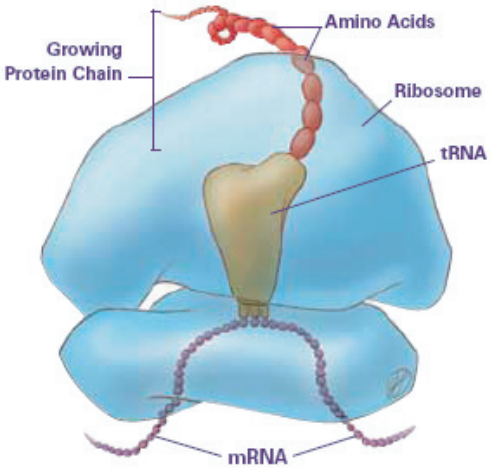


Figure 1.2: An overview of ribosomal protein synthesis. Figure (a) shows a group of ribosomes in action: they are connected by a strand of mRNA which they translate and synthesize the proteins, the chains of amino acids. In this figure, the growing peptide chain is radiating at right angles to the mRNA (taken from the University of Texas Medical Branch website [31]). Figure (b) is a cartoon of the translation process in terms of the various components that participate. Only one tRNA is shown but, in reality, there are 3 binding sites for the tRNAs that are all occupied during translation (see Figure 4.1). This figure is taken from the NIH website [32].

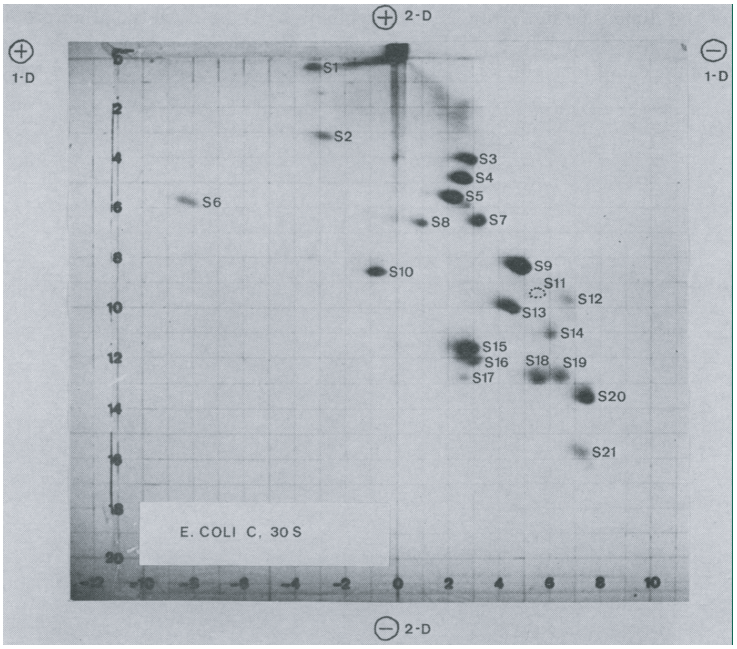


Figure 1.3: Two-dimensional gel electrophoresis of *E. coli* small ribosomal subunit proteins. First dimension (vertical): 8% acrylamide, pH 8.6; second dimension (horizontal): 18% acrylamide, pH 4.6. (taken from Kaltschmidt and Wittmann [40], which was the first to separate and number the proteins of the ribosomal subunits)

work of J. Frank who also determined the positions where the tRNA and the mRNA bind on the ribosome [52, 53, 54, 55, 56, 57, 58, 59, 60]. The resolution provided by cryoelectron microscopy was down to about 1 nm. In the years around 2000, the technology had advanced enough in the area of x-ray crystallography that allowed for the determination of the x-ray structure of several subunits at very high resolution [61, 62, 63, 64, 65]. Figure 1.4 presents the structure of the 70S ribosome of E. coli and that of the 30S subunit. The determination of the number of proteins and the molecular structure of the ribosome has provided some insight into the ribosomal function. However, now that the molecular structure of several ribosomes is well known, there are more questions that need to be answered. The mechanism of ribosomal self-assembly is one such question [66]. But why is it important that we study the ribosome and the process of self-assembly?

1.3.4 Importance of Studying Ribosomal Self-Assembly

From the historical overview described in the last section, it is clear that several decades were spent trying to establish ribosomal structure. The bacterial ribosome is highly complex and, consequently, the determination of the ribosomal structure at high resolution had to wait for the relevant technological developments to take place. However, why is it that the study of the ribosome started with the highly complicated 70S bacterial ribosome? The answer to this question lies in the fact that,

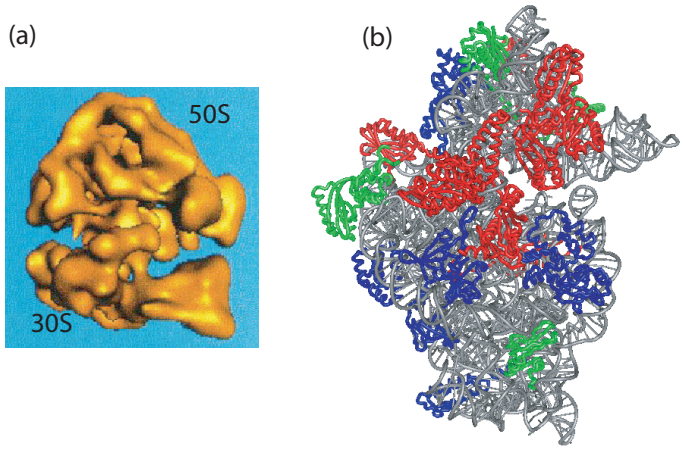


Figure 1.4: An overview of the ribosomal structure. Figure (a) shows a stereo image of the 23Å reconstruction of the *E. coli* 70S ribosome (taken from Stark et al. [55]). Figure (b) shows a ribbon diagram of the structure of the 30S subunit of the *E.coli* ribosome (taken from Wimberly et al. [63]).

in spite of its complexity, the bacterial ribosome, with the 3 rRNA molecules and the 55 proteins, is the simplest one to study [67]. The main function of the ribosome, protein synthesis, remains the same for all organisms and the conservation of ribosomal structure among lineages has implied that ribosomal structure has great evolutionary stability [68]. rRNA sequences have been conserved across various organisms and, in fact, it is this similarity that led to the discovery of the third domain of life, the Archaea [69]. The similarity of the ribosome across different organisms is, therefore, a great advantage in the sense that, our knowledge of the bacterial ribosome (function, behavior, self-assembly, etc.) can be carried across to the different organisms and give us insight into the organisms themselves.

In addition, studying the assembly of the ribosome is of great importance exactly because this cell structure is responsible for protein synthesis. Knowing the mechanism by which the ribosome assembles can be useful for the design of antibiotics. When bacterial cells infect an organism, they synthesize their own proteins and they grow and reproduce, so it is essential for antibiotics to stop this reproduction. A lot of antibiotics used in the treatment of bacterial infections target the process of bacterial protein synthesis via the ribosome [70, 30, 71, 72, 73]. Many bind to the 30S subunit and either hinder the decoding process or prevent the functional conformational changes of the ribosome [30]. Recently it has also been found that the aminoglycoside antibiotics inhibit not only the translation itself but also the formation of the small subunit [74]. Thus, in order for the antibiotics to target the ribosomal assembly, it is important to have an understanding of the interactions governing the assembly mechanism.

Understanding how the ribosome self-assembles is not only important for the design of antibiotics that can prevent bacteria from synthesizing their own proteins [70], but it will also give us infor-

mation about the RNA folding landscape and the RNA-protein interactions, which are central for understanding biological function and behavior. For example, it has been suggested that, in the case of large RNA molecules, the protein-dependent folding mechanisms are favored over the RNA independent mechanisms [75]. Consequently, the study of the ribosome will provide us with information about RNA and RNA-protein interactions that can be used for the study of other systems.

Macromolecular assembly is in itself a process of great interest. There are fundamental processes, such as transcription, DNA replication, cellular transport, viral capsid formation, etc., that are based on the self-assembly of macromolecular machines. Therefore, the study of ribosomal self-assembly will also give us an understanding about the process of self-assembly itself. The dynamic process of protein binding on the rRNA is central in Chapter 4, where we will present the literature in the topic and examine the rRNA folding and rRNA-protein binding of a fragment of the 30S ribosomal subunit as part of its self-assembly process.

1.4 Thesis Outline

It is clear at this point that the work presented here is divided into two parts, which are both under the umbrella of non-equilibrium dynamical systems. That is, the work of this thesis concentrates on systems that are not in equilibrium states by looking at measurable quantities that are dynamic, in the sense that they are changing with time. Such quantities include, for example, the flow and flux of matter and the forward and backward rates of a reaction, both of which result in concentrations (of particles or of reaction products) that are time dependent. In the chapters that follow, therefore, we have a theoretical and experimental evaluation of diffusion in the small-numbers limit and a determination of the reaction rates of the dynamic process of RNA-protein interactions. Below we discuss in more detail each of the chapters of this thesis.

In Chapter 2 we use E. T. Jaynes' principle of Maximum Caliber, which was described in detail in this chapter, in order to present a simple framework for the principles that underlie the dynamical laws of transport: Fick's law of diffusion, Fourier's law of heat flow, the Newtonian viscosity law, and the mass-action laws of chemical kinetics. In this chapter we make predictions about dynamical flux distributions and fluctuations associated with these laws of transport and we also introduce quantities, such as bad actors and potent trajectories, that are relevant when describing systems that contain small numbers of particles. The work in this chapter was in collaboration with K. Ghosh, M. Inamdar, K. Dill, and R. Phillips and has been published in Ghosh et al. [76].

In Chapter 3 we test the theoretical predictions about diffusion in the "small-numbers" limit that were derived in Chapter 2: we present the experiment conducted in order to verify or falsify the derived theoretical predictions. The experimental setup involves the use of microfluidics technology, as we observe the free expansion of a colloidal suspension inside a microfluidic channel. We also

present the image (i.e., data) acquisition and analysis process using Interactive Data Language (IDL) algorithms. We do some estimates to explore some of the experimental subtleties and ensure that the colloidal particles are indeed diffusing free of any interactions. The experimental error is also discussed, as well as the measures taken, both theoretical and experimental, in order to reduce and correct for it. Finally, we present all our experimental results for high, low, and medium concentrations of colloidal particles, as well as some results that were obtained using simulations. In the end we discuss the agreement between the experimental results and the theoretical predictions made at the beginning of the chapter. The work in this chapter was in collaboration with K. Ghosh, M. Inamdar, K. Dill, and R. Phillips and has been published in Seitaridou et al. [77].

Chapter 4 discusses the problem of the ribosomal self-assembly. We begin by reviewing the ribosome, its structure and importance for the proper function of the cell as well as the importance of the self-assembly process. We continue by presenting the background on the ribosomal assembly, the information we already know about the assembly of the 30S subunit as well as the dynamical information about the self-assembly process that is still unknown. The problem that we are interested in answering, namely finding the statistically weighted map of assembly, is explained in great detail. We discuss that in order to study the complete self-assembly process, automation is required. This automation will come via the use of microfluidics technology that we test by looking at a fragment of the assembly process, that of the folding of the 3-way junction of the central domain of the rRNA and the subsequent binding of the protein S15. This system has been extensively studied and, for this reason, presents a gold standard that we can use to test our experimental setup. The experimental goals and the way we address them are described. In addition to microfluidics, the experimental setup involves the use of two-photon spectroscopy and fluorescence resonance energy transfer, each of which is described in detail. We then present the theoretical background needed for analyzing the data and finding the reaction rates of the rRNA folding and S15 binding. We continue by presenting the obtained results and we compare them to the results that have been obtained using the more traditional techniques [78, 79, 80]. The work in this chapter was in collaboration with W. Ridgeway, J. Williamson, and R. Phillips and will be published in Ridgeway et al. [81].

1.5 The Use of Microfluidics in This Thesis

From Section 1.4 it can be seen that the experimental work presented in this thesis involves the use of microfluidic technology. Microfluidics refers to devices or flow configurations that have design features on the scale of a micron. The reason why microfluidics play an important role in this work is because of the many advantages they offer and their relation to the problems this thesis addresses. Some of the advantages of microfluidics include:

- Use of minimal amount of sample: Since the microfluidic channels have dimensions of the

order of microns, the resulting volumes are of the order of nanoliters. Consequently, this is the approximate volume of the samples used inside the microfluidic devices. This constitutes a significant improvement of three orders of magnitude over the volumes of samples used in traditional experiments that use plate readers or fluorimeters. In the case of the latter, the reactants are pipetted inside reaction tubes or plates via pipettes which require at least a microliter volume of sample. Therefore, with the use of microfluidics, the sample is not wasted: a lot more reactions can take place with a minimal sample volume.

- Automation: Just like in the case of traditional electronic circuits where switches have allowed for the use of logic and, consequently, for automation via programmable circuits, microfluidic devices offer the same capabilities. The microfluidic valves can be easily controlled by software and programs written by the user. This automation combined with the use of a small amount of sample results in a significant number of reactions taking place at a very small amount of time without demanding the user's time (other than the time it will take the user to write the program).

- User-friendly: Microfluidic devices are easy to fabricate. Several devices can be fabricated at once in the course of a day and, as a result, the troubleshooting and debugging process is very short. In addition, microfluidic devices are very easy to use since the devices themselves are transparent (so samples can be observed under the microscope) and the interface between the device and the user is such that the user has direct control over which samples are allowed into the device and when.

- Cost: Microfluidic devices are extremely inexpensive. There is an initial cost associated with creating the mold of the device, i.e., the cookie cutter that is used in order to create multiple copies of the device with the specific design, but the actual fabrication of the device does not cost more than a couple of dollars.

- Versatility: The devices are transparent and, therefore, can be used to look at biological samples under the microscope, just as if they were placed between a glass slide and a cover-slip (see Figure 1.5). In addition, the scaling down in size that microfluidics offers results in physics and physical phenomena that are considerably different from our macroscale everyday experiences. The physics of microfluidics resulted in a variety of different applications for this technology: such applications (with a reference example for each) are especially in the areas of chemistry, biology, and display technology where microfluidic devices have been used in order to find molecular diffusion coefficients [82], pH [83], and reaction kinetics [84]. Examples of other applications include capillary electrophoresis [85], immunoassays [86], flow cytometry [87], PCR amplification [88], protein crystallization [89], and cell manipulation [90]. For a review on the physics behind microfluidics and their applications see [91].

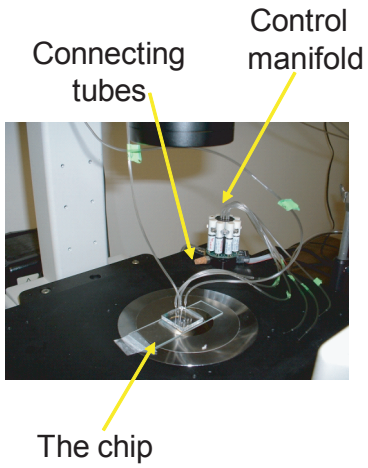


Figure 1.5: A microfluidics chip on top of a microscope. The picture shows the tubes used in order to load the sample and transfer them inside the microfluidic channels. The control manifold used to apply pressure to the control valves in order to open and close them is also shown. For details in the operation of microfluidic devices see Chapters 3 and 4.

The size of microfluidic devices (of the order of a micron) is comparable to the size of the biological cell. Therefore, they allow for the study of the "small-numbers" problem since it is possible to create a system of particles of the order of a hundred and then observe the system under the microscope. In addition, the physics of microfluidics and the automation this technology provides allow for the observation and measurement of many reactions as they are occurring in real time with the use of a minimal amount of sample. Thus, with the use of microfluidics it is easy to obtain the dynamics of a process, like self-assembly, that consists of many reactions each with its own rate. Consequently, we see that microfluidics can be useful for the study of non-equilibrium systems. It is important to note that the work of this thesis does not aim to improve microfluidics technology but only use it as a tool for the purposes of our experiments. Essentially, the problems discussed in this work and the experiments conducted to answer them also present applications for microfluidics.

Chapter 2

Using Maximum Caliber to Obtain Dynamical Laws of Particle Transport

[*Most of the work in this chapter appears in Ghosh et al. [76]*]

Recent advances in experimental biology have rendered the "small numbers" problem an important one to understand. Small particle numbers and their fluctuations are of great importance in nanotechnology and inside biological cells, which contain small numbers of proteins, at the order of a thousand [92]. In addition, there are now experimental techniques for studies at the single molecule level [93, 94, 95, 96, 97, 98, 15, 99] of ion channels, molecular motors, and in laser trap experiments [100, 101, 102] that have become possible partly due to advances in video microscopy, which have allowed the tracking of individual submicron particle trajectories [103, 104]. Examples of such single molecule experiments include the packing of DNA in viruses [97] and the force-induced folding and unfolding of DNA and RNA molecules [15, 99]. At the same time, it is important to note that the biological cell is a system that is away from equilibrium. This non-equilibrium state enables the flow of matter and energy and, therefore, maintains the cell at a stable and functional condition. [105] If the cell were in equilibrium, then nothing would ever happen.

Therefore, we see that many biological systems are systems away from equilibrium functioning at the small numbers limit. In this chapter we describe a simple framework for the principles that underlie the dynamical laws of transport: Fick's law of diffusion, Fourier's law of heat flow, the Newtonian viscosity law, and the mass-action laws of chemical kinetics, made by Ghosh et al. [76] using the "principle of Maximum Caliber", first described by E. T. Jaynes [7]. In the next chapter we will put the theoretical predictions on diffusion and Fick's law to the test by using a microfluidics set-up in order to look at diffusion of particles at the small-numbers limit. In addition to the experimental results directly related to the theory, some of the technical details that are related to the experiment will also be presented.

At this point, an argument might be posed against the theoretical and experimental results presented. Isn't diffusion already a well known and studied problem? What new results can be obtained with regards to this problem, both in the theoretical and experimental front that have not already been obtained using conventional non-equilibrium thinking and bulk diffusion experiments? Objections of such nature were addressed by Feynman in his Nobel lecture [106] who noted: "Theories of the known, which are described by different physical ideas may be equivalent in all their predictions and are hence scientifically indistinguishable. However, they are not psychologically identical when trying to move from that base into the unknown. For different views suggest different kinds of modifications which might be made and hence are not equivalent in the hypotheses one generates from them in one's attempt to understand what is not yet understood."

2.1 Theory of Maximum Caliber and Theoretical Predictions Based on Fick's Law

The "principle of Maximum Caliber" introduced by E. T. Jaynes provides the same type of foundation for systems of many degrees of freedom that the Second Law of Thermodynamics has provided for equilibrium systems. In order to illustrate this principle, Ghosh et al. [76] have used a slight variant of one of the oldest and simplest models in statistical mechanics, the Dog-Flea Model, or Two-Urn Model [107, 108]. Using this model Ghosh et al. examined dynamical fluctuations, as opposed to just average forces and flows, in systems with small numbers of particles. The focus is on microtrajectories as the basis for making predictions about dynamics.

We begin with the main principle embodied in Fick's Law. To keep it simple, we consider one-dimensional diffusion along a coordinate x. One statement of Fick's First Law is that the average particle flux is proportional to the average concentration gradient [3],

$$\langle J \rangle = -D \frac{\partial \langle c \rangle}{\partial x}, \tag{2.1}$$

where $\langle J \rangle$ is the observed macroscopic flux, D is the diffusion coefficient and $\langle c \rangle$ is the concentration of particles. We use brackets here, $\langle \ldots \rangle$, to make explicit that this phenomenological expression deals with averages over macroscopically large numbers of particles, and to indicate that the particle concentration and flux can be meaningfully represented as smooth functions of space and time in macroscopic systems. Fick's First Law is the basis for Fick's Second Law, also known as the diffusion equation. The combination of Fick's First Law with particle conservation:

$$\frac{\partial \langle c \rangle}{\partial t} = -\frac{\partial \langle J \rangle}{\partial x}, \tag{2.2}$$

gives the diffusion equation:

$$\frac{\partial \langle c \rangle}{\partial t} = D \frac{\partial^2 \langle c \rangle}{\partial x^2}. \tag{2.3}$$

Solving Eq. (2.3) subject to two boundary conditions and one initial condition gives both $\langle c(x,t) \rangle$, the average concentration in time and space, and the average flux $\langle J(x,t) \rangle$, when no other forces are present. The generalization to situations involving additional applied forces is the Smoluchowski equation [3]. These equations have been extensively verified in bulk gases and solutions with macroscopically large number of particles [109].

Fick's Law describes averages over a macroscopic number of particles; it does not describe small-number fluctuational quantities, such as $\langle J^2 \rangle - \langle J \rangle^2$, or any other aspect of the flux distribution function. Does Fick's Law hold in the limit of small numbers of particles? And, are there violations? That is, if Fick's Law predicts flow to the right, due to a concentration gradient sloping downward towards the right, does it ever happen that particles flow instead to the left? Such situations have been called "Second-Law violations" [110, 17] or, in classical thermal problems, they are expressed in terms of "Maxwell's Demon" [111]. Such fluctuations are, of course, not real violations of the Second Law, because the Second Law is only a statement about averages, not fluctuations [1]. Here we refer instead to such trajectories that go "against the grain" as *bad actors.* Our interest here is not just in average fluxes, but in the full flux distribution function.

To predict dynamical distributions of diffusion rates, we can use either classical random-flight theory or a maximum-entropy-like approach [76], called "Maximum Caliber" [7].

2.2 The Equilibrium Principle of Maximum Entropy

Because our strategy follows so closely the Jaynes derivation of the Boltzmann distribution law of equilibrium statistical mechanics [3, 5], we first show the equilibrium treatment. To derive the Boltzmann law, we start with a set of equilibrium microstates $i = 1, 2, 3, \ldots, N$ that are relevant to the problem at hand. Our aim is to calculate the probabilities p_i of these microstates in equilibrium. We define the entropy S of the system as

$$S(\{p_i\}) = -k_B \sum_{i=1}^{N} p_i \ln p_i, \tag{2.4}$$

where k_B is Boltzmann's constant. The equilibrium probabilities, $p_i = p_i^*$ are those values of p_i that cause the entropy to be a maximum, subject to two constraints:

$$\sum_{i=1}^{N} p_i = 1, \tag{2.5}$$

which is a normalization condition that insures that the probabilities p_i sum to one, and

$$\langle E \rangle = \sum_i p_i E_i, \tag{2.6}$$

which says that the energies, when averaged over all the microstates, sum to the macroscopically observable average energy.

By using Lagrange multipliers λ and β to enforce the first and second constraints, respectively, we obtain an expression for the values p_i^* that maximize the entropy [3, 5]:

$$\sum_i [-1 - \ln p_i^* - \lambda - \beta E_i] = 0. \tag{2.7}$$

The result is that

$$p_i^* = \frac{e^{-\beta E_i}}{Z}, \tag{2.8}$$

where Z is the partition function, defined by

$$Z = \sum_i e^{-\beta E_i}. \tag{2.9}$$

After a few thermodynamic arguments, the Lagrange multiplier β can be shown to be equal to $1/k_B T$ [5], This derivation, first given in this simple form by Jaynes,[5] identifies the probabilities that are both consistent with the observable average energy and that otherwise maximize the entropy. Jaynes justified this strategy on the grounds that it would be the best prediction that an observer could make, given the observable mean energy, if the observer is ignorant of all else. Although this derivation of the Boltzmann law is now popular, its interpretation as a method of prediction, rather than as a method of physics, is controversial. Nevertheless, for our purposes here, it does not matter whether we regard this approach as a description of physical systems or as a strategy for making predictions.

We switch from the principle of maximum entropy to the principle of maximum caliber [7]. In particular, rather than focusing on the probability distribution $p(E_i)$ for the various microstates, we seek $p[\{\sigma_i(t)\}]$, where $\sigma_i(t)$ is the ith microscopic trajectory of the system. Again we maximize an entropy-like quantity obtained from $p[\{\sigma_i(t)\}]$ to obtain the predicted distribution of microtrajectories. If there are no constraints, this maximization results in the prediction that all the possible microtrajectories are equally likely during the dynamical process. In contrast, certain microtrajectories will be favored if there are dynamical constraints, such as may be specified in terms of the average flux.

2.3 Fick's Law from the Dog-Flea Model

We want to determine the diffusive evolution of particles in a one-dimensional system. The key features of this system are revealed by considering two columns of particles separated by a plane, as shown in Figure 2.1. The left-hand column 1 has $N_1(t)$ particles at time t and the right-hand column 2 has $N_2(t)$ particles. This system is a simple variant of the famous "dog-flea" model the Ehrenfests introduced in 1907 [107, 108]. Column 1 corresponds to dog 1, which has N_1 fleas on its back at time t, and column 2 corresponds to dog 2, which has N_2 fleas at time t. In the time interval between time t and $t+\Delta t$, a flea can either stay on its current dog or jump to the other dog. This model has been used extensively to study the Boltzmann H-theorem and to understand how the time asymmetry of diffusion processes arises from the underlying time symmetry in the laws of motion [112, 107, 108]. We use this model for a slightly different purpose. In particular, our aim is to take a well-characterized problem like diffusion and to reveal how the principle of maximum caliber may be used in a concrete way.

First consider the equilibrium state of the dog-flea model. The total number of ways of partitioning the $(N_1 + N_2)$ fleas is

$$W(N_1, N_2) = \frac{(N_1 + N_2)!}{N_1! N_2!}. \tag{2.10}$$

The state of equilibrium is that for which the entropy, $S = k_B \ln W$ is a maximum. A simple calculation shows that the entropy is a maximum when the value $N_1 = N_1^*$ is as nearly equal to $N_2 = N_2^*$ as possible. In short, at equilibrium, both dogs will have approximately the same number of fleas, in the absence of any bias.

Our focus here is on how the system reaches equilibrium. We discretize time into a series of intervals Δt. We define a dynamical quantity p, which is the probability that a particle (flea) jumps from one column (dog) to the other in any time interval Δt. Thus, the probability that a flea stays on its dog during the time interval is $q = 1 - p$. We assume that p is independent of the time t and that all the fleas and jumps are independent of each other.

In equilibrium statistical mechanics, the focus is on the microstates. However, for dynamics we focus on *processes*, which at the microscopic level we call the microtrajectory. Characterizing the dynamics requires more than just information about the microstates; we must also consider the processes. Let m_1 represent the number of particles that jump from column 1 to 2 and m_2 represent the number of particles that jump from column 2 to 1 between time t and $t + \Delta t$. There are many possible values of m_1 and m_2: it is possible that no fleas will jump during the interval Δt, or that all the fleas will jump, or that the number of fleas jumping will be in between these limits. Each one of these different situations corresponds to a distinct microtrajectory of the system in this idealized dynamical model. We need a principle to tell us what number of fleas will jump during the time interval Δt at time t. Because the dynamics of this model is so simple, the implementation

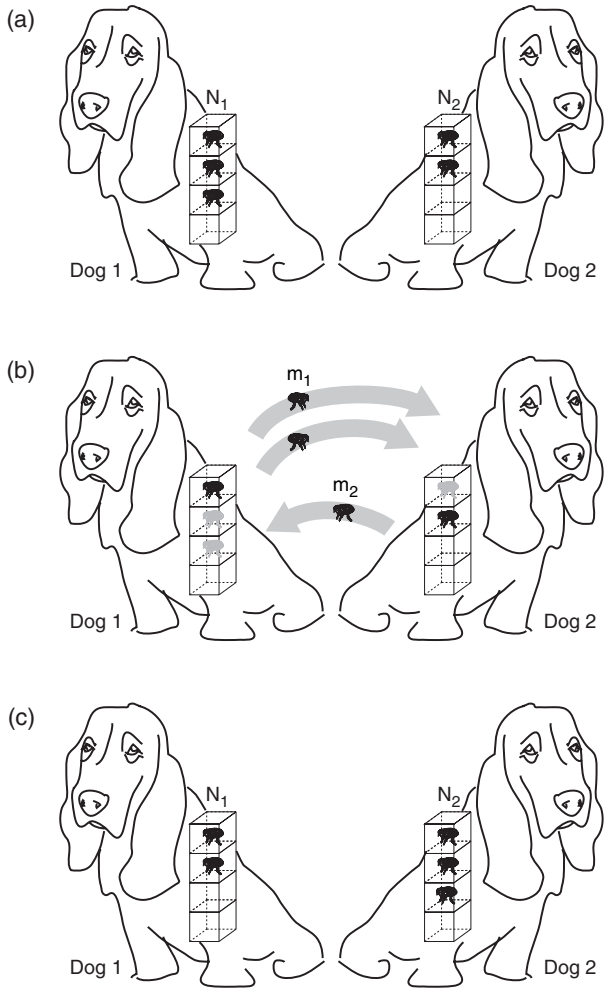


Figure 2.1: Schematic of the simple dog-flea model. (a) State of the system at time t; (b) a particular microtrajectory in which two fleas jump from the dog on the left and one flea jumps from the dog on the right; (c) occupancies of the dogs at time $t + \Delta t$

of the caliber idea is reduced to a simple exercise in enumeration and counting using the binomial distribution.

2.3.1 The Dynamical Principle of Maximum Caliber

The probability $W_d(m_1, m_2|N_1, N_2)$ that m_1 particles jump to the right and that m_2 particles jump to the left in a time interval Δt, given that there are $N_1(t)$ and $N_2(t)$ fleas on the dogs at time t, is

$$W_d(m_1, m_2|N_1(t), N_2(t)) = \underbrace{\left[p^{m_1} q^{N_1 - m_1} \frac{N_1!}{m_1!(N_1 - m_1)!}\right]}_{W_{d_1}} \underbrace{\left[p^{m_2} q^{N_2 - m_2} \frac{N_2!}{m_2!(N_2 - m_2)!}\right]}_{W_{d_2}}. \quad (2.11)$$

W_d counts microtrajectories in dynamics in the same spirit that W counts microstates for equilibrium. In the same spirit of the second law of thermodynamics, we maximize W_d over all the possible microtrajectories (that is, over m_1 and m_2) to predict the flux of fleas between the dogs. This maximization is the implementation of the principle of maximum caliber for this model. Maximizing W_d over all the possible processes (different values of m_1 and m_2) gives our prediction (right flux $m_1 = m_1^*$ and left flux $m_2 = m_2^*$) for the macroscopic flux that we should observe experimentally. (We follow the usual definition of flux, the number of particles transferred per unit time and per unit area. For simplicity, we take the cross-sectional area to be unity.)

Because the jumps of the fleas from each dog are independent, we find the predicted macroscopic dynamics by maximizing W_{d_1} and W_{d_2} separately, or for convenience their logarithms:

$$\frac{\partial \ln W_{d_i}}{\partial m_i}\Big|_{N, m_i = m_i^*} = 0, \quad (i = 1, 2). \tag{2.12}$$

The application of Stirling's approximation to Eq. (2.11) gives:

$$\ln W_{d_i} = m_i \ln p + (N_i - m_i) \ln q + N_i \ln N_i - m_i \ln m_i - (N_i - m_i) \ln(N_i - m_i). \tag{2.13}$$

We call $\mathcal{C} = \ln W_d$ the caliber. Maximizing $\mathcal{C}$ with respect to m gives

$$\frac{\partial \ln W_{d_i}}{\partial m_i} = \ln p - \ln q - \ln m_i^* + \ln(N_i - m_i^*) = 0. \tag{2.14}$$

This result may be simplified to yield

$$\ln\Big(\frac{m_i^*}{N_i - m_i^*}\Big) = \ln\Big(\frac{p}{1-p}\Big), \tag{2.15}$$

which implies that the most probable jump number is simply given by

$$m_i^* = pN_i. \tag{2.16}$$

Because our probability distribution W_d is nearly symmetric about the most probable value of flux, the average number and the most probable number are approximately the same. Hence, the average net flux to the right is

$$\langle J(t) \rangle = \frac{m_1^* - m_2^*}{\Delta t} = p\Big[\frac{N_1(t) - N_2(t)}{\Delta t}\Big] \approx -\frac{p\Delta x^2}{\Delta t}\frac{\Delta c(x,t)}{\Delta x}, \tag{2.17}$$

which is Fick's law for this simple model with the diffusion coefficient given by $D = p\Delta x^2/\Delta t$. We have rewritten $N_1 - N_2 = -\Delta c \Delta x$.

This approach gives us a simple explanation for why there is a net flux of particles diffusing across

a plane down a concentration gradient: more microscopic trajectories lead downhill than uphill. It shows that the diffusion coefficient D is a measure of the jump rate p. This model does not assume that the system is near-equilibrium, for example, it does not utilize the Boltzmann distribution law, and thus it indicates that Fick's law should also apply far from equilibrium. We might have imagined that for very steep gradients, Fick's law might have been only an approximation and that diffusion is more accurately represented as a series expansion of higher derivatives of the gradient. But at least for the present model, Fick's law is a general result that emerges from counting microtrajectories. On the other hand, we would expect Fick's law to break down when the particle density becomes so high that the particles start interacting with each other, thus spoiling the assumption of independent particle jumps.

2.3.2 Fluctuations in Diffusion

We have shown that the most probable number of fleas that jump from dog 1 to dog 2 between time t and $t+\Delta t$ is $m_1^* = pN_1(t)$. The model also tells us that sometimes we will have fewer fleas jumping during this time interval and sometimes we will have more fleas. These variations are a reflection of the fluctuations resulting from the system following different microscopic pathways.

We focus now on predicting the fluctuations. To illustrate, we first construct a table of W_d, the different numbers of possible microtrajectories, for all the values of m_1 and m_2. To keep the illustration simple, we consider the special case $N_1(t) = 4$ and $N_2(t) = 2$. We also assume $p = q = 1/2$. Table 2.1 lists the multiplicities of all the possible routes of flea flow. A given entry tells us how many microtrajectories correspond to the value of m_1 and m_2.

m_1 \ m_2	0	1	2
0	1	2	1
1	4	8	4
2	6	12	6
3	4	8	4
4	1	2	1

Table 2.1: Trajectory multiplicity for the case where $N_1(t) = 4$ and $N_2(t) = 2$. Each entry in the table corresponds to the total number of trajectories for the particular values of m_1 and m_2.

Table 2.1 confirms our previous discussion. The dynamical process for which W_d is a maximum (12 microtrajectories in this case) occurs when $m_1^* = pN_1 = 1/2{\times}4 = 2$, and $m_2^* = pN_2 = 1/2{\times}2 = 1$. You can calculate the probability of this flux by dividing $W_d = 12$ by the sum of entries in Table 2.1, which is $2^6 = 64$, the total number of microtrajectories. The result, which is the fraction of all the possible microtrajectories that have $m_1^* = 2$ and $m_2^* = 1$, is 0.18. We have chosen an example in which the particle numbers are very small, so the fluctuations are large; they account for more than 80% of the flow. In systems with large numbers of particles, the relative fluctuations are much

smaller.

Now look at the top right entry of Table 2.1. This entry says that there is a probability of 1/64 that both fleas on dog 2 will jump to the left while no fleas will jump to the right, implying that the net flux for this microtrajectory is backward relative to the concentration gradient. We call these "bad actor" microtrajectories. In these cases, particles flow to increase the concentration gradient, not decrease it. At time t, there are 4 fleas on the left dog, and 2 on the right. At the next instant, $t + \Delta t$, all 6 fleas are on the left dog, and no fleas are on the right-hand dog.

Similarly, if you look at the bottom left entry of Table 2.1, you see a case of superflux: a net flux of 4 particles to the right, whereas Fick's law predicts a net flow of only 2 particles to the right. Table 2.1 illustrates that Fick's law is only a description of the average or most probable flow and that Fick's law is not always exactly correct at the microscopic level. However, such violations of Fick's law are of low probability, a point that we will make more quantitative in the following. Such fluctuations have been experimentally measured in small systems [113].

$m_1 \backslash m_2$	0
0	1
1	6
2	15
3	20
4	15
5	6
6	1

Table 2.2: Trajectory multiplicity for $N_1(t) = 6$ and $N_2(t) = 0$ when the system is far from macroscopic equilibrium

We can elaborate on the nature of the fluctuations by defining the "potencies" of the microtrajectories. We define the potency to be the fraction of all the trajectories that lead to a substantial change in the macrostate. The potencies of trajectories depend on how far the system is from equilibrium. To see this, we continue our consideration of the simple system of 6 particles. The total number of microscopic trajectories available to this system at each instant in our discrete time picture is $2^6 = 64$. Suppose that at $t = 0$ all 6 of these particles are on dog 1. The total number of microscopic trajectories available to the system can be classified using m_1 and m_2, where in this case $m_2 = 0$ because there are no fleas on dog 2 (see Table 2.2).

What fraction of all microtrajectories changes the occupancies of both dogs by more than some threshold value, say $\Delta N_i > 1$? In this case, we find that 57 of the 64 microtrajectories cause a change greater than this value to the current state. We call these *potent* trajectories.

Now let us look at the potencies of the same system of 6 particles in a different situation, $N_1 = N_2 = 3$ when the system is in macroscopic equilibrium (see Table 2.3). In this case only the trajectories with (m_1, m_2) pairs given by (0,2), (0,3), (1,3), (2,0), (3,0), and (3,1) satisfy our

$m_1 \diagdown m_2$	0	1	2	3
0	1	3	3	1
1	3	9	9	3
2	3	9	9	3
3	1	3	3	1

Table 2.3: Trajectory multiplicity for $N_1(t) = 3$ and $N_2(t) = 3$ when the system is at macroscopic equilibrium

criterion. Summing over all of these outcomes shows that just 14 of the 64 trajectories are potent in this case.

There are two key observations conveyed by these arguments. For a system far from equilibrium the vast majority of trajectories at that time are potent and move the system significantly away from its current macrostate. Also when the system is near equilibrium, the vast majority of microtrajectories leave the macrostate unchanged. We now generalize from the tables to see when fluctuations will be important.

2.3.2.1 Fluctuations and Potencies

A simple way to characterize the magnitude of the fluctuations is to look at the width of the W_d distribution [3]. It is shown in standard texts that for a binomial distribution for which the mean and most probable value both equal $m_i^* = Np_i$, the variance is $\sigma_i^2 = N_i pq$. The variance characterizes the width. Moreover, if N_i is sufficiently large, a binomial distribution can be well-approximated by a Gaussian distribution

$$\mathcal{P}(m_i, N_i) = \frac{1}{\sqrt{2\pi N_i pq}} \exp\Big(- \frac{(m_i - N_i p)^2}{2N_i pq} \Big), \tag{2.18}$$

a convenient approximation because it leads to simple analytic results. However, this distribution function is not quite the one we want. We are interested in the distribution of flux, $P(J) = P(m_1 - m_2)$, not the distribution of right-jumps m_1 or left-jumps m_2 alone.

Due to a remarkable property of the Gaussian distribution, it is simple to compute the quantity we want. If you have two Gaussian distributions, one with mean $\langle x_1 \rangle$ and variance σ_1^2 and the other with mean $\langle x_2 \rangle$ and variance σ_2^2, then the distribution function, $P(x_1 - x_2)$ for the difference will also be a Gaussian distribution with mean $\langle x_1 \rangle - \langle x_2 \rangle$ and variance $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

For our binomial distributions, the means are $m_1^* = pN_1$ and $m_2^* = pN_2$ and the variances are $\sigma_1^2 = N_1 pq$ and $\sigma_2^2 = N_2 pq$. Hence the distribution of the net flux, $J = m_1 - m_2$ is

$$P(J) = \frac{1}{\sqrt{2\pi (pqN)}} \exp\Big(- \frac{(J - p(N_1 - N_2))^2}{2pqN} \Big), \tag{2.19}$$

where $N = N_1 + N_2$.

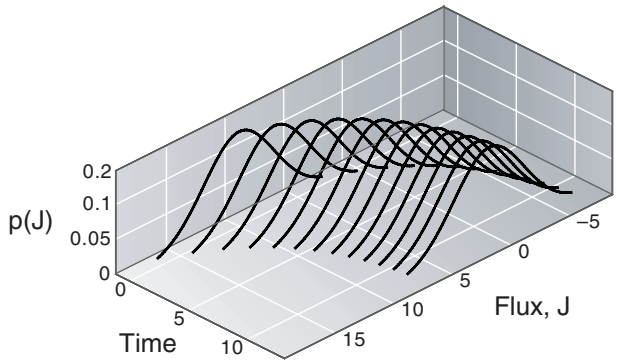


Figure 2.2: Schematic of the distribution of fluxes for different times as the system approaches equilibrium

Figure 2.2 shows an example of the distributions of fluxes at different times using $p = 0.1$ and starting from $N_1 = 100, N_2 = 0$. We update each time step using an averaging scheme, $N_1(t + \Delta t) = N_1(t) - N_1(t)p + N_2(t)p$. Figure 2.2 shows how the mean flux is large at first and decays toward equilibrium, $J = 0$. This result could also have been predicted from the diffusion equation. Equally interesting are the wings of the distributions, which show the deviations from the average flux; these deviations are not predictable from the diffusion equation.

One measure of the importance of the fluctuations is the ratio of the standard deviation σ to the mean, $\sqrt{\sigma^2}/J$. In the limit of large N_1, $\sqrt{\sigma^2/J}$ reduces to

$$\sqrt{\frac{\sigma^2}{J}} = \sqrt{\frac{Npq}{(N_1 - N_2)p}} \sim N^{-1/2}. \tag{2.20}$$

In a typical bulk experiment, the particle numbers are large, of the order of Avogadro's number 10^{23}. In such cases, the width of the flux distribution is exceedingly small, and it becomes overwhelmingly probable that the mean flux will be governed by Fick's law. However, within biological cells and in applications involving small numbers of particles, the variance of the flux can become significant. It has been observed that both rotary and translational single motor proteins sometimes transiently step backward relative to their main direction of motion [101].

As a measure of the fluctuations, we now calculate the variance in the flux. It follows from Eq. (2.19) that $\langle J^2 \rangle = Npq$, where $N = N_1 + N_2$. Thus, we can represent the magnitude of the fluctuations as

$$\delta = \sqrt{\frac{\langle (\Delta J)^2 \rangle}{\langle J \rangle^2}} = \frac{\sqrt{Npq}}{pfN} \propto \frac{1}{f}\sqrt{\frac{q}{p}N^{-1}}, \tag{2.21}$$

where $N = N_1 + N_2$ is the total number of fleas and $f = (N_1 - N_2)/(N)$ is the normalized concentration difference. The quantity δ is also a measure of the degree of backflux. In the limit of large N, δ goes to zero. That is, the noise diminishes with system size. However, even when N is large, δ can still be large (indicating the possibility of backflux) if the concentration gradient, $N_1 - N_2$, is

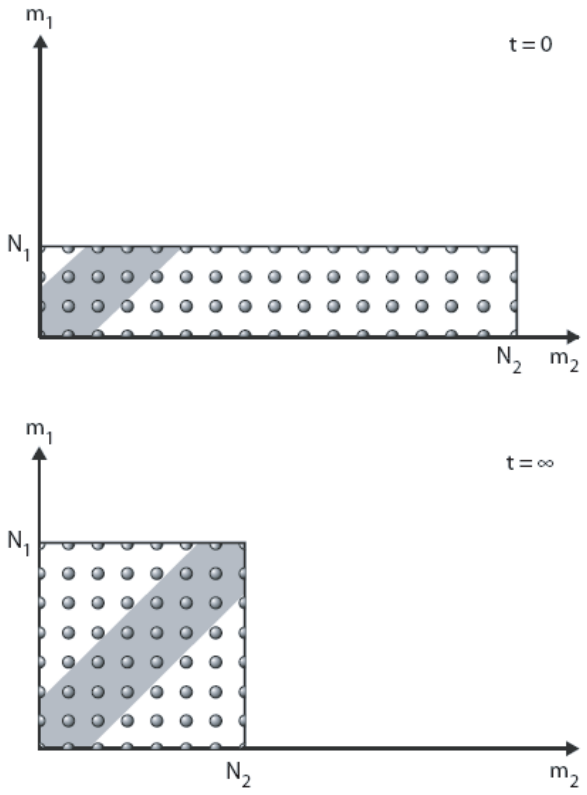


Figure 2.3: Schematic of which trajectories are potent and which are impotent. The shaded region corresponds to the impotent trajectories for which m_1 and m_2 are either equal or approximately equal and hence make relatively small change in the macrostate. The unshaded region corresponds to potent trajectories.

small.

Another measure of fluctuations is the potency. Trajectories that are not potent should have $|m_1 - m_2| \approx 0$ which corresponds to a negligible change in the current state of the system as a result of a given microtrajectory. In Figure 2.3 the impotent microtrajectories are shown as the shaded band for which $m_1 \approx m_2$. We define impotent trajectories as those for which $|m_1 - m_2| \leq h$ ($h \ll N$). In the Gaussian model, the fraction of impotent trajectories is

$$\Phi_{\text{impotent}} \approx \int_{-h}^{h} dJ \frac{1}{\sqrt{2\pi Npq}} \exp\left(\frac{-(J-(N_1-N_2)p)^2}{2Npq}\right) \tag{2.22}$$

$$= \frac{1}{2}\left(\text{erf}\left[\frac{h+(N_1-N_2)p}{\sqrt{2Npq}}\right] + \text{erf}\left[\frac{h-(N_1-N_2)p}{\sqrt{2Npq}}\right]\right) \tag{2.23}$$

and corresponds to summing over the subset of trajectories that have a small flux. To keep it simple, we take $p = q = 1/2$ for which the probability distribution for the microscopic flux $m_1 - m_2$ is given by Eq. (2.19). The choice of h is arbitrary and we choose h to be one standard deviation,

$\sqrt{N/4}$. Figure 2.4 shows the potencies for various values of N_1 and N_2. When the concentration gradient is large, most trajectories are potent, leading to a statistically significant change of the macrostate, whereas when the concentration gradient is small, most trajectories have little effect on the macrostate.

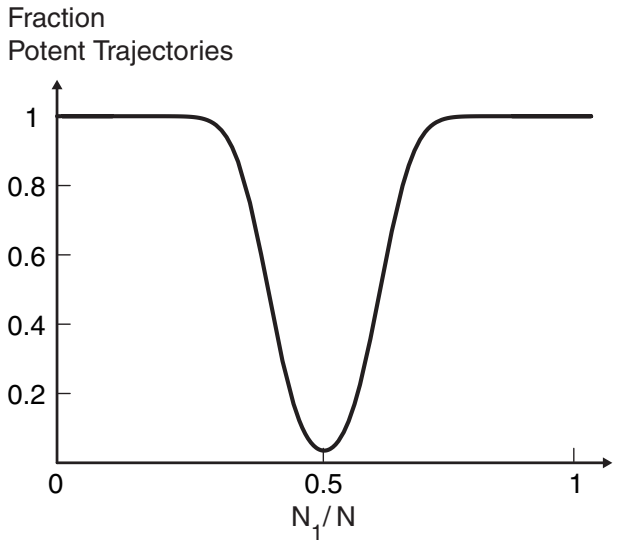


Figure 2.4: Illustration of the potency of the microtrajectories associated with different distributions of N particles on the two dogs. The total number of particles $N_1 + N_2 = N = 100$.

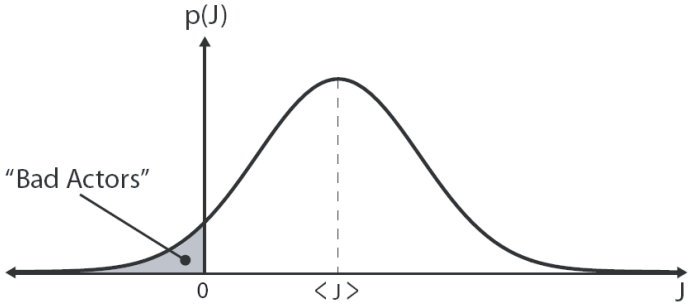


Figure 2.5: Illustration of the notion of bad actors. Bad actors are the microtrajectories that contribute net particle motion that has the opposite sign from the macroflux.

As another measure of the fluctuations, let us now consider the "bad actors" (see Fig. 2.5). If the average flux is in the direction from dog 1 to dog 2, what is the probability we will observe flux

in the opposite direction (bad actors)? We use Eq. (2.19) for $P(J)$ to obtain

$$\Phi_{\text{badactors}} \approx \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi Npq}} \exp\left(\frac{-(J-(N_1-N_2)p)^2}{2Npq}\right) \tag{2.24}$$

$$= \frac{1}{2}\left(1 - \text{erf}\left[\frac{(N_1-N_2)p}{\sqrt{2Npq}}\right]\right), \quad (N_2 > N_1) \tag{2.25}$$

which amounts to summing up the fraction of trajectories for which $J \leq 0$. Figure 2.6 shows the fraction of bad actors for $p = q = 1/2$. Bad actors are rare when the concentration gradient is large, and highest when the gradient is small. The discontinuity in the slope of the curve in Fig. 2.6 at $N_1/N = 1/2$ is a reflection of the fact that the mean flux abruptly changes sign at that value.

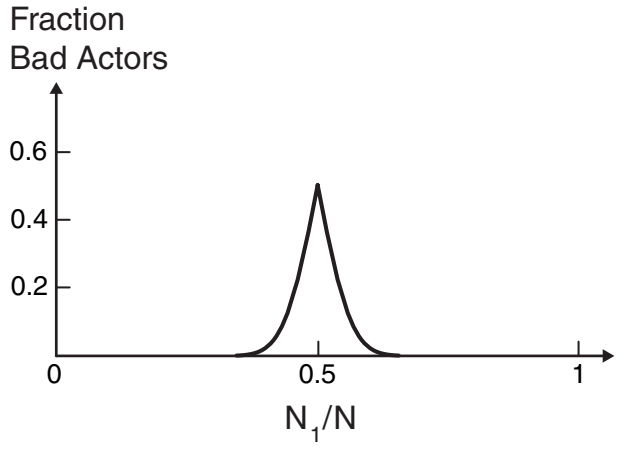


Figure 2.6: The fraction of all possible trajectories that go against the direction of the macroflux for $N = 100$. The fraction of bad actors is highest at $N_1 = N/2 = 50$.

2.4 Fourier's Law of Heat Flow

Although particle flow is driven by concentration gradients according to Fick's law, $\langle J \rangle = -D\partial c/\partial x$, energy flow is driven by temperature gradients according to Fourier's law [114]:

$$\langle J_q \rangle = -\kappa \frac{\partial T}{\partial x}. \tag{2.26}$$

Here, J_q is the energy transferred per unit time and per unit cross-sectional area and $\partial T/\partial x$ is the temperature gradient that drives it. The thermal conductivity [114] κ plays the role that the diffusion coefficient plays in Fick's law.

To explore a simplified version of Fourier's law that depends only on particle transport, we return to the dog-flea model as described in Section 2.3. Now columns 1 and 2 can differ not only in their particle numbers, $N_1(t)$ and $N_2(t)$, but also in their temperatures, $T_1(t)$ and $T_2(t)$. For simplicity,

we assume that each column is at thermal equilibrium and that each particle that jumps carries with it the average energy $\langle mv^2/2\rangle = k_BT/2$ from the column it left. We again take the cross-sectional area to be unity. In this simple model all energy is transported by hot or cold molecules switching dogs. The average heat flow at time t is

$$\langle J_q\rangle = \frac{m_1^*}{\Delta t}(k_BT_1/2) - \frac{m_2^*}{\Delta t}(k_BT_2/2) = \frac{pk_B}{\Delta t}[N_1T_1 - N_2T_2], \tag{2.27}$$

where m_1^* and m_2^* are, as defined in Section 2.3, the numbers of particles jumping from each column at time t. If the particle numbers are identical, $N_1 = N_2 = N/2$, then

$$\langle J_q\rangle = \frac{pk_BN}{\Delta t}(T_1 - T_2) = -\kappa\frac{\Delta T}{\Delta x}, \tag{2.28}$$

which is Fourier's law for the average heat flux for this two-column model. The model predicts that the thermal conductivity is $\kappa = (pk_BN\Delta x)/(\Delta t)$, which can be expressed in a more canonical form as $\kappa = pk_Bnv_{\rm av}\Delta x$ in terms of the particle density $n = N/\Delta x$ and the average velocity, $v_{\rm av} = \Delta x/\Delta t$. Our model gives a value for the thermal conductivity similar to that found in the kinetic theory of gases [114], $\kappa = (1/3)k_Bnv_{\rm av}\ell$, if Δx in our model corresponds to ℓ, the mean free path.

The factors of p and $1/3$ in our model and kinetic theory, respectively, can be reconciled by the following observation: Kinetic theory deals with the motion of particles in three dimensions so that each particle can move in 6 possible directions $(\pm x, \pm y, \pm z)$. As a result, only $1/6$ of the particles will contribute to the heat flux in our direction of interest. Also, on the plane of interest there are particles coming in from the positive and negative directions. So the factor of $1/6$ is increased by two to $1/3$. In contrast, in our model, the particles move with probability p or stay with probability $1 - p$. So the contribution to the flux comes from a fraction p of the particles. The origin of the numerical factors is thus clear.

The numerical factors of $1/3$ or p are not really important to the point we are trying to make. It is also not fair to read too much into this simple but illustrative model. The key point is that the simple model captures the main physical features of heat flow by appealing to the idea of summing over the weighted microtrajectories available to the system.

2.5 Newtonian Viscosity

Another phenomenological law of gradient-driven transport is that of Newtonian viscosity [114]

$$\tau = \eta\frac{dv_y}{dx}, \tag{2.29}$$

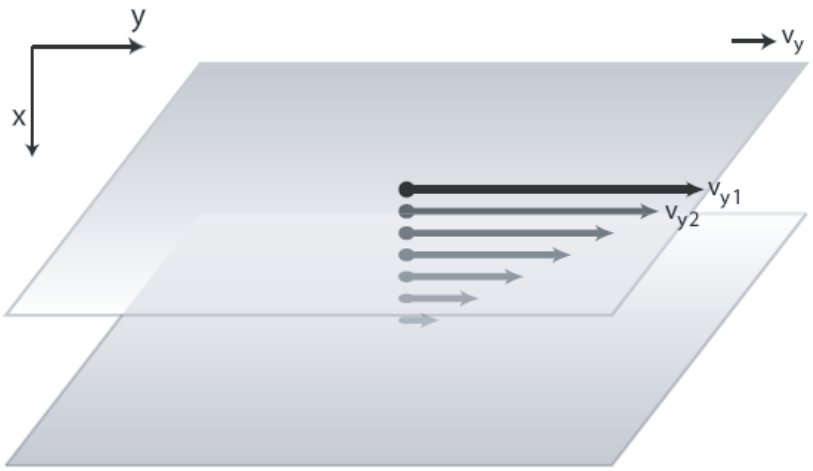


Figure 2.7: Illustration of Newton's law of viscosity. The fluid is sheared with a constant stress. The fluid velocity decreases continuously from its maximum value at the top of the fluid to zero at the bottom. There is thus a gradient in the velocity which can be related to the shear stress in the fluid.

where τ is the shear stress that is applied to a fluid, dv_y/dx is the resultant shear rate, and the proportionality coefficient η is the viscosity of a Newtonian fluid. Whereas Fick's law describes particle transport and Fourier's law describes energy transport, Eq. 2.29 describes the transport (in the x-direction, from the top moving plate toward the bottom fixed plate) of linear momentum in the y-direction (parallel to the plates) (see Figure 2.7). In the same spirit as our simplified treatment of Fourier's law, the dog-flea model can be used as the basis of a particle transfer version of momentum transport. Suppose each particle in column 1 of Figure 2.1 carries momentum mv_{y_1} along the y-axis, and m_1^* particles hop from column 1 to 2 at time t carrying with them some linear momentum. As before, we consider the simplest model for which every particle carries the same average momentum from the column it leaves to its destination column.

The flux J_p is the amount of y-axis momentum that is transported from one plane to the next in the x-direction per unit area:

$$\langle J_p \rangle = \frac{m_1^*}{\Delta t}(mv_{y_1}) - \frac{m_2^*}{\Delta t}(mv_{y_2}) = \frac{pm}{\Delta t}\left[N_1 v_{y_1} - N_2 v_{y_2}\right]. \tag{2.30}$$

If the number of particles is the same in both columns, $N/2 = N_1 = N_2$, Eq. 2.30 simplifies to

$$\langle J_p \rangle = \frac{pmN}{\Delta t}[v_{y_1} - v_{y_2}] = \eta \frac{\Delta v_y}{\Delta x}, \tag{2.31}$$

which is the Newtonian law of viscosity for this two-column model. The viscosity is predicted by this model to be $\eta = (pmN\Delta x)/(\Delta t)$. If we convert this result to the more canonical form, we have $\eta = pmnv_{\rm av}\Delta x$, where $n = N/\Delta x$ is the particle density, and $v_{\rm av} = \Delta x/\Delta t$ is the average velocity. This form is equivalent to the value given by the kinetic theory of gases [114], $\eta = (1/3)mn\ell v_{\rm av}$, if Δx from our model equals ℓ. The numerical factors of $1/3$ and p in the kinetic theory and our model, respectively, have the same origin as discussed in the context of heat flow. Note that this simple model based on molecular motions will clearly not be applicable to complex fluids where the underlying molecules possess internal structure.

2.6 Chemical Kinetics Within the Dog-Flea Model

Let us now look at chemical reactions using the dog-flea model. Chemical kinetics can be modeled using the dog-flea model when the fleas have preference for one dog over the other. Consider the reaction

$$A \underset{k_r}{\overset{k_f}{\rightleftharpoons}} B. \tag{2.32}$$

The time-dependent average concentrations, $[A](t)$ and $[B](t)$ are often described by chemical rate equations [3],

$$\frac{d[A]}{dt} = -k_f[A] + k_r[B] \tag{2.33a}$$

$$\frac{d[B]}{dt} = k_f[A] - k_r[B], \tag{2.33b}$$

where k_f is the average conversion rate of an A to a B, and k_r is the average conversion rate of a B to an A. These rate expressions describe only average rates; they do not give the distribution of rates. Some As will convert to Bs faster than the average rate $k_f[A]$ predicts, and some will convert more slowly. Again we use the dog-flea model and consider the average concentrations and the fluctuations in concentrations. A particularly fruitful area for applications of the $A \rightleftharpoons B$ dynamics considered here is to problems involving molecular motors and ion channels.

We assume that dog 1 represents chemical species A and dog 2 represents chemical species B. The net chemical flux from 1 to 2 is given by $J_c = m_1^* - m_2^*$. What is different about our model for these chemical processes than in our previous situations is that now the intrinsic jump rate from column 1 (species A), p_1, is different than the jump rate from column 2, p_2. This difference reflects the fact that a forward rate can differ from a backward rate in a chemical reaction. We assume that the fleas have a different escape rate from each dog. Fleas escape from dog 1 at rate p_1 and fleas escape from dog 2 at rate p_2. Maximizing W_d gives $m_1^* = N_1p_1$ and $m_2^* = N_2p_2$, so the average flux (which is almost the same as the most probable flux because of the approximately symmetric nature

of the binomial distribution) at time t is

$$\langle J \rangle = N_1 p_1 - N_2 p_2 = k_f[A] - k_r[B], \tag{2.34}$$

which is just the standard mass-action rate law, expressed in terms of the mean concentrations. The mean values satisfy detailed balance at equilibrium ($\langle J \rangle = 0$ implies that $N_2/N_1 = p_1/p_2 = k_f/k_r$).

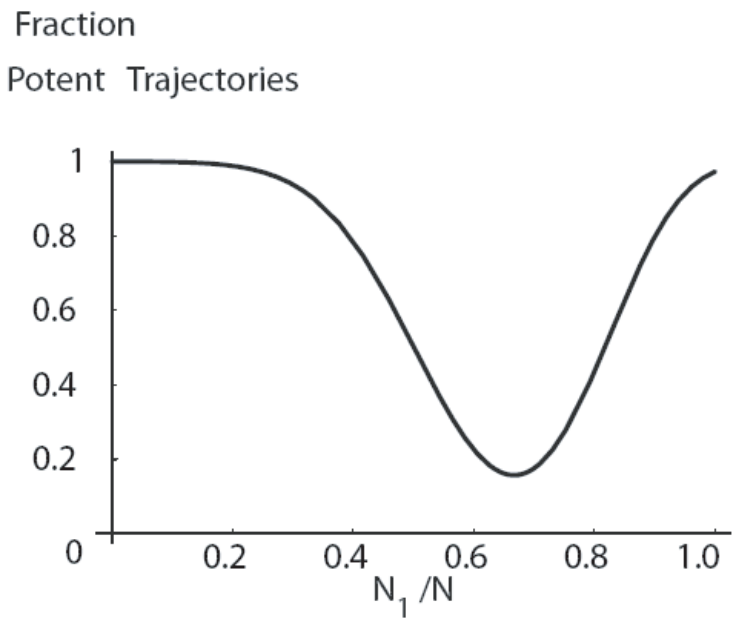


Figure 2.8: The fraction of potent trajectories Φ_{potent} as a function of N_1/N for $N_1 + N_2 = N = 100$, and $p_1 = 0.1$ and $p_2 = 0.2$. The minimum value of the potency does not occur at $N_1/N = 0.5$, but at $N_1/N = 0.66$. This value of N_1/N also corresponds to its equilibrium value given by $p_2/(p_1 + p_2)$.

More interesting than the behavior of the mean chemical reaction rate is the behavior of the fluctuations. For example, if the number of particles is small, then even when $k_f[A] - k_r[B] > 0$, indicating an average conversion of As to Bs, the reverse can happen occasionally. When will these fluctuations be large? As in Section 2.3, we first determine the probability distribution of the flux J. In this case, the probability distribution becomes:

$$P(J) = \frac{1}{\sqrt{2\pi(p_1 q_1 N_1 + p_2 q_2 N_2)}} \exp\Big(- \frac{(J - (N_1 p_1 - N_2 p_2))^2}{2(p_1 q_1 N_1 + p_2 q_2 N_2)} \Big). \tag{2.35}$$

We use this flux distribution function to consider the fluctuations in the chemical reaction. The relative variance in the flux is

$$\sqrt{\frac{\langle (\Delta J)^2 \rangle}{\langle J \rangle^2}} = \frac{\sqrt{N_1 p_1 q_1 + N_2 p_2 q_2}}{N_1 p_1 - N_2 p_2}. \tag{2.36}$$

As before, the main message is that when the system is not yet at equilibrium (that is, the denominator is nonzero), macroscopically large systems will have negligibly small fluctuations. The relative magnitude of the fluctuations scales approximately as $N^{-1/2}$. Let us also look at the potencies of microtrajectories as another window into the fluctuations. If we use Eq. (2.23) with p_1 and p_2, we find that the fraction of trajectories that are impotent is

$$\Phi_{\text{impotent}} \approx \int_{-h}^{h} dJ \frac{1}{\sqrt{2\pi(N_1 p_1 q_1 + N_2 p_2 q_2)}} \exp\left(\frac{-(J-(N_1 p_1 - N_2 p_2))^2}{2(N_1 p_1 q_1 + N_2 p_2 q_2)}\right) \tag{2.37}$$

$$= \frac{1}{2}\left(\text{erf}\left[\frac{h+(N_1 p_1 - N_2 p_2)}{\sqrt{2(N_1 p_1 q_1 + N_2 p_2 q_2)}}\right] + \text{erf}\left[\frac{h-(N_1 p_1 - N_2 p_2)}{\sqrt{2(N_1 p_1 q_1 + N_2 p_2 q_2)}}\right]\right). \tag{2.38}$$

For $N_1 + N_2 = N = 100$ and $p_1 = 0.1$ and $p_2 = 0.2$, $\Phi_{\text{potent}} = 1 - \Phi_{\text{impotent}}$ is shown in Fig. 2.8 as a function of N_1/N.

2.7 Derivation of the Dynamical Distribution Function from Maximum Caliber

We have used the binomial distribution function W_d as the basis for our treatment of stochastic dynamics. The maximum caliber assumption is that if we find the value of W_d that is a maximum with respect to the microscopic trajectories, this value will give the macroscopically observable flux. We now restate this assumption in more generally in terms of the probabilities of the trajectories.

Let $P(i)$ be the probability of a microtrajectory i during the interval from time t to $t + \Delta t$. A microtrajectory is a specific set of fleas that jump; for example microtrajectory $i = 27$ might be the case for which fleas 4, 8, and 23 jump from dog 1 to 2. We take as a constraint the average number of jumps, $\langle m \rangle$, the macroscopic observable. The quantity $m_i = 3$ in this case indicates that trajectory i involves 3 fleas jumping. We express the caliber $\mathcal{C}$ as

$$\mathcal{C} = \sum_i P(i) \ln P(i) - \lambda \sum_i m_i P(i) - \alpha \sum_i P(i), \tag{2.39}$$

where λ is the Lagrange multiplier that enforces the constraint of the average flux and α is the Lagrange multiplier that enforces the normalization condition that the $P(i)$s sum to one. Maximizing the caliber gives the populations of the microtrajectories,

$$P(i) = e^{-\alpha - \lambda m_i}. \tag{2.40}$$

Note that the probability $P(i)$ of the ith trajectory depends only on the total number m_i of the jumping fleas. Also, all trajectories with the same m_i have the same probabilities. In the same way

that it is sometimes useful in equilibrium statistical mechanics to switch from microstates to energy levels, we now express the population $P(i)$ of a given microtrajectory in terms of $\rho(m)$, the fraction of all the microtrajectories that involve m jumps during this time interval,

$$\rho(m) = g(m)Q(m), \tag{2.41}$$

where $g(m) = N!/[m!(N-m)!]$ is the density of trajectories with flux m (in analogy with the density of states for equilibrium systems), and $Q(m)$ is the probability $P(i)$ of microtrajectory i with $m_i = m$. In other words, i denotes a microtrajectory (a specific set of fleas jumping) and m denotes a microprocess (the number of fleas jumping). The total number of is associated with a given m is $g(m)$. It can also be easily seen that,

$$\sum_i P(i) = \sum_{m=0}^{N} g(m)Q(m) = \sum_{m=0}^{N} \rho(m) = 1 \tag{2.42}$$

$$\langle m \rangle = \sum_i m_i P(i) = \sum_{m=0}^{N} m g(m) Q(m) = \sum_{m=0}^{N} m\rho(m). \tag{2.43}$$

Thus the distribution of jump-processes written in terms of the jump number m is

$$\rho(m) = \frac{N!}{m!(N-m)!} e^{-\alpha} e^{-\lambda m}. \tag{2.44}$$

The Lagrange multiplier α can be eliminated by summing over all trajectories and requiring that $\sum_{m_i=0}^{N} \rho(m_i) = 1$, that is,

$$e^{\alpha} = \sum_{m}^{N} g(m) e^{-\lambda m} = \sum_{m}^{N} \frac{N!}{m!(N-m)!} e^{-\lambda m} = (1 + e^{-\lambda})^N. \tag{2.45}$$

We combine Eqs. (2.44) and (2.45) to obtain

$$\rho(m) = \frac{N!}{m!(N-m)!} \frac{e^{-\lambda m}}{(1+e^{-\lambda})^N}. \tag{2.46}$$

If we now let

$$p = \frac{e^{-\lambda}}{1+e^{-\lambda}}, \tag{2.47}$$

we obtain

$$p^m = \frac{e^{-\lambda m}}{[1+e^{-\lambda}]^m}, \tag{2.48}$$

and

$$(1-p)^{N-m} = \frac{1}{[1+e^{-\lambda}]^{N-m}}. \tag{2.49}$$

If we combine Eqs. (2.46), (2.48), and (2.49), we find the simple form

$$\rho(m) = \frac{N!}{m!(N-m)!} p^m (1-p)^{(N-m)}, \tag{2.50}$$

which appears in Eq. (2.11) and which we have used throughout this chapter.

2.8 Summary and Comments

We have shown how to derive the phenomenological laws of nonequilibrium transport, including Fick's law of diffusion, Fourier's law of heat conduction, the Newtonian law of viscosity, and the mass-action laws of chemical kinetics from a simple physical foundation. We used the dog-flea model for describing how particles, energy, or momentum can be transported across a plane. We combined this simple model with the principle of maximum caliber, a dynamical analog of the principle of maximum entropy for the laws of equilibrium. For dynamics we focus on microtrajectories rather than microstates and maximize a dynamical entropy-like quantity, subject to an average flux constraint. In this way maximizing the caliber is the dynamical equivalent of minimizing a free energy for predicting equilibria. A particular value of this approach is that it also gives us fluctuation information, not just averages. In diffusion, for example, sometimes the flux can be a little higher or lower than the average value expected from Fick's law. These fluctuations can be important for biology and nanotechnology, where the numbers of particles can be very small, and therefore where there can be significant fluctuations in rates around the average.

In the next chapter, we will test the theoretical predictions with regards to Fick's law of diffusion by devising an experiment that will observe the diffusion of particles in the small numbers limit.

Chapter 3

Measuring Flux Distributions for Diffusion in the Small-Numbers Limit

[*Most of the work in this chapter appears in Seitaridou et al. [77]. A big thank you goes to P. Wiggins for help with some of the figures and to M. Inamdar for help with the writing of this chapter which was part of the supplemental material of our paper.*]

For the classical diffusion of independent particles, Fick's Law gives a well-known relationship between the average flux and the average concentration gradient. What has not yet been explored experimentally, however, is the dynamical distribution of diffusion rates in the limit of small particle numbers. Here, we measure the distribution of diffusional fluxes using a microfluidics device filled with a colloidal suspension of a small number of microspheres and we compare these results to the predictions made in the previous chapter. Our experiments show that: (1) the flux distribution is accurately described by a Gaussian function; (2) Fick's Law—that the average flux is proportional to the particle gradient—holds even for particle gradients down to a single particle difference; (3) the variance in the flux is proportional to the sum of the particle numbers; and (4) there are backwards flows, where particles flow up a concentration gradient, rather than down it. In addition, in recent years, two key theorems about nonequilibrium systems have been introduced—Evans' Fluctuation Theorem for the distribution of entropies, and Jarzynski's work theorem. Here, we introduce a new fluctuation theorem, for the fluxes, and we find that it is confirmed quantitatively by our experiments.

3.1 Summary of Theoretical Predictions Based on Fick's Law

In order to proceed to the details regarding the experiment that will measure the dynamical distribution of diffusion rates and will test the prediction discussed in the last chapter, in this section we summarize those theoretical results that are relevant to this experiment.

To predict dynamical distributions of diffusion rates, we can use either classical random-flight theory or a maximum-entropy-like approach [76], called "Maximum Caliber" [7]. In short, if particles are independent, diffusing in one dimension, and if their jump rates are stationary in time, the distribution of particle fluxes, $P(J)$, at time t along an x-axis from bin 1 at x having N_1 particles, to an adjacent bin 2 at $x + \Delta x$ having N_2 particles, should follow the binomial distribution, or approximately a Gaussian function [76]. As we saw in the last chapter,

$$P(J) = \frac{1}{\sqrt{2\pi N p(1-p)/\Delta t^2}} \exp \frac{-(J - (N_1 - N_2)p/\Delta t)^2}{2Np(1-p)/\Delta t^2} \tag{3.1}$$

or

$$\begin{aligned} P(J) &= \frac{1}{\sqrt{2\pi \langle (\Delta J)^2 \rangle}} \exp\left(-\frac{(J - \langle J \rangle)^2}{2\langle (\Delta J)^2 \rangle} \right) \\ &= \frac{1}{\sqrt{2\pi (pqN)}} \exp\left(-\frac{(J - p\Delta N)^2}{2pqN} \right), \end{aligned} \tag{3.2}$$

where $\Delta N = N_1 - N_2$ is the difference in the number of particles associated with the two bins of interest, $N = N_1 + N_2$ is the total number of particles in these two bins, $\langle (\Delta J)^2 \rangle$ is the variance in the flux J, and $q = 1 - p$, with p being the probability that a particle jumps in the time interval Δt. Thus, the time step Δt provides the time during which all the action happens, i.e., the particle will move a distance Δx or stay put. The second line of Eq. 3.2 resulted from direct substitution of Eqs. 3.3 and3.4 explained in the next paragraph.

Various moments of the distribution function are readily obtained from this model. First, the model predicts that J, the average net number of particles that jump per unit time at time t is [76]

$$\langle J \rangle = \langle j_1 - j_2 \rangle = -\frac{p}{\Delta t} \Delta N, \tag{3.3}$$

where j_1 is the flux from bin 1 at x to bin 2 at $x + \Delta x$, and j_2 is from bin 2 to 1. This proportionality of the average flux $\langle J \rangle$ to ΔN simply predicts Fick's Law, where the diffusion coefficient D is related to p by $D = p\Delta x^2/\Delta t$, and where Δx is the bin size and Δt is the unit time step.

For the flux fluctuations, i.e., the second moment, the model predicts

$$\langle \Delta J^2 \rangle = \langle (J - \langle J \rangle)^2 \rangle = \frac{p(1-p)}{\Delta t^2} N. \tag{3.4}$$

Hence, the key prediction here is that the flux fluctuations are proportional to the total particle number, N. We now provide a quick derivation of the formula for the third moment, $\langle J^3 \rangle = 3\langle \Delta J^2 \rangle \langle J \rangle + \langle J \rangle^3$. The distribution for $P(J)$, as seen from Eq. 3.2, is symmetric about the mean value of flux $\langle J \rangle$. Thus, the odd moment $\langle (J - \langle J \rangle)^3 \rangle$ is equal to zero, $\langle (J - \langle J \rangle)^3 \rangle = 0$. Expansion

of the left-hand side of this equation and a couple of mathematical manipulations give:

$$\begin{aligned}
& \langle (J - \langle J \rangle)^3 \rangle = 0, \\
\implies \quad & \langle J^3 \rangle - \langle J \rangle^3 - 3\langle J^2 \rangle \langle J \rangle - 3\langle J \rangle^3 = 0, \\
\implies \quad & \langle J^3 \rangle - \langle J \rangle^3 - 3(\langle J^2 \rangle - \langle J \rangle^2)\langle J \rangle = 0, \\
\implies \quad & \langle J^3 \rangle - \langle J \rangle^3 - 3\langle \Delta J^2 \rangle \langle J \rangle = 0, \\
\implies \quad & \langle J^3 \rangle = 3\langle \Delta J^2 \rangle \langle J \rangle + \langle J \rangle^3, \qquad (3.5)
\end{aligned}$$

where we have made use of Eq. 3.4 in the second to last line.

The bad-actors are defined to be those trajectories which go against the direction of the macroscopic flux. To obtain the fraction of bad-actors we note (Ghosh et al. [76]) that,

$$\Phi_{\text{badactors}} \approx \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi N \frac{p(1-p)}{\Delta t}}} \exp\left(\frac{-(J - (N_1 - N_2)\frac{p}{\Delta t})^2}{2N\frac{p(1-p)}{\Delta t}} \right). \qquad (3.6)$$

If we define as positive the macroscopic flux, then the above equation amounts to summing up the fraction of trajectories for which $J \leq 0$. Evaluation of this integral gives:

$$\begin{aligned}
\Phi_{\text{badactors}} &= \frac{1}{2}\left(1 - \text{erf}\left[\frac{(N_1 - N_2)p}{\sqrt{2Np(1-p)}}\right]\right), N_2 > N_1 \qquad (3.7) \\
&= \frac{1}{2}\left(1 - \text{erf}\left(\frac{\langle J \rangle}{\sqrt{2\langle (\Delta J)^2 \rangle}}\right)\right), \qquad (3.8)
\end{aligned}$$

where, again, we have made use of Eqs. 3.3 and 3.4. For small values of $\frac{\langle J \rangle}{\sqrt{\langle (\Delta J)^2 \rangle}}$ this expression can be linearized with respect to its argument to give

$$\Phi_{\text{badactors}} \approx \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\langle J \rangle}{\sqrt{\langle (\Delta J)^2 \rangle}} \ldots + \mathcal{O}\left\{ \left(\frac{\langle J \rangle}{\sqrt{\langle (\Delta J)^2 \rangle}} \right)^3 \right\}. \qquad (3.9)$$

In the expression above, the next higher term (the cubic term), is an order of magnitude smaller than the linear term for the values of $\frac{\langle J \rangle}{\sqrt{\langle (\Delta J)^2 \rangle}}$ used in our experiments described below (see Figure 3.20 in Section 3.4)

3.1.1 A "Flux" Fluctuation Theorem

Recently, a useful description of nonequilibrium dynamics has involved fluctuation theorems. Fluctuation theorems characterize the extent to which the system deviates from its dominant flow behavior [17, 115, 116, 117, 118]. In the diffusive dynamics case of interest here, if the number of

particles N_1 in bin 1 is greater than the number of particles N_2 in bin 2, then particles, on average, will flow from 1 to 2. Fluctuation theorems describe the amount of reverse flow. Ours is a *flux* fluctuation theorem, i.e., it is expressed in terms of the quantity $P(J)/P(-J)$, where J is the flux. This differs from fluctuation theorems [115, 17] expressed in terms of entropies, $P(\Delta S)/P(-\Delta S)$ and from the work theorems of Jarzynski and Crooks, which are expressed in terms of the work w as $P[w/(kT)]/P[-w/(kT)]$ [16].

In our approach, the ratio $P(J)/P(-J)$, obtained from Eq. 3.1, gives the ratio of probabilities of fluxes in the forward and backward directions [76],

$$\begin{aligned}
\frac{P(J)}{P(-J)} &= \frac{\exp[-\frac{(J-\Delta Np/\Delta t)^2}{2Np(1-p)/\Delta t^2}]}{\exp[-\frac{(-J-\Delta Np/\Delta t)^2}{2Np(1-p)/\Delta t^2}]}, \qquad (3.10)\\
&= \exp[-\frac{(J-\Delta Np/\Delta t)^2-(-J-\Delta Np/\Delta t)^2}{2Np(1-p)/\Delta t^2}],\\
&= \exp[\frac{4J\Delta Np/\Delta t}{2Np(1-p)/\Delta t^2}],\\
&= \exp(\frac{2\Delta t}{1-p}\frac{\Delta N}{N}J). \qquad (3.11)
\end{aligned}$$

Taking the logarithm of $P(J)/P(-J)$ and substituting for the quantities $\langle J\rangle$ and $\langle \Delta J^2\rangle$, we find the fluctuation theorem for the dog-flea model to be,

$$\ln\frac{P(J)}{P(-J)} = \frac{2\langle J\rangle}{\langle(\Delta J)^2\rangle}J. \qquad (3.12)$$

Thus, the quantity $\ln[P(J)/P(-J)]$ is predicted to be proportional to the normalized flux $\langle J\rangle/\langle(\Delta J)^2\rangle\times J$. In situations having large flux, the back-flow becomes exponentially negligible.

3.2 The Experimental Setup and Preliminary Data Acquisition and Analysis

To study the dynamical distributions in diffusion and test the theoretical predictions, we devised a microfluidics experiment. Using a microfluidic chamber, we filled it with a colloidal suspension of particles, which executed Brownian motion under the microscope. The concentration of the colloidal particles inside the chamber was small enough so that we can observe diffusion in the small numbers limit. In what follows, we describe the design of our microfluidics chip, the setup associated with it, the data acquisition process, and some preliminary analysis.

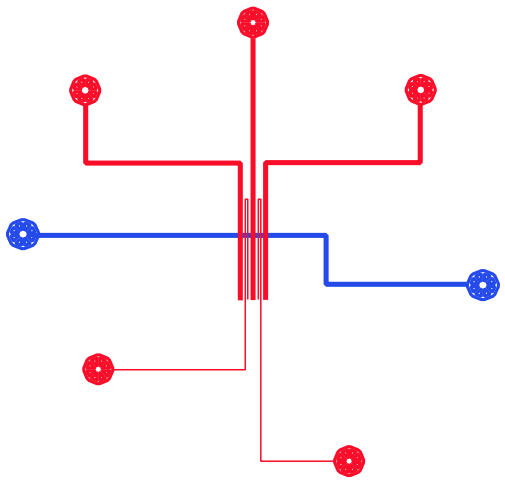


Figure 3.1: The microfluidics chip design. The microfluidics chip was designed in Adobe Illustrator. The flow layer is shown in blue while the control layer is shown in red. The circles at the end of each of the lines indicate where holes need to be punched. These holes are the communication ports between the chip and its user.

3.2.1 The Microfluidics Design

The idea was to create the colloidal suspension analog of the gas diffusion experiment in classical thermodynamics, where an isolated chamber is divided in two parts by a partition. One half of the chamber is filled with particles at a uniform concentration, while the other half is filled with solution but no particles. After the partition is removed, the particles are free to diffuse from one side of the chamber to the other. The experimental implementation involved the use of microfluidics, as the small volume associated with them allows the observation of systems with small numbers of particles. In addition, the chip fabrication allows for the observation to be done under the microscope. In Figure 3.1 we show the whole design that we created using Adobe Illustrator. Figure 3.2 is a zoom into the area of interest (i.e., the experimental chamber) and in the section that follows we describe each one of its elements. Our chips consisted of two layers, the flow layer and the control layer.

3.2.2 The Flow Layer

The flow layer is essentially one horizontal (flow) line, 100 μm in width, that contains the chamber of interest. In order for the flow line to fill with the solution of interest, pressure needs to be applied.

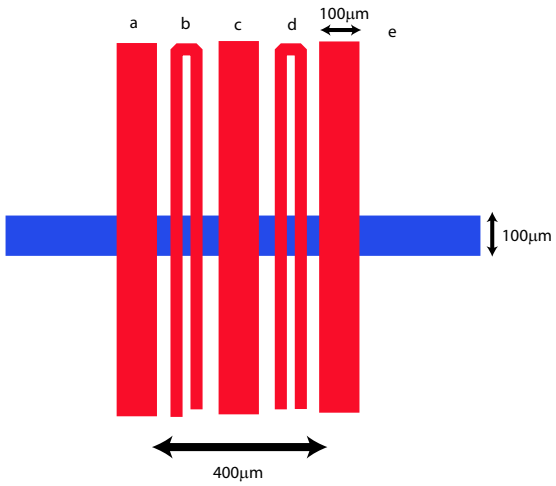


Figure 3.2: Zooming in the microfluidics design. The flow layer is drawn in blue while the control layer is in red. The outer control valves (a and e) are used for the isolation of the chamber while the middle one (valve c) is used as the partition. Valves b and d are used as heat bath, in order to maintain the room temperature of the suspension underneath.

In our case, the air pressure used to force the suspension of colloidal particles into the chamber was controlled via an air pressure gauge and regulator and was approximately 4.5 psi.

3.2.3 The Control Layer

As its name reveals, the control layer is used to control the flow of the samples that fill the flow layer. In our case, we are using the push-down configuration, where the control layer is located on top of the flow layer. When pressure is applied to the control valves, the valves expand and, thus, collapse down onto the flow line, which closes. In order for this to occur, the control valves have to be of the same width as the flow line that they need to close. If their width is smaller than that, then these control valves will have no effect on the opening/closing of the flow line below. In order to close the control valves, we used a separate air pressure gauge and regulator and adjusted the air pressure to approximately 12 psi.

The chamber of the flow layer is isolated from the rest of the flow line via two outer control valves (valves a and e in Figure 3.2), which, for the reason explained above, also have to be 100 μm wide. These lines are spaced 400 μm apart. There is also a middle control valve (valve c), which plays the role of the partition. The smaller control valves (valves b and d, 20 μm wide) that are in between the larger control valves and directly above the two halves of the chamber serve the purpose of a heat bath. These lines are filled with water which is at room temperature and are pressurized in order to ensure that the water in them will not evaporate. Since the duration of the

experiment is long (more than 6 hours), the water that fills these smaller control valves maintains the colloidal suspension beneath them at the same constant room temperature. If the heat bath is not there, then, during the experiment, the colloidal suspension slowly evaporates/dries out (the PDMS material the chips are made of is permeable to gas. Therefore, molecules from the colloidal suspension that jump into the gas phase can easily escape out of the chip). The volume sample is, thus, continuously reduced which, in turn, causes the middle control valve to collapse (the chamber is isolated and, therefore, the empty space caused by the evaporation has to be occupied by the control valves that are directly above it). The heat bath valves are always under a pressure of 12 psi during the experiment. Therefore, in the section that follows we only describe the operation of the larger control valves (valves a, c, and e).

3.2.4 Chip Fabrication and Operation

Using the techniques of soft lithography, chip fabrication [119], and the Sylgard 184 Silicone elastomer kit (Dow Corning Corporation), we fabricated a chip as follows:

Soft Lithography: For the fabrication of the two molds, one for the flow and one for the control layers, the silicon wafers were treated as follows: The silicon wafers were spun with photoresist 5740 at 2000 rpm for 60 s and were soft baked at 105°C for 90 s. The photoresist was then exposed to UV light for 45 s and was developed using developer 2401 at $1:5$ dilution. The wafers were then rinsed with H_2O and dried with N_2. In order to round the flow channels the wafer for the flow layer was hard baked at 125°C for 20 min. The final wafers (i.e. mold) had channels with a height of approximately 10 μm, as determined by a profilometer.

Chip fabrication: Using the Sylgard 184 Silicon Elastomer kit (PDMS parts A and B from the supplier) we fabricated the push-down configuration chip as follows:

- Flow Layer: We prepared a mixture of $A:B$ in $20:1$ (15 g part A and 0.75 g part B). The mixture was then poured on the flow layer mold and spun at 2500 rpm for 1 min (acceleration at 15 s). The wafer was then placed in an 80°C oven for 20 min to harden.

- Control Layer: We prepared a mixture of $A:B$ in $5:1$ (25 g part A and 5 g part B). The mixture was then poured on the control layer mold and was placed in a vacuum chamber for approximately 30 min to degas it and remove any air bubbles from the mixture. The wafer was then placed in an 80°C oven for 20 min to harden.

After removing both molds from the oven, we peeled off the hardened PDMS from the control layer mold. Holes were punched at the input ports of the control layer, which was then aligned on top of the flow layer mold that still had the thin layer of hardened PDMS on it. The aligned chips were now incubated at 80°C oven for 1.5 hours to bind. After peeling off the chips from the flow layer mold, we punched the holes at the flow layer input ports, placed the chips on top of No 1 cover

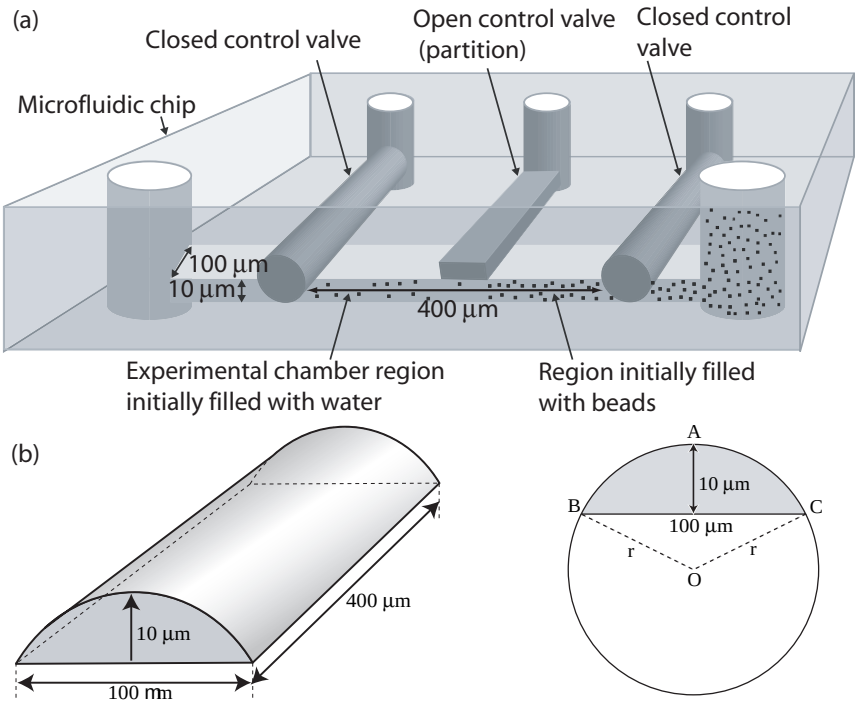


Figure 3.3: The microfluidics experiment. Colloids corralled on one side of a gate begin to diffuse at time $t = 0$ when the gate is opened. (a) Schematic of the microfluidic chip (see text for details). (b) The geometry of the microfluidic chamber (not drawn to scale)

slips and incubated them overnight in the 80°C oven in order for the PDMS (i.e., chips) to bond to the cover slip. After removing the chips from the oven the next day, they were ready to be used.

Based on the design and the above fabrication method, we made a microfluidics chamber having approximate dimensions 400 μm by 100 μm, partitioned into two regions, as described above (see Figure 3.3a for an additional cartoon of the chip. In this cartoon the heat bath valves are omitted to reduce cluttering). Due to the fabrication process, the cross section of this chamber is a segment of a circular disc, with a maximum depth of 10 μm (see Figure 3.3b). With only the middle control valve (valve c) closed so that no mixing can occur, one side of the flow line is filled with water while the other is filled with a solution containing about 200 colloidal, green fluorescent polystyrene particles 0.29 μm in diameter (Duke Scientific, Cat. No. G300) (see Figure 3.3a). As will be explained in detail below, the beads are at an optimized concentration so that the interactions are negligible while at the same time permitting sufficient statistics over a wider range of ΔN and N (also see Sections 3.3.6 and 3.3.8). When the flow line is completely filled, we close the outer control valves (a and e) in order to isolate the chamber from the rest of the flow line. After that happens, we open the middle control line, i.e., the partition is removed. We define the moment when the partition is removed as time $t = 0$.

3.2.5 Data Acquisition

At time $t = 0$ we remove the partition, allowing particles to diffuse from one side of the chamber to the other, and we begin taking periodic snapshots at a $10\times$ magnification under an Olympus IX71 inverted microscope in order to record the particle positions at periodic time intervals. (We performed the same experiment under equilibrium conditions where the initial concentration was uniform across the whole chamber (results not shown, see Section 3.4)). We take 3 snapshots of the beads in the chamber every time interval of $\Delta t = 10$ seconds for 6 hours. Since there is a possibility that within one snapshot some particles are temporarily overlapping and/or are out of focus, the 3 snapshots are used to minimize that error, as will be seen in Section 3.3.14.3 (the effect of this error will be described in great detail in the following section, especially in Section 3.4). The snapshots are taken using fluorescence microscopy with a SONY DFW-V500 camera. (During the time when no snapshots are taken, a shutter prevents the experimental chamber from being exposed to the incident light, to prevent photobleaching and heating of the chamber (See Section 3.3.9).) We then determine the particle positions at each snapshot using a computerized centroid tracking algorithm [103].

3.2.6 Image Analysis Using IDL for Obtaining the Quantities of Interest: N, J, and $(\Delta J)^2$

Our data consists of a set of images which depict the colloidal particles inside the microfluidics chamber. In order to be able to extract the particle positions and, thus, the time-dependant density distributions from these images, we created an Interactive Data Language (IDL) code. The sections that follow describe the algorithms used in order to extract the quantities of interest.

3.2.6.1 Algorithm for Finding the Concentration Matrix, N

We begin by finding the concentration matrix, the matrix that contains information about how many particles there are at each position in the chamber at all times. The concentration matrix, **N**, is required for the calculation of all the other parameters of interest, such as the fluctuations in the number of particles, the fluctuations in the flux, the bad actors, etc. In this section we explain in detail how each element of the concentration matrix is obtained.

Our data set contains a series of pictures taken at an interval Δt apart. Let's assume that one of our data sets contains k frames. Each frame is colored and depicts the chamber of interest, within which the diffusion of the beads is taking place. Each image is imported in IDL. In that software, it is transformed to a black-and-white image and is cropped so that only the region of interest (i.e., the chamber) is shown. In order to identify the beads and their positions, we use IDL code developed by Erik Weeks and John Crocker for tracking particles [103]. Based on this code, the image is smoothed

and, after the user has input parameters about the size of the beads, the brightness threshold, etc., the beads and their positions are identified. So, the IDL output at this point is a file with the particle positions.

After the particle positions have been found, the time-dependent particle density is determined by dividing the chamber into a number of equal-sized bins of width Δx, each along the longest dimension of 400 μm, and by computing the number of particles in each bin (for a detailed discussion on the choice of the value for the bin size Δx see Section 3.2.6.2). Although the microfluidic chamber is three-dimensional, it can be shown that, in the case of weak particle-particle and particle-wall interactions, the problem can be collapsed to a one-dimension diffusion problem (see Section 3.3.10). Therefore, we bin only along the x-axis, the direction of the concentration gradient. Each bin then corresponds to a column in the concentration matrix. That is, each row of the concentration matrix contains i elements (i.e., i columns) equal to the number of bins. Each one of the i elements has the value of the number of particles that belong to that bin. For example, the first element of the row is equal to the number of particles in bin 1, the second element is equal to the number of particles in bin 2, etc. Since we have assumed that our data set consists of k frames, that means that the concentration matrix $\mathbf{N}$ will contain k such rows. As a result, we have a concentration matrix with dimensions $\mathbf{N} = \mathbf{N}(i, k)$, where we have used the IDL notation for defining matrices (the first index corresponds to the number of columns of the matrix and the second index corresponds to the number of rows). The matrix $\mathbf{N}$ is, therefore, defined as follows:

$$\mathbf{N} = \begin{pmatrix} n_{11} & n_{21} & \cdots & n_{i1} \\ n_{12} & n_{22} & & \vdots \\ \vdots & & \ddots & \\ n_{1k} & n_{2k} & \cdots & n_{ik} \end{pmatrix}. \tag{3.13}$$

For example, element n_{pq} is equal to the number of particles in bin p of the q^{th} frame (i.e., at time $t = q\Delta t$).

3.2.6.2 The Importance of the Bin Size Value, Δx

When looking at two consecutive bins, we define the number of particles in bin 1 and bin 2 as N_1 and N_2, respectively. As is expected, the range of values for N_1 and N_2 depends on the value of the bin size Δx. The choice of the bin size will affect the statistics for each combination of N_1 and N_2 as well as the range of their sum $N = N_1 + N_2$ and their difference $\Delta N = N_1 - N_2$ themselves. If the bin size is too big, then there will not be a lot of bins and, consequently, there will not be enough statistics. In addition, the range of values for N will not include small numbers (since it will be rare to have 1 or 2 particles in a single bin). On the other hand, if the bin size is too small, we may not

have a sufficient range of values for N and ΔN (since it will be rare to have more than a couple of particles in the small bin). Also, in this case, there will be an increased probability for a particle to have multiple jumps across bins within the time interval Δt. As was explained in Section 3.1, the discrete value of Δt provides the a time interval during which all the action happens, i.e., the particle will jump the distance Δx or stay at its current location. Experimentally, this discrete value is the result of the image acquisition process, as was described in Section 3.2.5. Therefore, the optimal choice of the bin size was made based on the bead's expected mean excursion within the time-interval Δt, which is $\sqrt{2D\Delta t}$. This is the only relevant microscopic length scale. Here, D is the diffusion coefficient for an individual bead given by the Stokes's formula [114]. For a bead of 0.29 μm in diameter suspended in water at room temperature, the Stokes's formula gives a diffusion coefficient D of approximately 1.5 μm^2/s. As will be described in greater detail in Section 3.3.12, this value, within experimental error, is equal to the one we obtain by fitting our data of the concentration profile at different times to the one-dimensional diffusion equation using D as our fitting parameter (i.e., $D = 1.3 \pm 0.27$ μm^2/s. This gives a bin size of $\Delta x \approx 5$ μm. By observing all the consecutive bin pairs for all the frames taken we were able to obtain, on average, about 5000 points for each combination of N_1 and N_2. Given the bead concentration in the microfluidics channel, N_1 and N_2 ranged from 0 through 6. The choice of bin size determines the value of the jump probability p, as discussed in 3.3.13.

In the following section we will use the concentration matrix **N** in order to find the flux **J** and its higher moments. In Section 3.1 we made predictions about the flux distribution, $P(J)$, the average flux, $\langle J \rangle$, the fluctuations in the flux, $\langle \Delta J^2 \rangle$, etc. Therefore, in order to be able to compare the experimental results with these theoretical predictions, we need to acquire the fluctuation matrix **J** given the concentration matrix **N**.

3.2.6.3 Algorithm for Finding the Flux Matrix, J

Just like the concentration matrix, **N**, is the matrix that contains information about how many particles there are at each position in the chamber at all times, the fluctuation matrix **J** is the matrix that has information about the flux of particles at a given plane during all time intervals Δt. In this section, we explain how the flux matrix **J** is obtained.

As we have explained before, each of the data sets gives a matrix, **N**. The number of rows of this matrix is equal to the number of frames in the data set while the number of columns of the matrix is equal to the number of bins that we have divided our chamber into. So if a data set contains k frames and the chamber in which diffusion occurs is divided into i bins, then the matrix will have i

columns and k rows. That is, $\mathbf{N} = \mathbf{N}(i,k)$.

$$\mathbf{N} = \begin{pmatrix} n_{11} & n_{21} & \cdots & n_{i1} \\ n_{12} & n_{22} & & \vdots \\ \vdots & & \ddots & \\ n_{1k} & n_{2k} & \cdots & n_{ik} \end{pmatrix} \tag{3.14}$$

After the matrix **N** has been found, the flux matrix **J** can be found via the continuity equation. First, let us find the dimensions of this matrix. The flux can only be defined when there is a time interval involved, that is, we can only have flux between two different frames (two frames correspond to one time interval). Therefore, the flux matrix will have one less row than **N**. In addition, flux is defined at a plane, but there in only one plane separating two bins. Thus, the number of planes is equal to one less than the number of bins. However, we also need to include the planes corresponding to the boundaries. For the example of two bins, that means that there is a total of three planes; two correspond to the boundaries and one to the plane that separates them. Thus, **J** will have one more column than **N**. Consequently, $\mathbf{J} = \mathbf{J}(i+1, k-1)$.

Now, let's apply the continuity equation for bin l. Since the total number of particles is conserved, then for the given bin l the number of particles at time $(t + \Delta t)$, $n_l(t + \Delta t)$, is equal to the particles that were in the same bin at time t, $n_l(t)$, plus the number of particles that came in during the interval Δt, $n_{\text{in}}(t) = J_{\text{in}}(t)\Delta t$, minus the number of particles that left the bin during the time interval Δt, $n_{\text{out}}(t) = J_{\text{out}}(t)\Delta t$.

We have:

$$n_l\left(t + \Delta t\right) = n_l\left(t\right) + \left(J_{\text{in}}\left(t\right) - J_{\text{out}}\left(t\right)\right)\Delta t \tag{3.15}$$

$$\Rightarrow J_{\text{out}}\left(t\right) = -\frac{n_l\left(t + \Delta t\right) - n_l\left(t\right)}{\Delta t} + J_{\text{in}}\left(t\right), \tag{3.16}$$

where, in our experiments, Δt is the time difference (in seconds) between two consecutive frames. The same expression, written in IDL language (reminder that the first index corresponds to the column (i.e., bin) and the second index to the frame (i.e., time)), is:

$$J\left(l+1, m\right) = -\frac{N\left(l, m+1\right) - N\left(l, m\right)}{\Delta t} + J\left(l, m\right). \tag{3.17}$$

From this equation it is easy to calculate the **J** matrix if we remember the boundary conditions. The first term on the right-hand side of Eq. 3.17 can be easily calculated since the matrix **N**, given by Eq. 3.14, is known. Also, since we cannot have flux of particles into and out of the chamber for all times t, $J\left(1, t\right) = J\left(i+1, t\right) = 0$. Thus, using the first of the boundary conditions and Eq. 3.14, one can find $J\left(2, t\right)$ from Eq. 3.17 and then use this result to find $J\left(3, t\right)$, etc. In this way, a flux

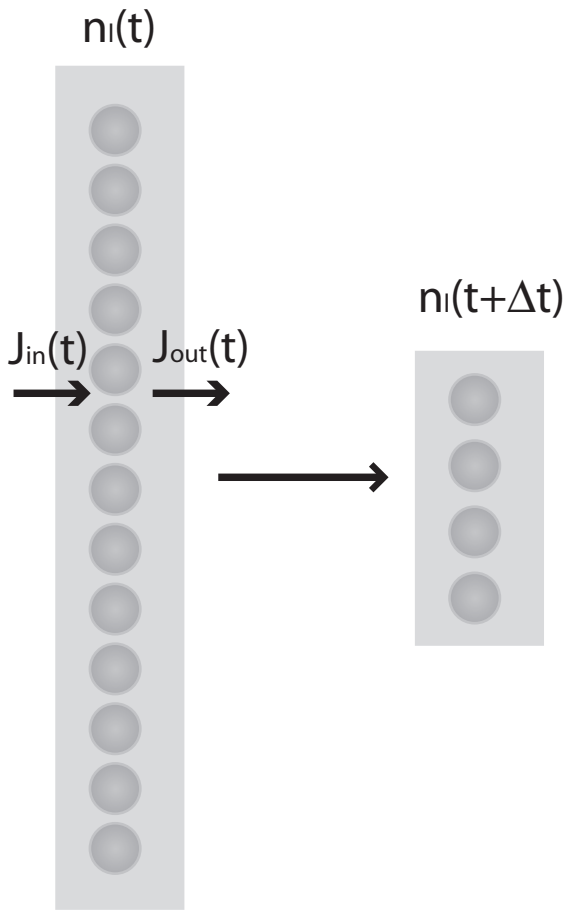


Figure 3.4: Schematic that describes the continuity equation given by Eq. 3.16. The number of particles, $n_l(t + \Delta t)$, in a given bin l at time $(t + \Delta t)$ is equal to the number of particles that were in the bin initially, $n_l(t)$, plus the particles that came in, $J_{\rm in}(t)\Delta t$, minus the particles that left, $J_{\rm out}(t)\Delta t$, during the time interval Δt.

matrix **J** can be found for each data set represented by the matrix **N**.

When considering two consecutive bins with particles N_1 and N_2, respectively, then there is a flux J associated with them. For a given combination of N_1 and N_2 there is a whole distribution of J values, $P(J)$. By averaging these J values we, thus, get the average flux $\langle J\rangle$ for that specific combination of N_1 and N_2. From that distribution we can also get the fluctuations in the flux, $\langle J^2\rangle - \langle J\rangle^2$. Figures 3.5 and 3.6 show examples of the normalized probability distribution $p(J)$ for the specific cases of $(N_1, N_2) = (4, 2)$ and $(N_1, N_2) = (4, 1)$, respectively.

As in Eq. 3.2, each point of the distribution $P(J)$ has, therefore, $N = N_1 + N_2$, $\Delta N = N_1 - N_2$, $\langle J\rangle$, and $\langle(\Delta J)^2\rangle = \langle J^2\rangle - \langle J\rangle^2$ values associated with it. By combining the data for all N_1 and N_2 combinations, we can then get results like those presented in Section 3.4 (see, for example, Figure 3.19).

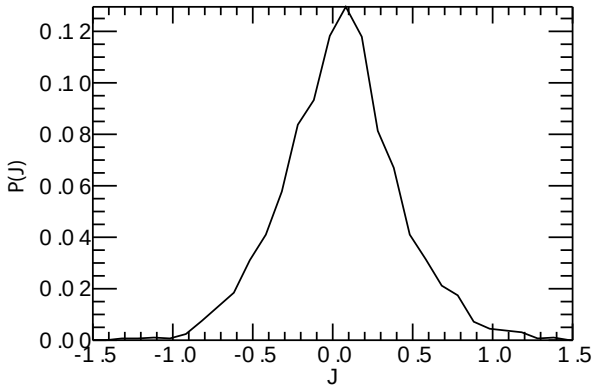


Figure 3.5: The normalized flux probability distribution, $P(J)$, for the specific case of $N_1 = 4$ and $N_2 = 2$. The probability distribution $P(J)$ was found by evaluating all the flux values J for this (N_1, N_2) combination.

It is worth mentioning that the values for the fluctuations in the flux are expected to be positive. This is because $\langle J_i^2 \rangle > \langle J_i \rangle^2$. The proof for that is straightforward. The square of any quantity is a positive number:

$$\langle (\langle J_i \rangle - J_i)^2 \rangle > 0. \tag{3.18}$$

If we expand the square and distribute the average among all the terms we get:

$$\langle \big(\langle J_i \rangle^2 - 2\langle J_i \rangle J_i + J_i^2 \big) \rangle > 0 \tag{3.19}$$

$$\Rightarrow \langle J_i^2 \rangle - 2\langle J_i \rangle \langle J_i \rangle + \langle J_i \rangle^2 > 0 \tag{3.20}$$

$$\Rightarrow \langle J_i^2 \rangle - \langle J_i \rangle^2 > 0. \tag{3.21}$$

Thus, indeed, the values of the fluctuations in the flux are positive quantities.

In order to ensure that the code is correct (i.e., it indeed calculates the fluctuations in the flux), at the end of our algorithm, we perform a check on the final fluctuation matrix. By shifting (by one column) the fluctuation matrix, subtracting it from the original matrix, and adding to this result the difference between two consecutive frames of the **N** matrix divided by Δt, we get a final matrix, **G**, which, if the fluctuation matrix was found correctly, should be 0 (essentially, we just plugged our found matrix **J** back to Eq. 3.17. In IDL language, the above description translates as follows (for the purposes of this example, it is done only for the n^{th} row of the J matrix. The i^{th} column is

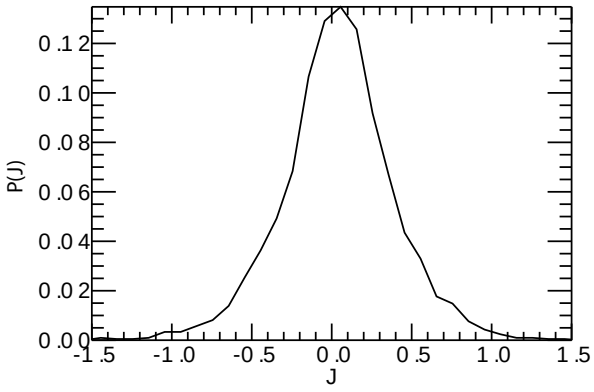


Figure 3.6: The normalized flux probability distribution, $P(J)$, for the specific case of $N_1 = 4$ and $N_2 = 1$. The probability distribution $P(J)$ was found by evaluating all the flux values J for this (N_1, N_2) combination.

excluded because we have defined it to be 0):

$$G = (J(1:(i-1),n) - J(0:(i-2),n)) + \frac{N(0:(i-2),(n+1)) - N(0:(i-2),n)}{\Delta t}. \qquad (3.22)$$

We conclude this section by saying that in this algorithm, in order to find **J**, we have invoked only our data, given by the matrix **N**, and the continuity equation, which is an expression for mass conservation, with our experimental boundary conditions. No other assumptions have been made.

3.3 Data Analysis and Some Estimates

There are a number of issues that arise in the implementation of this experiment that could contaminate our ability to do a precise measurement. To that end we perform estimates as a sanity check to explore some of the subtleties about experimental design. More specifically, the estimates presented in this section were done in order to verify that the way the experiment is set up ensures that the particles are diffusing free from any interactions from the moment the partition is removed at time $t = 0$. At the end of this section we also present a description of the experimental error and the steps that were taken in order to reduce and correct it.

3.3.1 Estimation of the Reynolds Number for Our Microfluidic Design

Since we are using microfluidics, an important number to estimate is the Reynolds number. We know that the Reynolds number, Re, is given by the formula:

$$\mathrm{Re} = \frac{\rho \mathrm{v} \mathrm{l}}{\mu} \tag{3.23}$$

where ρ is the density of the fluid, v is the fluid velocity, l is the characteristic length and μ is the fluid viscosity. We know that if $\mathrm{Re} < 2300$ the flow is laminar and, thus, inertia can be neglected [120]. In this case we have a water solution and, thus,

$$\nu = \frac{\mu}{\rho} \Rightarrow \nu = 0.01 \text{ cm}^2/\text{s}. \tag{3.24}$$

Also, in our case, as was seen in Figure 3.3, $l = 10$ μm (for the channel height) or $l = 100$ μm (for the channel width). When filling up the channels of the microfluidic chip with the colloidal suspension, we can see under the microscope how long it takes for the channels to fill. An approximate value for the velocity of the fluid when it is filling up the channels is $v = 100$ μm/s. Let's consider the worst case scenario by taking the largest value for l. We have:

$$\mathrm{Re} = \frac{\rho \mathrm{v} \mathrm{l}}{\mu} \Rightarrow \mathrm{Re} = \frac{\mathrm{vl}}{\frac{\mu}{\rho}} \Rightarrow \mathrm{Re} = \frac{10^{-4}\,\frac{\mathrm{m}}{\mathrm{s}} \cdot 100 \cdot 10^{-6}\mathrm{m}}{10^{-6}\,\frac{\mathrm{m}^2}{\mathrm{s}}} \Rightarrow \mathrm{Re} = 0.01 << 2300. \tag{3.25}$$

Therefore, indeed, inertia can be neglected. This conclusion is helpful when estimating the time it takes for the water solution to reach equilibrium after the partition is removed, shown in the following section.

3.3.2 Estimation of the Time It Takes for the Water Solution to Reach Equilibrium After the Middle Control Line Is Opened

At the beginning of the experiment the middle control valve is opened and there is a sudden rush of fluid from both sides of the chamber. For the side of the chamber that is filled with the colloidal suspension, this sudden rush results in particle transport. So, the question that arises is: After the partition is removed, how much time does it take for the suspension to reach equilibrium in the microfluidic channel? After this time the particles can move only via free diffusion. We consider the complete Navier-Stokes equation:

$$\rho \frac{\partial \vec{v}}{\partial t} + \underbrace{\rho \vec{v} \nabla \vec{v}}_{\text{inertia}} = \underbrace{\rho \vec{g}}_{\text{gravity}} - \underbrace{\nabla \vec{P}}_{\text{pressure}} + \underbrace{\mu \nabla^2 \vec{v}}_{\text{viscosity}}, \tag{3.26}$$

which is Newton's second law for fluids. In the above equation, $\vec{v}$ is the fluid velocity, ρ is the fluid density, μ is the fluid viscosity, $\vec{g}$ is the acceleration due to gravity, $\vec{P}$ is the external pressure exerted on the fluid and t is the time. However, for the case of our experiment the following conditions are met:

1. Gravity does not contribute since flow takes place in the horizontal direction, parallel to the ground.

2. Since there is no external pressure applied (i.e., the chamber is isolated), the pressure term can be neglected.

3. Due to the low Reynolds number (see previous section) the inertia term can also be neglected.

Thus, the previous equation reduces to:

$$\rho \frac{\partial \vec{v}}{\partial t} = \mu \nabla^2 \vec{v}. \tag{3.27}$$

From this expression, using simple dimensional analysis, we get:

$$\rho \frac{v}{t} = \mu \frac{v}{l^2} \Rightarrow t = \frac{\rho l^2}{\mu} \Rightarrow t = \frac{l^2}{\nu}, \tag{3.28}$$

where $\nu = \frac{\mu}{\rho} = 0.01$ cm^2/s for water, as we found in the previous section. Also, in our case, $l = 10$ μm (for the channel height) or $l = 100$ μm (for the channel width). Considering the channel width (the worst case scenario), we have:

$$t = \frac{l^2}{\nu} \Rightarrow t = \frac{10^{-4} \text{ cm}^2}{0.01 \text{ cm}^2/\text{s}} \Rightarrow t = 0.01 \text{ s}. \tag{3.29}$$

This means that the equilibrium is reached almost instantaneously after the middle control line is opened and, consequently, the motion of the beads is only due to diffusion after this time has passed. Indeed, to test this estimate experimentally, we took videos of the opening of the partition at a rate of 30 frames per second, the fastest rate permitted by our camera. Analysis of those frames (data not shown) shows that, on average, it takes 1 frame (i.e., 33 ms = 0.03 s) before the suspension reaches equilibrium (the time point of equilibrium was defined as the time point when there was no transport of particles, i.e., the particles were only executing Brownian motion). This experimental result is dependent on our camera. Had we been able to take frames at a faster rate, for example at 60 frames per second, the time it takes for equilibrium to be reached could be found to be shorter. Nevertheless, we see that our experimental result is of the same order of magnitude as the theoretical estimate found in this section.

3.3.3 Estimation of the Volume of the Chamber

The colloidal particle concentration inside the microfluidics chamber is important to know. On one hand, we need it to be low enough that the particles are not interacting but high enough that we get enough combinations for N_1 and N_2 (i.e., the number of particle in two consecutive bins). We described this condition in Section 3.2.4 as optimizing the particle concentration. In this and the sections that follow we estimate parameters that describe how important the value of concentration or, equivalently, the mean particle distance is.

In order to be able to estimate the mean spacing between the particles in the chamber, we need to know its volume. The shape of the chamber, whose volume we want to find, is shown in Figure 3.7. In order to calculate the volume we first need to find the area of the "circular-like" entrance to the chamber. We calculate this area by considering part of a circle of radius r (Figure b, shaded area). We use Figure c as an easy reference for the calculations that follow. First we need to find the

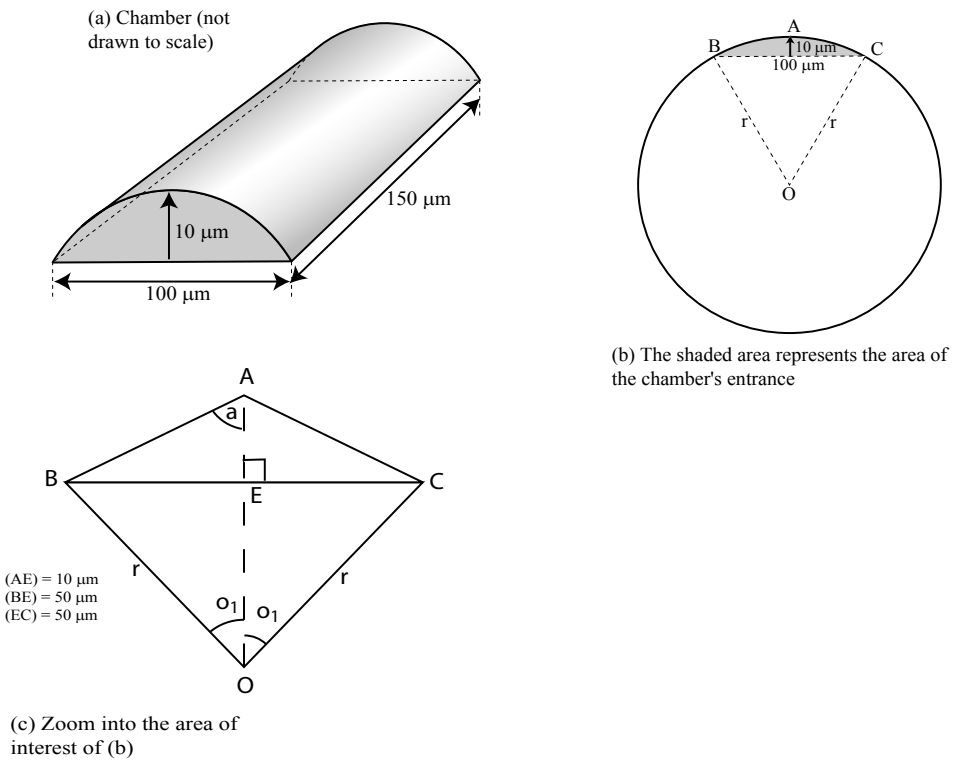


Figure 3.7: The geometry of the microfluidics chamber. This cartoon depicts the microfluidic chamber inside of which the experiment of free diffusion takes place. The chamber is part of a cylinder (a). The entrance to the chamber is part of a circular disk (b) whose area we find with the help of figure (c).

radius, r, of the circle. With the help of Figure 3.7c, we see that we can find the radius by using triangles ABE, OBA, and OBE, since the dimensions of AE, BE, and EC are known. The goal is to find the distance $r = AO = AE + EO$. First we use $\triangle(ABE)$ to find the angle a:

$$\tan a = \frac{50}{10} \Rightarrow a = 78.7^{\circ}. \tag{3.30}$$

Then we use triangle $\triangle(OBA)$ to find the other angle O_1:

$$\hat{O_1} = 180^{\circ} - 2 \times 78.7 \Rightarrow \hat{O_1} = 22.62^{\circ}. \tag{3.31}$$

Finally, we use $\triangle(OBE)$ to find EO:

$$\tan \hat{O_1} = \frac{BE}{EO} \Rightarrow \tan 22.6^{\circ} = \frac{50}{EO} \Rightarrow EO = 120\ \mu\text{m}. \tag{3.32}$$

Thus, the radius is:

$$r = AE + EO \Rightarrow r = 10 + 120 \Rightarrow r = 130\ \mu\text{m}. \tag{3.33}$$

Therefore, the area of the circular section (OBAC) will be given by:

$$\text{Area(OBAC)} = \frac{2 \times 22.62^{\circ}}{360^{\circ}} \pi \text{r}^2 \Rightarrow \text{Area(OBAC)} = 6.67 \times 10^{-9}\ \text{m}^2. \tag{3.34}$$

The area of the triangle OBC is given by:

$$\triangle(OBC) = \frac{1}{2} BC \times OE \Rightarrow \triangle(OBC) = \frac{1}{2} 100 \times 10^{-6} \times 120 \times 10^{-6} \Rightarrow \triangle(OBC) = 6 \times 10^{-9}\ \text{m}^2. \tag{3.35}$$

Therefore, the area of the shaded region of the circle, which corresponds to the area of the entrance to the chamber, is given by:

$$\text{Shaded Area} = \text{Area(OBAC)} - \triangle\text{(OBC)} \Rightarrow \text{Shaded Area} = 6.7 \times 10^{-10}\ \text{m}^2. \tag{3.36}$$

So, for the volume, V, of the chamber we have:

$$V = \text{Shaded Area} \times \text{length} \Rightarrow \text{V} = 6.7 \times 10^{-10}\ \text{m}^2 \times 150 \times 10^{-6}\ \text{m} \Rightarrow \text{V} = 1.005 \times 10^{-13}\ \text{m}^3 \approx 10^5\ \mu\text{m}^3. \tag{3.37}$$

3.3.4 Estimate of the Mean Spacing Between the Beads in the Chamber

Using the result for the volume, V, of the chamber that we found in the previous section, we can calculate the mean spacing between the beads in the chamber as follows. The mean distance, d,

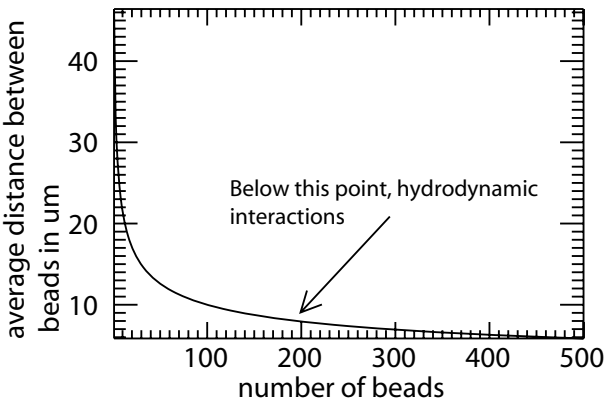


Figure 3.8: Average distance between beads trapped in the chamber vs. the number of beads. This plot is based on Eq. 3.39. The arrow indicates the distance (about 8 μm) below which the hydrodynamic interactions become critical [121].

between the beads that are trapped in the chamber is given by the formula:

$$d = \sqrt[3]{\frac{\text{chamber volume}}{\text{number of beads}}}. \tag{3.38}$$

Since the chamber volume is approximately 10^5 μm^3, the above formula becomes

$$d = \sqrt[3]{\frac{10^5}{\text{number of beads}}}, \tag{3.39}$$

where d will be given in μm. Figure 3.8 shown is a plot of d versus the number of beads. On that plot we show the point (200 beads, 8 μm) below which hydrodynamic interactions become significant [121].

In our experiment the number of particles is about 200. So, for the mean spacing, d, we have

$$d = \sqrt[3]{\frac{10^5}{\text{number of beads}}} \approx \sqrt[3]{\frac{10^5}{200}} \approx 8 \ \mu\text{m}. \tag{3.40}$$

We will show in the following sections that this number for the mean spacing permits us to neglect electrostatic and hydrodynamic interactions between the beads. The magnitude of these electrostatic and hydrodynamic interactions will also be shown. First, let's find the mass of the polystyrene particles we used in our experiments, since we will need to use this value in later sections.

3.3.5 Estimation of the Mass of the Polystyrene Beads

Taken from the manufacturer's (Duke Scientific) data sheet, the density of the polystyrene beads, ρ, is $\rho = 1.05$ g/cm^3. For our experiments, we used two different size particles, $r = 0.5$ μm and

$r = 0.145$ μm. In Section 3.4 we only present the results for the latter size of particles (the reason for this is described at the end of Section 3.3.8). However, here we will present estimates for both cases.

Then, since $m = \rho V$ we have for our beads:

a) for $r = 0.5$ μm

$$m = \rho V \Rightarrow m = 5.5 \times 10^{-13} \text{ g} \tag{3.41}$$

b) for $r = 0.145$ μm

$$m = \rho V \Rightarrow m = 1.34 \times 10^{-14} \text{ g} \tag{3.42}$$

In the following section, we will use these values for the mass of the beads in order to find what the charge per bead is.

3.3.6 Estimation of the Distance Above the Charged Bead Surface After Which the Bead Appears Neutral

In Section 3.2.4 we briefly mentioned that inside the microfluidics chamber the beads are at an optimized concentration so that bead interactions are negligible while at the same time there is a wide range of $N = N_1 + N_2$ and $\Delta N = N_1 - N_2$ (N_1 and N_2 being the number of particles in two consecutive bins) to allow for sufficient statistics. We just found that the bead concentration is such that the mean particle spacing is about 8 μm. For such large particle separations, electrostatic [122] and hydrodynamic interactions [121, 123] are negligible. Although the polystyrene particles have a small negative surface charge (given by the manufacturer, Duke Scientific, in the case of an aqueous solution of particles), it is readily shown using the Poisson-Boltzmann equation that there is little interaction at these separations. Also, the data of Meiners and Quake [121] and Dufresne et al. [123], applied to our system, show that hydrodynamic interactions should be small. In this and the following sections, we present the estimates that back up these claims.

When beads are in a solution, they present a negative charge on their surface. This is because the surface atoms can wander away from the bead due to thermal motion leaving some of the electrons behind [124]. The same happens when, for example, DNA is put in water: some of the loose atoms detach and leave the electrons behind which render the DNA molecule negatively charged. This surface charge is not all seen by the neighboring beads because the liberated positively charged atoms screen it. The liberated positive charges come from the beads only. This is because, in our experiment, we are using pure water and, thus, no other charges can contribute. In order to ensure that the particles in the experiments diffuse freely, any electrostatic interactions have to be negligible.

In this section we are trying to find the distance, x_0, above the bead's surface that the positive charges extend. For any neighboring beads that are located at distance greater than x_0 the bead appears neutral and, thus, electrostatic interactions need not be taken into account when considering the beads' motion inside the water solution.

The manufacturer (Duke Scientific) gives the charge on the beads to be approximately 10 μeq/g of beads ("eq" stands for equivalent charge). By definition, 1 eq is given by the formula:

$$1 \text{ eq} = N_A \times e^-, \tag{3.43}$$

where N_A is Avogadro's number, $N_A = 6.023 \times 10^{23}$, and e^- is the charge of the electron, $e^- = 1.6 \times 10^{-19}$ C. Consequently, we have that 1 eq = 96368 C and, therefore,

$$10 \ \mu\text{eq/g} \simeq 1 \text{ C/g}. \tag{3.44}$$

It is easy to find what the charge per bead (q) is, since in the previous section we found what the mass of each different size bead is. So, we have:
a) for $r = 0.5$ μm, $m = 5.5 \times 10^{-13}$ g

$$q = 1 \text{ C/g} \times 5.5 \times 10^{-13} \text{ g/bead} \Rightarrow q = 5.5 \times 10^{-13} \text{ C/bead}. \tag{3.45}$$

So, the surface density σ_q is given by dividing the charge q by the area of the spherical particle:

$$\sigma_q = \frac{q}{4\pi r^2} \Rightarrow \sigma_q = 0.175 \text{ C/m}^2.$$

b) for $r = 0.145$ μm, $m = 1.34 \cdot 10^{-14}$ g

$$q = 1 \text{ C/g} \times 1.34 \times 10^{-14} \text{ g/bead} \Rightarrow q = 1.34 \times 10^{-14} \text{ C/bead} \tag{3.46}$$

and the surface charge density is:

$$\sigma_q = \frac{q}{4\pi r^2} \Rightarrow \sigma_q = 0.051 \text{ C/m}^2. \tag{3.47}$$

Knowing the surface charge q, we can now compute the distance x_0. This calculation is based on Nelson's *Biological Physics*, pp. 263–272 [124]. The assumptions made are that we have a planar geometry, i.e., that the charged surfaces are planar, and are immersed in pure water, which is the case for our experiment. In addition, we are looking at the problem in one dimension, so we will work along the x axis. We will first find the electrostatic potential energy V of a counterion as a

function of the distance away from the surface, x (scalar). That is, $V \equiv V(x)$. We will also use the mean field approximation and define $c_+(x)$ as the concentration of counterions. If we assume the ions are moving independently of each other in the potential $V(x)$, then the concentration $c_+(x)$ of counterions is given by the Boltzmann distribution:

$$c_+(x) = c_0 e^{-eV(x)/k_B T}, \tag{3.48}$$

where c_0 is a constant, $eV(x)$ is the electrostatic potential energy of a counterions and k_B and T have their usual meanings of the Boltzmann constant and temperature, respectively. If we choose $V(0) = 0$ then we will have that $c_+(0) = c_0$; c_0 is the concentration of counterions at the surface. From Poisson's equation we know that:

$$\frac{d^2V}{dx^2} = \frac{-\rho_q}{\varepsilon}, \tag{3.49}$$

where ρ_q is the charge density. The charge density is related to the charge concentration via the equation:

$$\rho_q = ec_+(x) = ec_0 e^{-eV(x)/k_B T}. \tag{3.50}$$

Let's define a length, the Bjerrum length, as:

$$l_B \equiv \frac{e^2}{4\pi\varepsilon k_B T} \tag{3.51}$$

in water. l_B is the charge separation distance or, in other words, an indication of how close we can push together two ions with the same charge if their thermal energy is $k_B T$. For monovalent ions in water at room temperature, $l_B = 0.71$ nm. Let's also define a dimensionless potential, $\overline{V}(x)$ as:

$$\overline{V}(x) = \frac{eV(x)}{k_B T}. \tag{3.52}$$

So, if we combine the Poisson equation with the Boltzmann distribution using the definition of the dimensionless potential we have:

$$\frac{d^2\overline{V}}{dx^2} = -4\pi l_B c_0 e^{-\overline{V}}. \tag{3.53}$$

We need to solve the above equation subject to the boundary conditions that

1. From Gauss' law, at the surface (x=0)

$$\left.\frac{d\overline{V}}{dx}\right|_{surface} = 4\pi l_B \frac{\sigma_q}{e}, \tag{3.54}$$

2. Since no charge is located at infinity,

$$\left.\frac{d\overline{V}}{dx}\right|_{\infty} = 0, \tag{3.55}$$

3. From convention, $\overline{V}(0) = 0$.

We guess a solution of the form,

$$\overline{V}(x) = B\ln(1 + \frac{x}{x_0}). \tag{3.56}$$

Plugging this solution into Eq. 3.53 we can find the constants B and x_0 by comparing same order terms. We have:

$$B = 2 \tag{3.57}$$

$$x_0 = \frac{1}{\sqrt{2\pi l_B c_0}}. \tag{3.58}$$

Also, this solution has to satisfy the first boundary condition. If we plug the solution we have that:

$$c_0 = 2\pi l_B(\frac{{\sigma_q}^2}{e}). \tag{3.59}$$

So, now, if we go back to our original potential, $V(x)$, we get:

$$V(x) = 2\frac{k_B T}{e}\ln(1 + \frac{x}{x_0}), \tag{3.60}$$

with $x_0 = e/(2\pi l_B \sigma_q)$.

The layer of thickness x_0 is formally known as the Gouy-Chapman layer and is the constant that we set out to find at the beginning of this section. Plugging in all the values for the various parameters in the expression for x_0 we have:
a) for $r = 0.5$ μm, $m = 5.5 \times 10^{-13}$ g

$$\begin{aligned} & x_0 = e/(2\pi l_B \sigma_q), \\ \Rightarrow \quad & x_0 = \frac{1.6 \times 10^{-19}\text{ C}}{2\pi \times 0.71 \times 10^{-9}\text{ m} \times 0.175\text{ C/m}^2}, \\ \Rightarrow \quad & x_0 = 0.205\text{ nm} \end{aligned} \tag{3.61}$$

b) for $r = 0.145\ \mu$m, $m = 1.34 \times 10^{-14}$ g

$$\begin{aligned} & x_0 = e/(2\pi l_B \sigma_q), \\ \Rightarrow \quad & x_0 = \frac{1.6 \times 10^{-19}\ \text{C}}{2\pi \times 0.71 \times 10^{-9}\ \text{m} \times 0.051\ \text{C/m}^2}, \\ \Rightarrow \quad & x_0 = 0.703\ \text{nm}. \end{aligned} \tag{3.62}$$

As was mentioned, the distance x_0 is the distance above which the beads appear to be neutral. For distances greater than 0.2 nm and 0.7 nm there are not electrostatic interactions between the beads. These values we found for x_0 are so small compared to the mean spacing between the beads, which we found to be 8 μm, that electrostatic interactions between the beads can be neglected. Other than this simple estimate we also conducted the experiment by adding BSA (Bovine Serum Albumin) at a concentration of 1 mg/ml of water as an example of a different salt condition. We observed no change in the corresponding analysis of particle flux (data not shown). This was expected since x_0, as estimated above, was very small to begin with.

Note: If we were to solve this problem in spherical coordinates, we would have to solve the Poisson equation, which now has the following form:

$$\frac{1}{r}\frac{d^2(rV(r))}{dr^2} = \frac{ec_0 e^{-eV(r)/k_B T}}{\varepsilon} \tag{3.63}$$

which can be solved numerically, but will essentially give the same order of magnitude for the Gouy-Chapman layer thickness.

We just found that during the experiment the beads do not interact electrostatically. However, this is not the only type of interaction that can take place. As we mentioned, the beads can also interact hydrodynamically. In what follows, we will look at the hydrodynamics interactions associated with the experiment. Just like with the electrostatic interactions, we wish to find the distances above which hydrodynamic interactions can be neglected. In our estimates for these interactions, we will find how the interactions affect the value for the diffusion coefficient D. Therefore, firstly we will find what this diffusion coefficient D is, based on Stokes' equation.

3.3.7 Estimation of the Diffusion Coefficient, D, Based on Stokes Equation

We know from the Stokes equation that the diffusion coefficient, D, is given by:

$$D = \frac{k_B T}{6\pi r \eta}, \tag{3.64}$$

where r is the radius of the bead, η is the viscosity of the medium (for water in this case $\eta = 10^{-3}$ kg/(m·s)), T is the temperature (in this case taken to be room temperature, i.e., 300K) and k_B is the Boltzmann constant ($k_B = 1.38 \times 10^{-23}$ J/K). Plugging in these values to the equation above gives:

a) for $r = 0.5$ μm

$$D = \frac{1.38 \times 10^{-23} \cdot 300}{6\pi \times 0.5 \times 10^{-6} \times 10^{-3}} \Rightarrow D = 4.39 \times 10^{-13} \text{ m}^2/\text{s} \Rightarrow D = 0.44 \ \mu\text{m}^2/\text{s} \tag{3.65}$$

b) for $r = 0.145$ μm

$$D = \frac{1.38 \times 10^{-23} \times 300}{6\pi \times 0.145 \times 10^{-6} \times 10^{-3}} \Rightarrow D = 1.5147 \times 10^{-12} \text{ m}^2/\text{s} \Rightarrow D = 1.5147 \ \mu\text{m}^2/\text{s}. \tag{3.66}$$

In the next section we will see what the correction to these values is due to the hydrodynamic interactions.

3.3.8 Estimation of the Hydrodynamic Interaction Between the Beads and Between the Beads and the Chamber Walls

Colloidal particles in solutions interact not only electrostatically (due to their surface charge) but hydrodynamically as well. As a particle moves, a flow field is induced in the solvent which will be experienced by all the other particles in the solution. This force that the other particles experience due to the motion of the original particle is defined as a hydrodynamic interaction. When the particle's motion is taking place close to the container's wall, then the wall's no-flow boundary condition will break the symmetry of the particle's flow field and increase the drag on the sphere. The method of reflections was introduced by Faxen in 1927 to address this question. In this section we will see what the magnitude of these hydrodynamic interactions is, both between the particles and between the particles and the chamber walls.

In a previous section we found that the mean spacing, d, between two beads (and, consequently, between a bead and a wall) is about 10 μm. Meiners and Quake [121] observed the hydrodynamic cross correlations between two particles at varying distances. They found that the cross-correlation function of the bead positions shows a time-delayed anticorrelation. For their data representation, they defined a dimensionless quantity $\varepsilon = r/d$, where r is the bead's radius and d is the mean bead separation. They then found the magnitude of the anticorrelation as a function of this quantity ε. As expected, they found that the smaller the value of ε, i.e., the bigger the mean separation d or the smaller the radius r, then the smaller the hydrodynamic interaction (i.e., the smaller the magnitude of the anticorrelation).

Applying their observations to our experiment, we have that, since the radius of the beads is $r = 0.5$ μm, then the strength of the hydrodynamic coupling, ε, is given by:

$$\varepsilon = r/d \Rightarrow \varepsilon = 0.5/10 \Rightarrow \varepsilon = 0.05. \tag{3.67}$$

For the case of $r = 0.145$ μm we have that

$$\varepsilon = r/d \Rightarrow \varepsilon = 0.145/10 \Rightarrow \varepsilon = 0.0145. \tag{3.68}$$

With this strength value at hand, we see from Figures 2 and 3 of the Meiners and Quake paper [121] that at the distance of $d = 8$ μm there is hardly any correlation (anticorrelation, to be exact) between the beads. The same results are presented in a paper by Dufresne et al. [123]. In this paper, the authors examine the motion of two beads separated by a distance d at a height h above a wall. As was explained at the beginning of this section, there are also hydrodynamic interactions between the beads and the walls due to the walls' no-flow boundary condition. Again, in this paper the main ratio of interest is r/d or r/h. Considering that $d = h = 10$ μm, those ratios have the same numerical values that we found above. The authors then find the correction to the Stokes value of the diffusion coefficient, D_0 for two collective (C), where the beads move together, and two relative (R) modes along directions perpendicular ($\perp$) and parallel ($\parallel$) to the initial separation, as well as normal (z) and parallel (xy) to the wall. More specifically, they look at the ratio of the actual diffusion coefficient D over the Stokes coefficient D_0. (**Note:** Here I am considering the two walls at the sides of the chamber (400 μm by 100 μm). The walls at the top and bottom of the chamber will be considered below.)

In the case of the two spheres being far from the walls, the corrections are the same as in the case of the Meiners and Quake paper and are given by the equations:

$$\frac{D_{\perp}^{C,R}(d)}{2D_0} = 1 \pm \frac{3r}{4d} + ..., \tag{3.69}$$

$$\frac{D_{\parallel}^{C,R}(d)}{2D_0} = 1 \pm \frac{3r}{2d} + ..., \tag{3.70}$$

where the positive corrections correspond to the collective modes while the negative corrections correspond to the relative modes. The corrections say that the collective diffusion coefficients $D_{\perp}^{C}$ and $D_{\parallel}^{C}$ are enhanced by hydrodynamic coupling (fluid displaced by one particle entrains the other) while the relative diffusion coefficients $D_{\perp}^{R}$ and $D_{\parallel}^{R}$ are suppressed (since fluid needs to be transported in and out of the particles).

In order to describe the effect of the walls, the following equations are used:

$$\frac{D_z(h)}{D_0} = 1 - \frac{9r}{8h} + ..., \tag{3.71}$$

when the bead moves perpendicular to the wall and

$$\frac{D_{xy}(h)}{D_0} = 1 - \frac{9r}{16h} + ..., \tag{3.72}$$

when it moves parallel to the wall. As expected, both corrections show that the diffusion coefficient D is suppressed due to the presence of the wall (i.e., the drag on the particle is increased). From these equations we observe that, again, the change in the diffusion constant turns out to be of the order of $r/h = 0.05$ for the 1 μm beads and $r/h = 0.0145$ for the 0.29 μm beads. Hence, it would be reasonable to conclude that there is no hydrodynamic coupling between the beads and between the beads and the walls at the sides of the chamber.

As far as the top and bottom walls of the chamber are concerned, the situation is slightly different. The height of the chamber is about 10 μm, meaning that that is the distance between the top and bottom walls. This implies that a bead that is in that chamber can have a maximum distance from either of the two walls of 5 μm. In this case, the numerical value of the above ratio is modified to give:

$$\varepsilon = r/h \Rightarrow \varepsilon = 0.5/5 \Rightarrow \varepsilon = 0.1, \tag{3.73}$$

and for the case of the $r = 0.145$ μm beads we have that

$$\varepsilon = r/d \Rightarrow \varepsilon = 0.145/5 \Rightarrow \varepsilon = 0.029. \tag{3.74}$$

Considering that the Stokes values for the diffusion coefficients D_0 are 0.44 μm^2/s and 1.51 μm^2/s, respectively (found in Section 3.3.7), this is about 23% correction in the first case and 2% correction in the second case. Of course, this correction changes drastically as the bead moves closer to either of the two walls.

It is clear from these estimates that in our experimental set-up hydrodynamic interactions do not play a significant role. However, more importantly, our estimates show that, as expected, the smaller the bead size, the smaller the magnitude of the hydrodynamic interactions. More specifically, we see that for the case of the beads with radius 0.145 μm hydrodynamic interactions between the beads and between the beads and the walls are non-existent (the correction to the diffusion coefficient for the smaller beads is insignificant). Therefore, in order to eliminate the possibility of hydrodynamic interactions, we use the results taken from those ($d = 0.29$ μm) beads. This is why in Section 3.4 we only present the results for the smaller size particles.

With this section we have concluded the discussion on the particle interactions. We have found that for the concentration used in our experiments, which results in a mean particle separation of about 8 μm, electrostatic and hydrodynamic interactions are negligible, especially for the case of the smaller size particle with radius 0.145 μm. In the following section we look at possible temperature changes inside the microfluidic chamber.

3.3.9 Estimation of the Temperature Increase During the Experimental Time

For the purposes of our experiment, we need to shine light, which comes from a mercury arc lamp, to the chamber in order to excite the particles, since they are observed with fluorescence microscopy. As was mentioned in Section 3.2.5, the chamber is not continuously illuminated, since there is a shutter which only opens during the time when a snapshot is taken. However, for the purposes of this estimate, we will consider the worst case scenario in which the sample is illuminated all the time. The goal of this estimate is to see whether there is sufficient heating of the sample to alter the diffusion.

Let's assume, then, that the chamber is continuously illuminated by light during the experimental time and, thus, it is expected that the temperature of the colloidal suspension will increase. It is important to know what this increase in temperature is since the diffusion coefficient, D, that we found in Section 3.3.7 depends on the temperature. In this calculation we will estimate the temperature increase at the end of a 5 hour experiment. We are considering the worst case scenario since, in what follows, we are also not taking into account the heat dissipation to the non-illuminated area/environment.

We know that for a material that is illuminated by light, the energy, I_a, of the absorbed light is given by:

$$I_a = I_0(1 - e^{-\alpha d}), \tag{3.75}$$

where I_0 is the initial light energy, α is the absorption coefficient and d is the distance within the material that the light has travelled. We know that the laser illuminates a glass area of radius $R = 1.5$ mm. The light power, P, that comes to the microscope stage is $P = 10$ mW. Taking the duration of the experiment to be 5 hours, we find the total energy, I_{tot}, to be:

$$I_{tot} = P \times \text{time} \Rightarrow \text{I}_{\text{tot}} = 10 \text{ mW} \times 3600 \times 5 \text{ s} \Rightarrow \text{I}_{\text{tot}} = 180 \text{ J} \tag{3.76}$$

and the illumination area, A_{tot}, is:

$$A_{\text{tot}} = \pi R^2 \Rightarrow A_{\text{tot}} = \pi 1.5 \text{ mm}^2 \Rightarrow A_{\text{tot}} = 7.07 \text{ mm}^2. \tag{3.77}$$

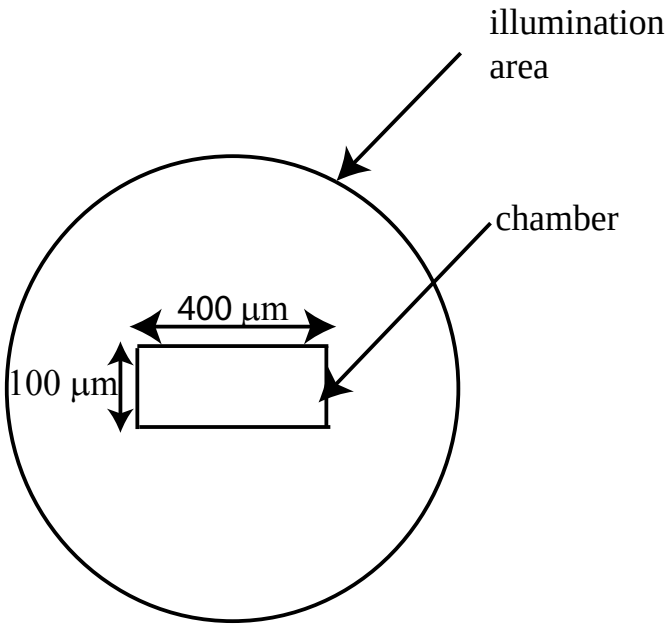


Figure 3.9: The illuminated area of the experiment. This cartoon depicts the total area that the mercury lamp illuminates which includes the experimental microfluidics chamber. The illuminated area has a radius of $R = 1.5$ mm but the chamber dimensions are 400 μm by 100 μm (figure not drawn to scale).

The area of the glass, A_{chamber}, that corresponds to the chamber is:

$$A_{\text{chamber}} = \text{width} \times \text{length} \Rightarrow \text{A}_{\text{chamber}} = 400\ \mu\text{m} \times 100\ \mu\text{m} \Rightarrow \text{A}_{\text{chamber}} = 0.04\ \text{mm}^2. \tag{3.78}$$

So, for I_0 we have:

$$I_0 = I_{\text{tot}} \frac{A_{\text{chamber}}}{A_{\text{tot}}} \Rightarrow I_0 = 180\ \text{J} \frac{0.04\ \text{mm}^2}{7.07\ \text{mm}^2} \Rightarrow I_0 = 1.02\ \text{J}. \tag{3.79}$$

From Figure 3.10 we see that for glass the absorption coefficient, α for a wavelength of 500 nm (the excitation wavelength of the fluorescent particles) is: $\alpha = 0.012\ \text{cm}^{-1}$. Also, for the glass slide we know the following:

average thickness = 0.15 mm,

weight = 360 mg,

length = 6 cm,

width = 2.2 cm.

So, the area of the glass slide is: $A_{\text{slide}} = 6.6 \times 2.2 = 13.2$ cm.

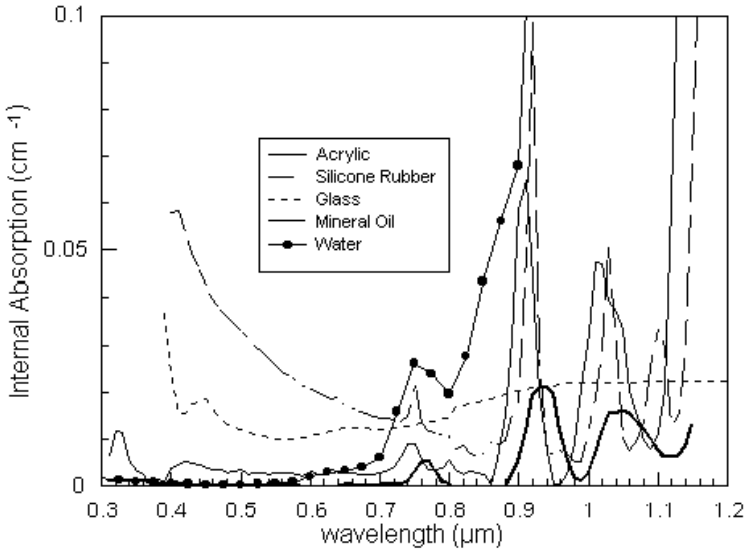


Figure 3.10: Internal absorption coefficient of a number of candidate optical materials as a function of wavelength (taken from *www.pv.unsw.edu.au/thesis/bowden/ch*06/6$_1$*.htm*). Using this graph we can find the absorption coefficient for glass at different wavelengths and, thus, estimate the temperature increase during the experiment.

Therefore, d corresponds to the average thickness and, thus, Eq. 3.75 now becomes:

$$I_a = I_0(1 - e^{-\alpha d}) \Rightarrow I_a = 1.02 \text{ J}(1 - e^{-0.012 \text{ cm}^{-1} 0.015 \text{ cm}}) \Rightarrow I_a = 1.8 \cdot 10^{-4} \text{ J}. \tag{3.80}$$

It is important to note here that we have only considered the energy absorbed by the glass and ignored the energy absorbed by the water. The reason is that, from Figure 3.10, we find the absorption coefficient for water to be almost zero at 500 nm. Consequently, the increase in temperature of the particle and water solution will come only from the direct contact that this solution has with the glass slide.

We know that for a material of mass m and specific heat c, the change in temperature ΔT when heat Q is supplied is:

$$\Delta T = \frac{Q}{mc}. \tag{3.81}$$

We find that the specific heat for glass is $c = 0.84$ J/g and Q corresponds to I_a that we found above. Also, for the mass of the glass that gets illuminated and corresponds to the chamber we have:

$$m = \text{weight} \times \frac{\text{A}_{\text{chamber}}}{\text{A}_{\text{slide}}} \Rightarrow \text{m} = 360 \text{ mg} \frac{0.0004 \text{ cm}^2}{13.2 \text{ cm}^2} \Rightarrow \text{m} = 0.011 \text{ mg}. \tag{3.82}$$

Then, for the change in temperature we have:

$$\Delta T = \frac{1.8 \times 10^{-4} \text{ J}}{0.011 \times 10^{-3} \text{ g} \times 0.84 \text{ J/g}} \Rightarrow \Delta T = 19.5 \text{ K}. \tag{3.83}$$

This temperature change in the illuminated area of the glass slide and, thus, in the particle solution, is quite significant. But how much of an effect would this increase in temperature have on the diffusion coefficient D? In Section 3.3.7 we found the diffusion coefficient using the Stokes equation. Now that the temperature has increased to a value $(T + \Delta T)$, the viscosity of the water that the particles are in has also changed. For $T+\Delta T = 300+19.5 = 319.5$ K the viscosity is $\eta' = 0.65\times10^{-3}$. At this new temperature the diffusion coefficient $D(T + \Delta T)$ is

a) for $r = 0.5$ μm, $D = 0.44$ μm^2 at $T = 300$ K

$$D(T+\Delta T) = \frac{k_B(T + \Delta T)}{6\pi r\eta'} \Rightarrow D(T+\Delta T) = \frac{1.38 \times 10^{-23} \times 319.5}{6\pi \times 0.5 \times 10^{-6} \times 0.65 \times 10^{-3}} \Rightarrow D(T+\Delta T) = 0.72 \ \mu\text{m}^2/\text{s}. \tag{3.84}$$

So, the diffusion coefficient has changed by $\Delta D = 0.28$ μm^2/s.

b) For $r = 0.145$ μm, $D = 1.5147$ μm^2 at $T = 300$ K

$$D(T+\Delta T) = \frac{k_B(T + \Delta T)}{6\pi r\eta'} \Rightarrow D(T+\Delta T) = \frac{1.38 \times 10^{-23} \times 319.5}{6\pi \times 0.145 \times 10^{-6} \times 0.65 \times 10^{-3}} \Rightarrow D(T+\Delta T) = 2.5 \ \mu\text{m}^2/\text{s}. \tag{3.85}$$

Thus, in this case the diffusion coefficient has changed by $\Delta D = 0.97$ μm^2/s.

We can see that this change in the diffusion coefficient ΔD due to the temperature increase is significant. However, as a reminder, here we have assumed that the chamber is being illuminated continuously for the total duration of 5 hours and we did not take into account the heat dissipation into the environment. Needless to say that, based on the results we found for this worst case scenario, the fact that we use a shutter so that the chamber is illuminated for only a couple of seconds (i.e., while a picture is taken) combined with the heat being dissipated completely eliminates the possibility of any temperature increases other than those coming from changes in the temperature of the room itself.

3.3.10 Collapse of the Three-Dimensional Problem to One Dimension

As was explained above, the microfluidic chamber where the experiment takes place is part of a cylinder and, thus, three-dimensional. Therefore, the diffusion of the particles takes place in three dimensions. However, for the theoretical predictions presented in Section 3.1, the problem involved particles jumping between bins along one dimension, which we called x. Thus, in order to make our experimental results compatible with the theoretical predictions, we need to show that the experimental setup, under the right conditions, can be analyzed as if it were one dimensional.

Indeed, it turns out that for weak particle-particle and particle-wall interactions, i.e., when the diffusion coefficient is independent of position, the problem can be collapsed to a one-dimensional diffusion problem. This is why, as was described in Section 3.2.6, we divide the microfluidics chamber into bins only along the x axis. In this section we demonstrate why this three-dimensional problem collapses to one-dimensional diffusion.

The three-dimensional, governing diffusion equation is,

$$\frac{\partial c(x,y,z,t)}{\partial t} = D(\frac{\partial^2 c(x,y,z,t)}{\partial x^2} + \frac{\partial^2 c(x,y,z,t)}{\partial y^2} + \frac{\partial^2 c(x,y,z,t)}{\partial z^2}). \tag{3.86}$$

As always, $c(x,y,z,t)$ is the particle concentration as a function of position (x,y,z) and time t. In order to reduce the problem to one dimension, we need to show that the y and z dimensions can be eliminated. For this reason we start by integrating the above equation. At a given value of x and t the following operation is performed:

$$\int_{y,z} \frac{\partial c(x,y,z,t)}{\partial t} dydz = \int_{y,z} (D(\frac{\partial^2 c(x,y,z,t)}{\partial x^2} + \frac{\partial^2 c(x,y,z,t)}{\partial y^2} + \frac{\partial^2 c(x,y,z,t)}{\partial z^2}))dydz. \tag{3.87}$$

The last two terms of this equation are

$$\int_{y,z} D(\frac{\partial^2 c(x,y,z,t)}{\partial y^2} + \frac{\partial^2 c(x,y,z,t)}{\partial z^2})dydz, \tag{3.88}$$

which, in order to make the problem one-dimensional, we need to show that they are equal to 0. At a given value of x and t these terms can be rewritten as

$$D\int_{y,z} \nabla.\nabla c(x,y,z,t)dydz, \tag{3.89}$$

where ∇ is the gradient operator on y and z. Using Gauss divergence theorem, it becomes

$$D\int_{S} \nabla c(x,y,z,t).\hat{n}dS, \tag{3.90}$$

where $\hat{n}$ is the normal to the boundary S of the cross section at constant x. But from our experimental configuration (i.e., boundary conditions) we also know that the flux on the boundary is zero, i.e.,

$$\nabla_{(x,y,z)} c(x,y,z,t).\hat{n} = 0. \tag{3.91}$$

On S, $\hat{n}$ only has y and z components and so becomes

$$\nabla c(x,y,z,t).\hat{n} = 0, \tag{3.92}$$

where ∇ is the gradient operator on y and z. Hence Eq. 3.87 becomes

$$\int_{y,z} \frac{\partial c(x,y,z,t)}{\partial t} dydz = \int_{y,z} (D(\frac{\partial^2 c(x,y,z,t)}{\partial x^2} dydz. \tag{3.93}$$

If we replace now $\int_{y,z} c(x,y,z,t)dzdy = \bar{c}(x,t)$ then it is the simple one-dimensional diffusion. Thus, we have shown that at a given point on the x axis, the axis of the concentration gradient, the other two dimensions can be integrated out! The problem has collapsed to one dimension. Therefore, since in our experiments we are observing one-dimensional diffusion, we need to solve the one-dimensional diffusion equation for our experimental configuration.

3.3.11 Solving the Diffusion Equation for the Experimental Configuration

From the estimates presented thus far, it is seen that in our experiment we observe the free diffusion of particles. However, in order to make sure that the observed diffusion is consistent with the diffusion Eq. 2.3, we wanted to fit the data for the time dependent particle distribution to the solution of the diffusion equation for our experimental configuration. We present an example of this fit of the data to the solution in the next section. In this section, we seek to obtain the solution.

We solve the diffusion equation subject to the boundary conditions of our experiment in order to find how the concentration profile depends on the position within the chamber and the time. Figure 3.11 shows graphically the initial conditions of our experiment. That is, at time $t = 0$, half of the microfluidics chamber has particles at a constant concentration c_0 while the other half is empty. The two graphs shown in this figure are equivalent, so only the second one will be used when identifying the boundary conditions. Here we repeat the diffusion Eq. 2.3, which has the form:

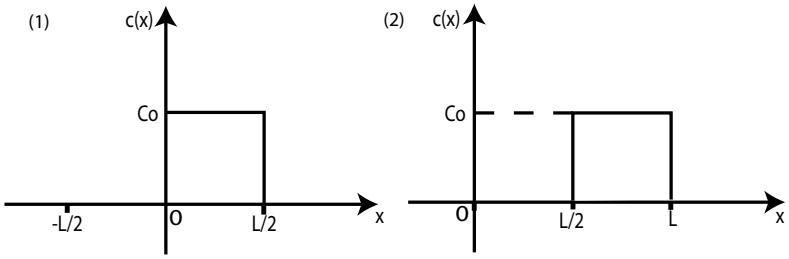


Figure 3.11: The initial conditions of our experiment shown graphically. c is concentration and x is position. L is the length of the chamber. Both graphs are equivalent since only the origin has been shifted. We are using this graphic representation of the experiment in order to identify the boundary conditions.

$$\frac{\partial c(x,t)}{\partial t} = D \frac{\partial^2 c(x,t)}{\partial x^2}, \tag{3.94}$$

where $c(x,t)$ is the concentration of particles as a function of position x and time t, respectively, and D is the diffusion coefficient. From Fick's first law we know that

$$\vec{J} = -D\frac{\partial c}{\partial x}, \tag{3.95}$$

where $\vec{J}$ is the flux of particles. From this equation we get the boundary conditions since at all times t there is no flux J in or out of the chamber (the chamber is bound at $x = 0$ and $x = L$). So, we have:

$$J(x=0) = 0 \Rightarrow \left.\frac{\partial c}{\partial x}\right|_{x=0,\ t} = 0, \tag{3.96}$$

$$J(x=L) = 0 \Rightarrow \left.\frac{\partial c}{\partial x}\right|_{x=L,\ t} = 0. \tag{3.97}$$

We will solve the diffusion equation using separation of variables. We define the solution of the above equation to be of the form:

$$c(x,t) = X(x)T(t), \tag{3.98}$$

where we see that X is only a function of position and T is only a function of time. If we plug this solution to the diffusion equation we get:

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} \tag{3.99}$$

$$\Rightarrow X(x)\frac{dT}{dt} = DT\frac{d^2X}{dx^2}. \tag{3.100}$$

In this last equation we see that the derivative with respect to x only affects $X(x)$, since $T(t)$ is not a function of x and, therefore, it is considered a constant. Similarly, the derivative with respect to t only affects $T(t)$. If we divide both sides of the above equation by (DXT) we get:

$$\frac{1}{DT}\frac{dT}{dt} = \frac{1}{X}\frac{d^2X}{dx^2}. \tag{3.101}$$

The last equation shows that each side has to be equal to a constant since the left side is a function of t only and the right side is a function of x only. Let's name this constant $-k^2$. Then, for each equation, we have:

$$\frac{1}{DT}\frac{dT}{dt} = -k^2 \Rightarrow \frac{dT}{T} = -Dk^2dt \Rightarrow \ln T = -Dk^2t \Rightarrow T(t) = e^{-k^2Dt} \tag{3.102}$$

$$\frac{1}{X}\frac{d^2X}{dx^2} = -k^2 \Rightarrow \frac{d^2X}{dx^2} + k^2X = 0 \Rightarrow X(x) = \begin{cases} \sin(kx) & \text{or} \\ \cos(kx) & \end{cases} \tag{3.103}$$

At this point we can evaluate k. We can see that the function at $t = 0$ that is shown in Figure 3.11

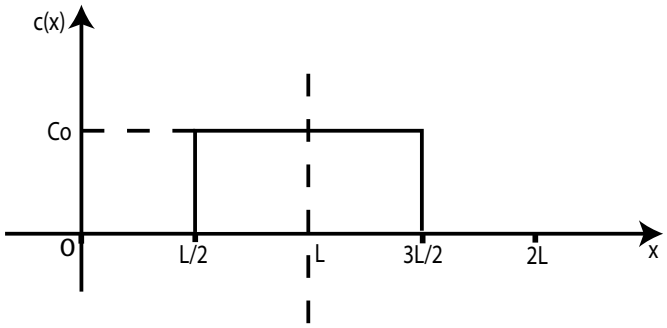


Figure 3.12: The initial conditions of the experiment shown graphically. The difference between this figure and Figure 3.11 is that the latter has been replicated so that the period is $2L$. The replication is based on the mirror image of the function with respect to the axes $x = L$.

can be replicated so that its period is $2L$, as is shown in Figure 3.12. However, as always, we are only interested in the $x \epsilon [0, L]$ interval. Therefore, $k = \frac{n\pi}{L}$ and the above equation now becomes:

$$X(x) = \begin{cases} \sin(\frac{n\pi x}{L}) & \text{or} \\ \cos(\frac{n\pi x}{L}) \end{cases} \tag{3.104}$$

One of the two solutions (i.e., sine or cosine) satisfies the boundary conditions. Using the possible results for $X(x)$ that we found above, we can test the boundary conditions as follows:

$$\frac{dX}{dx} = \begin{cases} \frac{n\pi}{L}\cos(\frac{n\pi x}{l}) \Rightarrow \frac{dX}{dx}\big|_{x=0,L} = \pm\frac{n\pi}{L} \\ -\frac{n\pi}{L}\sin(\frac{n\pi x}{l}) \Rightarrow \frac{dX}{dx}\big|_{x=0,L} = 0 \end{cases} \tag{3.105}$$

We see that the second solution (i.e. $X(x) = \cos\frac{(n\pi x)}{L}$) satisfies the boundary conditions, as these were determined in Eqs. 3.96 and 3.97. If we combine the solutions for $T(t)$ and $X(x)$ our solution for $c(x, t)$ is of the form:

$$c(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n cos(\frac{n\pi x}{L}) e^{-(\frac{n\pi}{L})^2 Dt}. \tag{3.106}$$

Therefore, we now need to solve for the coefficients a_n. For $t = 0$ we can use Figure 3.12 and the fact that $c(x, 0) = X(x) = c_0$ for $x \epsilon [L/2, 3L/2]$ (it is equal to 0 everywhere else). In addition, $c(x, t)$ is an even function, since it is in terms of cosines. Thus, by integrating over the period $[0, 2L]$, we have:

$$a_n = \frac{1}{L}\int_0^{2L} X(x)\cos(\frac{n\pi x}{L})dx \Rightarrow a_n = \frac{1}{L}\int_{\frac{L}{2}}^{\frac{3L}{2}} c_0\cos(\frac{n\pi x}{L})dx. \tag{3.107}$$

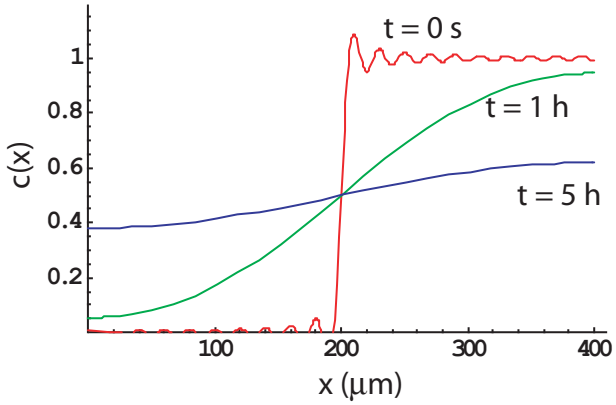


Figure 3.13: The particle concentration profile as a function of position inside the chamber at various time t. We plot the normalized $c(x)$ in the case of the 0.29 μm in diameter beads as a function of position in the chamber x for three different times, $t = 0$ (red), $t = 1h$ (green) and $t = 5h$ (blue). For the plots, we used the first 20 terms in the summation of Eq. 3.110.

In order to evaluate this last integral, we will consider the two cases:

$$n = 0 : a_0 = \frac{1}{L}\int_{\frac{L}{2}}^{\frac{3L}{2}} c_0 dx \Rightarrow a_0 = \frac{c_0}{L}\frac{2L}{2} \Rightarrow a_0 = c_0 \tag{3.108}$$

$$n \neq 0 : a_n = \frac{c_0}{n\pi}\left[\sin\left(\frac{n\pi x}{L}\right)\right]\Big|_{\frac{L}{2}}^{\frac{3L}{2}} \Rightarrow a_n = \frac{c_0}{n\pi}\left(\sin\left(\frac{3n\pi}{2}\right) - \sin\left(\frac{n\pi}{2}\right)\right) \Rightarrow a_n = \frac{2c_0}{n\pi}\sin\left(\frac{n\pi}{2}\right)\cos(n\pi)$$

$$a_n = \begin{cases} 0 & \text{for n-even} \\ \frac{-2c_0}{n\pi}(-1)^{\frac{n-1}{2}} = (-1)\frac{2c_0}{n\pi}(-1)^{\frac{n-1}{2}} = \frac{2c_0}{n\pi}(-1)^{\frac{n+1}{2}} & \text{for n-odd} \end{cases}. \tag{3.109}$$

Due to these results, the final solution is:

$$c(x,t) = \frac{c_0}{2} + \sum_{n-odd}^{\infty} \frac{2c_0}{n\pi}(-1)^{\frac{n+1}{2}}\cos\left(\frac{n\pi x}{L}\right)e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \tag{3.110}$$

for $x \in [0, L]$. In Figure 3.13 we plot the normalized particle concentration $c(x,t)$ as a function of position inside the chamber for three different times, $t = 0$ s, $t = 1$ h, and $t = 5$ h using only the first 20 terms in the summation over n. As an example, this plot corresponds to the particles that are 0.29 μm in diameter ($D = 1.51$ μm^2/s). From the graph we see that if we start with a particle concentration that has the form of a step function, as is the case in our experiments, then for this particle size it is expected that the system will be almost at equilibrium at approximately 5 h. More specifically, at $x = 0$ the concentration will be at 80% of its final value of $c_0/2$ after approximately 5.5 h.

If we are indeed observing the free diffusion of particles in our experiments, then it is expected that the concentration profiles will be described by Eq. 3.110 and will have the form shown in Figure 3.13. In the following section, we explore the form of the concentration profiles obtained experimentally.

3.3.12 Example of Fitting the Data to the Solution of the Diffusion Equation Using D as the Fitting Parameter

As was mentioned in the previous section, it is essential to confirm that what is observed under the microscope is, indeed, free particle diffusion. For this reason, fitting the data for the density distribution to the solution of the diffusion equation given by Eq. 3.110 using the diffusion coefficient D as the fitting parameter was essential. The goal was to see if the value that we got for D from the fitting was, within error, equal to the one given by Stokes equation as was found in Section 3.3.7.

We did the fitting using IDL's non-linear least squares fit algorithm. The algorithm works for an arbitrary number of fitting parameters but, as has been explained, in our case the only fitting parameter was the diffusion coefficient D. After fitting the data for the density distribution for various times and for different size particles (1 μm and 0.29 μm in diameter), we found that the values for D obtained from the fits were, within error, equal to the value given by the Stokes formula. In Figure 3.14 we show one such example of the fitting process. In this figure, the density distribution data of 1 μm in diameter particles at time $t = 90$ min is shown as a solid black line. On top of that we superpose the solution to the diffusion equation that we found (see Eq. 3.110) plotted for two different values of D: The solid red line corresponds to the solution when the value that we used for D came from the fit. In this case we had found $D = 0.45$ μm^2/s. The dashed blue line corresponds to Eq. 3.110 plotted using the value for D given by the Stokes formula, $D = 0.44$ μm^2/s. We see that the two lines overlap, confirming that the experimentally observed diffusion complies to the diffusion equation.

After averaging all the values that we got for D from fits of the density distribution at various times, we found the values to be: $D = (1.3 \pm 0.27)$ μm^2/s for the 0.29 μm and $D = (0.48 \pm 0.06)$ μm^2/s. As a reminder, we note that the respective values we found in Section 3.3.7 were $D = 1.51$ μm^2/s and $D = 0.44$ μm^2/s. Therefore, since the values of D from the fits are, within experimental errors, equal to the ones we found using the Stokes' equation, we can safely conclude that the observed diffusion can, indeed, be modeled by Eq. 3.110, which is the solution to the diffusion equation.

In the sections that follow we present some estimates that will help with the interpretation of the experimental results, which will be presented in Section 3.4.

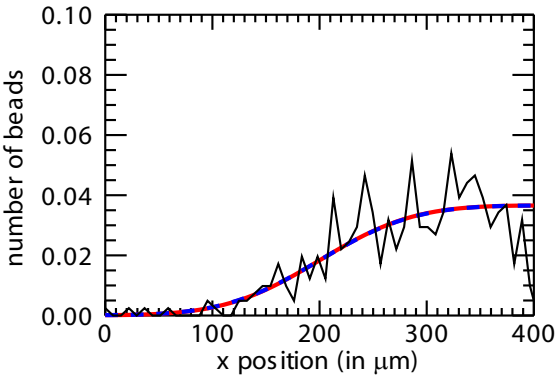


Figure 3.14: Normalized density distribution of 1 μm beads at time $t = 90$ min after the partition was removed. The solid black line shows the actual experimental data. After fitting the data to the solution given by Eq. 3.110 we got a value of $D = 0.45\ \mu\text{m}^2/\text{s}$. The solid red line shows the solution of Eq. 3.110 plotted using the value of D taken from the fit. The dashed blue line shows the same Eq. 3.110 but plotted with the value of D given by the Stokes formula, $D_S = 0.44\ \mu\text{m}^2/\text{s}$. It can be seen that the two lines overlap.

3.3.13 Discussion of the Bin Size Dependence on the Estimation of Flux

In Section 3.1 we presented theoretical predictions for various quantities, like $\langle J \rangle$ and $\langle (\Delta J)^2 \rangle$, etc. These quantities were found to depend on the probability that a particle will jump, p, in the time interval Δt. In Section 3.2.6.2 it was briefly mentioned that p depends on the value of the bin size, i.e., the width of the bin into which we divided the microfluidics chamber. In this section we discuss the role of the bin size and the corresponding transition rate of particle hopping from one bin to the other. In short, we find the slope p of the plot $\langle J \rangle$ versus ΔN (see Figure 3.15).

Let the bin size be Δx. If the time step is Δt, the standard deviation for the displacement for the bead with diffusion constant D is given by $\sigma = \sqrt{2D\Delta t}$. Let's consider a bin i; we are interested in finding the transition probability that a particle from bin $i-1$ (on the left side of bin i) enters bin i. The probability that a particle is found some place in bin $i-1$ is given by,

$$p_{i-1} = \frac{1}{\Delta x}. \tag{3.111}$$

If the particle starts its diffusion at position $x = x_1$, then the probability that the particle will diffuse

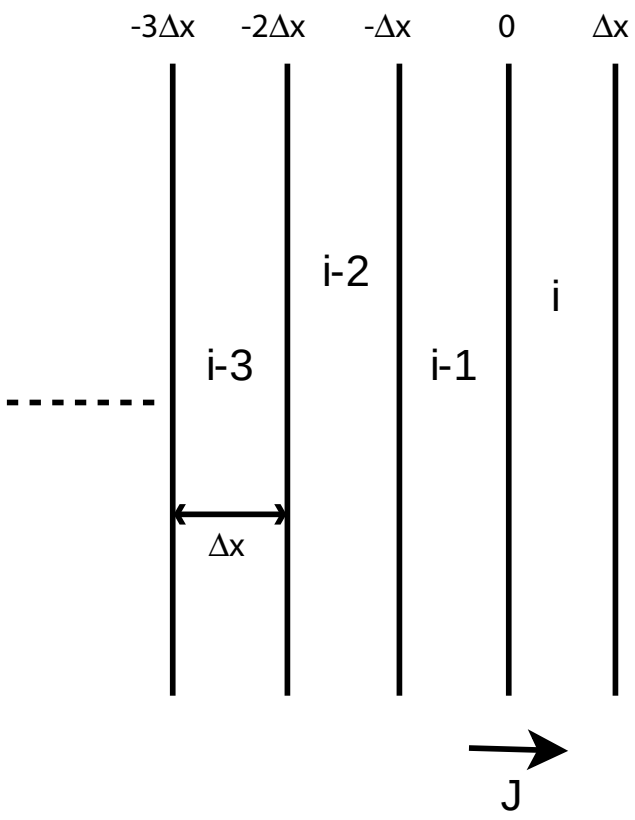


Figure 3.15: Schematic that shows the contribution of the flux into bin i from the bins to the left. Adjacent bins of width Δx. Particles that are in bins $i-1$, $i-2$, etc., all contribute to the flux of particles J that cross the plane at $x = 0$ into bin i

a distance $|x - x_1|$ is given by the spreading Gaussian distribution:

$$P = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-x_1)^2}{2\sigma^2}\right). \tag{3.112}$$

Consequently, the probability that the particle will reach some place x in bin i between $x = 0$ and $x = \Delta x$ is given by the product of the two probabilities:

$$\begin{aligned} q_1 &= p_{i-1} \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-x_1)^2}{2\sigma^2}\right) \\ &= \frac{1}{\Delta x} \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-x_1)^2}{2\sigma^2}\right). \end{aligned} \tag{3.113}$$

The total probability that a given particle located *anywhere* in bin $i-1$ at an initial position x_1 goes to *any* position x bin-i is given by,

$$p_1 = \frac{1}{\Delta x} \times \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{\Delta x} \int_{-\Delta x}^{0} dx_1 dx \exp\left(-\frac{(x-x_1)^2}{2\sigma^2}\right). \tag{3.114}$$

Note that the limits of integration of the integral over x_1 are $x_1 = -\Delta x$ and $x_1 = 0$, while the limits of the integral over x are $x = 0$ and $x = \Delta x$ (see, for reference, Figure 3.15). It must also be noted that the contribution will also come from bins $i-2$, $i-3$, etc., from the left of bin-i. It is not too difficult to see that the total contribution to the jump probability is addition of the respective contributions, and is given by,

$$\begin{aligned} p(\Delta x, \sigma) &= \sum p_k, \\ &= \frac{1}{\Delta x} \times \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{k=0}^{\infty} \int_{-(k+1)\Delta x}^{-k\Delta x} dx_k dx \exp\left(-\frac{(x-x_k)^2}{2\sigma^2}\right). \end{aligned} \tag{3.115}$$

This sum is a function of the bin size Δx and the standard deviation of the particle displacement $\sigma = \sqrt{2D\Delta t}$. After summing all the jump probabilities from all the neighboring bins, the effective probability p is found to be $p \simeq 0.32$, for the given experimental values of $\Delta x \approx 5$ μm and $\Delta t = 10$ s and $D = 1.51$ μm^2/s. This value for p matches the slope for the plot J vs. ΔN obtained by running the simulation for different values of σ and Δx (see, for example, Figure 3.31) as well as the experimental data (see Figure 3.40), as will be shown in Section 3.4.

Starting with the next section, we address the issue of experimental error. We first describe the methods that we used, both experimentally and analytically, in order to reduce the error and/or subtract it off. Later, in Section 3.4, we discuss how this error has affected the obtained experimental results.

3.3.14 Trying to Solve the Problem of the Error in the Particle Number

As was mentioned in Section 3.2.5, our data consists of a set of images taken at time intervals Δt. However, at any given time t, particles may be overlapping and/or drifting out of focus. Based on the specifications of the objective that we are using for the experiment (10x, $N.A. = 0.25$), the Rayleigh resolution limit, i.e., the minimum separation between two particles before they become unresolvable, is about 1 μm. That corresponds to an apparent radius for the particles that is bigger than the actual radius. For example, the apparent radius for the case of the 0.29 μm in diameter beads is $r_{\text{app}} = 1.145$ μm. If two particles come to a distance that is closer than this apparent radius, then the particles appear as one. Consequently, if the number of particles inside the chamber is about 200, then at the initial stages of the experiment (when all the particles are located inside half of the chamber where the length is only 200 μm), the probability of an overlap and, therefore, the estimate of the error in the particles number is found to be about 4%. Indeed, when the images obtained during the experiments are analyzed, there is an error in the number of particles found by the IDL algorithm described in Section 3.2.6, which was evaluated to be about 5-6% of the total number of particles from frame to frame. That is, from one frame to the next the total number

of particles is found to differ by an average amount of 5%, very close to the estimated error of 4%. However, it is important to note that, when considering all the frames, the average change in the total number of particles is zero, $\langle \Delta N_{\text{total}} \rangle = 0$. This statement just shows that the error is random and the particles both appear and disappear as they overlap and drift in and out of focus. In its essence, this error is not big since it corresponds to an error of about 1 particle for every 10 bins. Nevertheless, it has an extremely important effect on our results, especially that of the second moment for the flux, as will be seen in Section 3.4. This is because the flux is calculated via the use of the continuity Eq. 3.17 and, therefore, it gets propagated as we move from one bin to the next, resulting in both overestimation and underestimation in the value of the flux J. Due to the symmetry of the flux distribution, examples of which we saw in Section 3.2.6.3 (see Figures 3.5 and 3.6), when these overestimated and underestimated values of J are raised to an odd power, they cancel each other out and, thus, the odd moments of the flux remain unaffected by this error. However, this is not the case for the even moments, where the overestimated and underestimated values are raised to an even power. This will become clear in Section 3.4, where we will present the results of our data. For example, we will see that the slope of the plot of $\langle \Delta J^2 \rangle$ vs. N (Figure 3.24) was higher by an order of magnitude than the corresponding slope obtained from the simulation in which the number of particles was conserved (Figure 3.34). In this section, however, we present the several attempts that were made in order to experimentally minimize and/or quantify the error and subtract it off.

3.3.14.1 Adjustment of the Particle Size and/or the Intensity Threshold Parameters

When the images are imported in IDL, the particle size and intensity are the two user input parameters that the algorithm uses to distinguish the particles from the noise (see Section 3.2.6). We have tried to adjust the particle diameter (size) and the intensity cutoff. If the particle diameter and brightness are below a certain cutoff one starts collecting noise. As a result, the IDL program will estimate a far greater value (of the order of thousands) for the total number of particles. On the other hand, a very high cutoff in the particle size and the intensity will filter noise, but will also cost us some genuine particle-counts. The adjustment of these two parameters did not allow us to find a sharp transition between the noise and the non-noise regimes in order to clearly eliminate 100% of the former and not exclude the genuine counts. If such a transition were found, we could have chosen the right value for the cutoff diameter and the intensity. This smooth transition sets an arbitrariness in deciding the bead diameter and the intensity cutoff, which could easily lead to a 5% error in the particle number. Thus, our attempts on this front did not meet with success.

3.3.14.2 Using the Intensity Information as a Way for Identifying the Particles

Another way of eliminating the error in the particle number was to use the intensity information. Since energy is conserved, all images should have the same total intensity. We could use this intensity information in order to identify the particles, since they are the ones that contribute to the image intensity (the particles are fluorescent, so they are the only ones seen under the microscope) . Thus, we could integrate over the particle intensities by looking at the pixels that correspond to each particle as well as obtain the total intensity of the entire image. In this way we can get the total particle number. However, during implementation, we observed that even though the intensity (and therefore the energy) of the entire image is conserved in the short time scale, that was not the case in the long time scale. This was because both the background intensity and the individual particle intensities kept changing. A reason for the change in the total background intensity would be that the ambient lights were being switched on and off during the long duration of the experiment. It could also be due to leakage of the particle dye to the medium. The change in the individual particle intensity was mainly due to photo-bleaching, as well as the leakage of the dye to the medium, as has already been mentioned. Therefore, in order to be able to find the number of particles based on intensity information, we need much fewer fluctuations in intensity, which was not the case in our experiments.

In addition, it was found that the particles did not have equal intensities. After integrating over the intensities of several different beads, it was observed that the values were not equal (data not shown), which was unexpected based on the fact that the particles were identical in all other respects. Furthermore, when we looked at bright versus dim particles, it was found that the sum of the intensities of two dim particles had about the same value as just one bright particle! Ideally, a dim particle should have the same total intensity as a bright particle, but spread over a larger number of pixels. Thus, even though this observation was puzzling, we attributed it to two main factors. Firstly, it could be that some particles photo-bleach faster than others. Secondly, it could also be that the particles are not uniform as far as the number of fluorophores (i.e., amount of fluorescent dye) is concerned and, thus, some emit more light and have higher intensity than others. As a result, when this intensity algorithm was used for the particle identification, we found that the change in the number of particles from frame to frame was much greater than 5-6%.

3.3.14.3 Conducting Experiments by Taking Multiple Images at Each Time Interval

Since the particles execute Brownian motion, it is expected that the problem of particles overlapping and/or drifting out of focus can only be temporary. Therefore, as has already been explained in Section 3.2.5, one solution to the problem of not accurately identifying the total number of particles would be to take more than one picture at each time interval Δt. These frames can then be compared

in order to more accurately identify the total number of particles. Implementation of this solution reduced the error in the total particle number from 5-6% to 1-2%. For this reason, we conducted our experiments using this method of capturing multiple images at each time interval. After this experimental correction was implemented, the error was then completely eliminated via the use of a theoretical correction described in the following section.

3.3.14.4 Theoretical Attempt to Correct the Error

A) Via a simulation: In order to better explore the effect of the error in the particle number on the calculation of the quantities of interest ($\langle J \rangle$, $\langle (\Delta J)^2 \rangle$, etc.), we simulated the experiment by running the complete simulation with the entire set of particles, and then removing around 5% of the total particles at random. The results of this simulation will be shown in Section 3.4. However, at this point it must be noted that the experimental process of the disappearance of the beads need not be completely random. For example, in the bins with a larger number of particles it may be more likely for the error to occur due to the increased probability of particles overlapping. As a result, though the simulations give us a very good quantitative match with the experimental results, one needs to be careful while trying to remove the errors from the experiments by comparing them with the simulations.

B) Via a theoretical calculation: In the experimental setup discussed, we start with a step function initial condition. The concentration at time $t = 0$ is such that the left side of the chamber is devoid of particles. As explained in Section 3.2.6.3, we calculate the flux J by using the fact that $J = 0$ at the extreme ends of the chamber and the continuity equation with the assumption that the total number of particles is conserved. Unfortunately, due to the problem of particle overlapping we cannot keep track of all the beads at all times. Hence, the total number of particles N_{total} changes from one time frame to the other.

The continuity equation (Eq. 3.17) can be started from either the left end of the microfluidic chamber or the right end, and then propagated inside the chamber. If the total number of particles were conserved during all the time frames, it would be immaterial, from the point of view of obtaining the flux J, as to which end we start propagating the continuity equation from. Since the total number of particles changes from frame to frame, though, we will get different values of flux at a given plane depending on whether we start from the right or the left end of the chamber. To see this point we write down the following equations for the flux J at plane i:

$$J^i_{\text{right}} = J^i_{\text{actual}} + \sum_k \Delta n_k, \tag{3.116}$$

$$J^i_{\text{left}} = J^i_{\text{actual}} - (\Delta N_{\text{total}} - \sum_k \Delta n_k). \tag{3.117}$$

Here, $J^i_{\rm right}$ is the flux obtained when we start from the right, $J^i_{\rm left}$ is the flux obtained when we start from the left, $\sum \Delta n_k$ is the total discrepancy in the number of particles *till* plane i, while starting from the right, and $\Delta N_{\rm total}$ is the discrepancy in the total number of particles between these two given consecutive time frames. It should be noted that $\Delta N_{\rm total} - \sum \Delta n_k$ is the discrepancy in the total number of particles when we come from the left. Also, the negative sign in the right-hand side (RHS) of Eq. 3.117 is because, when we obtain the flux starting from the left, we calculate the flux in the direction opposite to that of the flux when we start from the right. In what follows we drop the index i on the plane number and the subscript "actual" from the true flux $J_{\rm actual}$.

Taking the difference of the two equations, we obtain

$$J_{\rm right} - J_{\rm left} = \Delta N_{\rm total}. \tag{3.118}$$

We can make a few observations from the above equation. Taking an average for a given value of N_1 and N_2, we obtain

$$\langle J_{\rm right}\rangle - \langle J_{\rm left}\rangle = \langle \Delta N_{\rm total}\rangle. \tag{3.119}$$

But it is seen from the experimental data that $\langle \Delta N_{\rm total}\rangle = 0$ (see Section 3.3.14). Thus we obtain that $\langle J_{\rm right}\rangle = \langle J_{\rm left}\rangle$. Thus, we can add Eqs. 3.116 and 3.117 and use this result to get that

$$\frac{\langle J_{\rm right} + J_{\rm left}\rangle}{2} = \langle J_{\rm act}\rangle + \langle \sum_k \Delta n_k\rangle. \tag{3.120}$$

If we assume that the change in the particle number is symmetric for the positive and the negative values, we can take $\langle \sum_k \Delta n_k\rangle = 0$. Using this fact we obtain that the mean value of flux is independent of the error in the number of particles.

We can use a similar analysis for $\langle (\Delta J)^2\rangle$. It can be safely assumed that $J^i_{\rm actual}$ is independent of Δn_k. We can also further assume that the errors in different planes are uncorrelated. This explicitly means that $\langle \Delta n_p \Delta n_q\rangle = \langle \Delta n_p\rangle \langle \Delta n_q\rangle$. From Eq. 3.116 and from the fact that, if we have two independent random variables X_1 and X_2, such that $X = X_1 + X_2$, variances add up, i.e., $\langle \Delta X^2\rangle = \langle \Delta X_1^2\rangle + \langle \Delta X_2^2\rangle$, we obtain

$$\langle \Delta J^2_{\rm right}\rangle \quad = \quad \langle \Delta J^2\rangle + \sum_{k=1}^{i} \langle (\Delta(\Delta n_k))^2\rangle. \tag{3.121}$$

We can rewrite

$$\Delta N = \sum_{k=1}^{i} \Delta n_k + \sum_{k=i+1}^{\text{bin total}} \Delta n_k, \tag{3.122}$$

where, "bin total" is the total number of bins we have. Using this result and the additivity property of variances for the sum of two independent random variables $X = X_1 + X_2$ mentioned above, we can also write

$$\langle(\Delta(\Delta N))^2\rangle = \sum_{k=1}^{i}\langle(\Delta(\Delta n_k))^2\rangle + \sum_{k=i+1}^{\text{bin total}}\langle(\Delta(\Delta n_k))^2\rangle \Rightarrow \tag{3.123}$$

$$\sum_{k=i+1}^{\text{bin total}}\langle(\Delta(\Delta n_k))^2 = \langle(\Delta(\Delta N))^2\rangle - \sum_{k=1}^{i}\langle(\Delta(\Delta n_k))^2\rangle, \tag{3.124}$$

where $\langle\Delta(\Delta N)^2\rangle$ is the variance in $\langle\Delta N\rangle$ and likewise for $\langle\Delta(\Delta n_k)^2\rangle$. If we also take the variances of Eq. 3.117 we are only left with the terms:

$$\langle\Delta J^2_{\text{left}}\rangle = \langle\Delta J^2\rangle + \sum_{k=i+1}^{\text{bin total}}\langle(\Delta(\Delta n_k))^2\rangle \tag{3.125}$$

which, after substitution of the above result for $\sum_{k=i+1}^{\text{bin total}}\langle(\Delta(\Delta n_k))^2\rangle$ gives:

$$\langle\Delta J^2_{\text{left}}\rangle = \langle\Delta J^2\rangle + \langle(\Delta(\Delta N))^2\rangle - \sum_{k=1}^{i}\langle(\Delta(\Delta n_k))^2\rangle. \tag{3.126}$$

Using Eq. 3.121 and Eq. 3.126, we obtain that

$$\frac{\langle\Delta J^2_{\text{right}}\rangle + \langle\Delta J^2_{\text{left}}\rangle}{2} - \frac{\langle\Delta N^2\rangle}{2} = \langle\Delta J^2\rangle$$

where, in the above, we replace $\langle(\Delta(\Delta N))^2\rangle$ by $\langle(\Delta N)^2\rangle$ as $\langle\Delta N\rangle = 0$, as verified experimentally (see Section 3.3.14). In short, what this means is that

$$\begin{aligned}\langle\Delta J^2\rangle &= \frac{\langle\Delta J^2_{\text{right}}\rangle + \langle\Delta J^2_{\text{left}}\rangle}{2} \\ &- \underbrace{\frac{\langle\Delta N^2\rangle}{2}}_{\text{correction}}.\end{aligned} \tag{3.127}$$

This exercise is just to show that the actual value of $\langle(\Delta J)^2\rangle$ is related to the average values obtained starting from the left and the right of the chamber plus a certain correction related to the discrepancy in the number of particles from frame to frame. We first obtain $\langle\Delta J^2_{\text{right}}\rangle$ and $\langle\Delta J^2_{\text{left}}\rangle$ for the flux calculated starting from the right and left, respectively. We then plot these fluxes vs. the total number of particles in the two adjacent bins, N (see Figures 3.16 and 3.17). Even though these plots will be described in great detail in Section 3.4, it is worth noting here for the sake of completeness of this section that it can be seen that the plots obey almost linear equations, as is expected from theory. If we take the average of $\langle\Delta J^2_{\text{right}}\rangle$ and $\langle\Delta J^2_{\text{left}}\rangle$, we obtain a resulting plot,

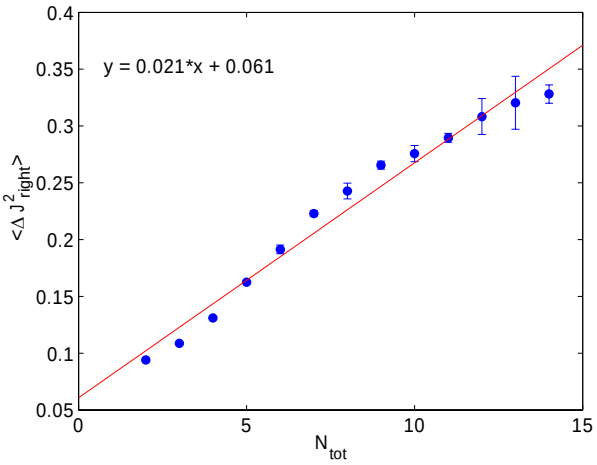


Figure 3.16: The fluctuations in the flux $\langle \Delta J^2_{\text{right}} \rangle$ plotted against the total number of particles between two consecutive bins, N. For this plot, which was obtained via the analysis of a data set with a high concentration of particles (around 300), the flux is calculated starting from the right side of the chamber, which is empty at $t = 0$. The solid line is the fit to the data, which are shown as blue dots. As expected from theory, there is a linear relation between the two variables. However, the slope of this plot is found to be about an order of magnitude higher than the expected value of 0.0022 (see Section 3.4.1).

as shown in Figures 3.18. The slope of this plot is around 0.004, which is much closer to the slope obtained from the simulation where the number of particles is conserved (see Figure 3.34 where the slope is 0.0028), while the intercept is around 0.23. The value of $\langle \Delta N^2 \rangle / 2$ (correction term in Eq. 3.127), as seen from the experiment is around 0.2, which is almost the same as the value of the intercept. Therefore, by subtracting this correction term off, we get a plot that agrees with the theoretical predictions, as will be described in Section 3.4. Thus, using this not so rigorous procedure, we can obtain a better $\langle \Delta J^2 \rangle$ vs. N curve, with a slope and an intercept that are of the same order of magnitude as those obtained by the theoretical predictions.

3.4 Results

In this section we present a detailed account of all the experimentally obtained results for the cases when the initial particle distribution (time $t = 0$) is a step function. It is important to note that the same experiments described below were also performed in the equilibrium setting, i.e., when the whole microfluidics chamber is filled with particles. The results obtained under the equilibrium setting were similar to those presented in this section. However, the equilibrium measurements were not shown, since the local concentration gradient was obviously rather small, thus limiting the range

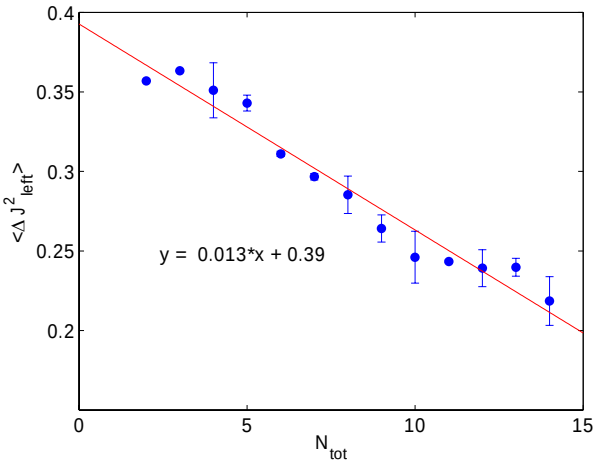


Figure 3.17: The fluctuations in the flux $\langle \Delta J^2_{\text{left}} \rangle$ plotted against the total number of particles between two consecutive bins, N. For this plot, which was obtained via the analysis of a data set with a high concentration of particles (around 300), the flux is calculated starting from the left side of the chamber, which is filled with particles at $t = 0$. The solid line is the fit to the data, which are shown as blue dots. There is still a linear relation between the two variables but now the slope of this plot is found to be negative and still about an order of magnitude higher than the expected value of 0.0022 (see Section 3.4.1).

of ΔN values.

We start off with a reminder of what the theoretical predictions are. Then we move on to present data for experiments done at high (about 300 particles) and low (about 100 particles) concentrations. We also compare these results with the ones obtained via simulations where the number of particles is both conserved and changes from frame to frame. Finally, we present the results for the experiments done at the optimized concentration of around 200 particles.

3.4.1 Theoretical Expectations

Before we present all the experimental results, it is important to summarize here all the predictions made by the theory, as were presented in Section 3.1. In that section, we made the following predictions

1. The principal prediction of the model is that the particle flux distribution function is Gaussian. Based on Eq. 3.2, we can obtain the probability that the flux is between J and $J + \Delta J_{\text{bin}}$ by multiplying that equation with ΔJ. We have defined that incremental value in the flux as ΔJ_{bin} because it is the bin size used to obtain the histogram in the $P(J)$ plots that will follow. This value is equal to $\Delta J_{\text{bin}} = 0.1\,\text{s}^{-1}$, i.e., 1 particle per unit time. Here, for simplicity, we

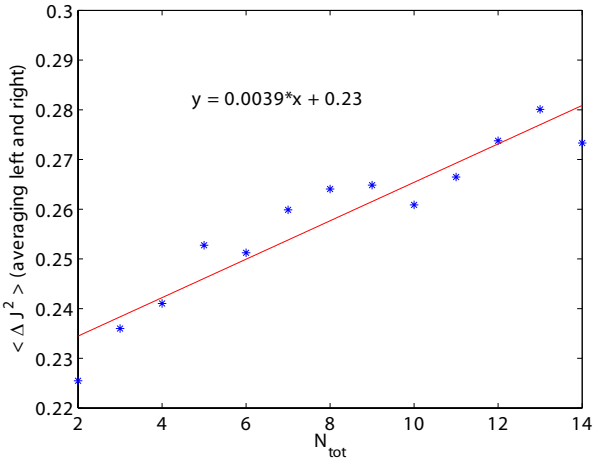


Figure 3.18: The fluctuations in the flux $\langle \Delta J^2 \rangle$ plotted against the total number of particles between two consecutive bins, N. This graph is the average of Figures 3.16 and 3.17, where the flux was evaluated starting from the right and the left of the chamber, respectively. The solid line is the fit to the data, which are shown as blue asterisks. As was theoretically predicted, the relation is linear and now the slope (0.004) is very close to that obtained from the simulation 0.0028 (Figure 3.34) and the expected value of 0.0022. The intercept (0.23) is very close to the value $\langle \Delta N^2 \rangle / (2 * 100) = 0.20$, the correction term in Eq. 3.127 (the factor 100 is the square of $\Delta t = 10$ s).

will just take this value to be $\Delta J_{\rm bin} = 1$ (i.e., we are not dividing by $\Delta t = 10$ s). Therefore,

$$P(J, J + \Delta J_{\rm bin}) = \frac{\Delta J_{\rm bin}}{\sqrt{2\pi \langle (\Delta J)^2 \rangle}} \exp(-\frac{(J - \langle J \rangle)^2}{2\langle (\Delta J)^2 \rangle}). \tag{3.128}$$

So, essentially, the above expression can be rewritten as:

$$\ln(P(J)\sqrt{\langle (\Delta J)^2 \rangle}) = \ln(\Delta J_{\rm bin}/\sqrt{2\pi}) - (\frac{J - \langle J \rangle}{\sqrt{2\langle (\Delta J)^2 \rangle}})^2. \tag{3.129}$$

If we define

$$x = \frac{J - \langle J \rangle}{2\langle (\Delta J)^2 \rangle} \tag{3.130}$$

$$y = \ln P(J) + \frac{1}{2} \ln(\langle (\Delta J)^2 \rangle), \tag{3.131}$$

then the above equation takes the form of:

$$y = -x^2 + \ln(\Delta J_{\rm bin}/\sqrt{2\pi}). \tag{3.132}$$

Therefore, we see that the theory predicts that the coefficient of the quadratic term should

be -1, of the linear term should be 0 and the constant term should be $\ln(1/\sqrt{2\pi}) = -0.92$. (Note: For the case of the results obtained at high concentration using a time step of $\Delta t = 10$ s, we used $\Delta J_{\text{bin}} = 0.2$ s^{-1}, i.e., 2 particles per unit time. In this case the constant term is approximately -0.22 (see Section 3.4.2).)

2. The bad actor trajectories can be predicted. Based on Eq. 3.9, we see that the fraction of bad actors should be linearly proportional to $\langle J\rangle/\sqrt{\langle(\Delta J)^2\rangle}$, for small values of that quantity. The theory also predicts that the coefficient of the linear term is $1/\sqrt{2\pi} \approx 0.4$ and the intercept should be 0.5. Because of the relations between J and ΔN and $\langle(\Delta J)^2\rangle$ and N, the fraction of bad actors should also be proportional to $\Delta N/\sqrt{N}$. In this case the coefficient of the linear term should be $\frac{p}{\sqrt{2\pi p(1-p)}} = 0.27$ (because in Section 3.3.13, we found that $p = 0.32$) while the intercept is still 0.5.

3. The flux fluctuation theorem predicts the ratio of probabilities of fluxes in the forward and backward directions. From Eq. 3.12 we see that $\ln \frac{P(J)}{P(-J)}$ is proportional to $\frac{J\langle J\rangle}{\langle(\Delta J)^2\rangle}$ and the constant of proportionality is equal to 2. Equivalently, $\ln \frac{P(J)}{P(-J)}$ is proportional to $J \times \Delta t$ with a constant of proportionality equal to $\frac{2\Delta N}{(1-p)N}$ (see Eq. 3.11). Thus, this constant is itself proportional to $\Delta N/N$ with a slope of $\frac{2}{1-p} = 3$, since in Section 3.3.13, we found that $p = 0.32$.

4. Fick's Law holds even in the limit of small particle numbers. As we saw from Eq. 3.3, $\langle J\rangle$ is proportional to ΔN and the slope should be equal to $p/\Delta t$. For our experiments, it is important to remember that the time interval is $\Delta t = 10$ s. Also, in Section 3.3.13, we found that $p = 0.32$. Therefore, the slope is expected to be 0.032 and the intercept 0.

5. The second moment of particle flux is proportional to the sum of particle numbers. As was seen in Eq. 3.4, $\langle(\Delta J)^2\rangle$ is proportional to N with the proportionality constant equal to $\frac{p(1-p)}{\Delta t^2}$. Since $p = 0.32$ and $\Delta t = 10$ s, then the slope is equal to 0.0022. The intercept should be 0.

6. The third moment of particle flux is given by Eq. 3.5, which says that $\langle J^3\rangle$ is equal to $3\langle(\Delta J)^2\rangle\langle J\rangle + \langle J\rangle^3$.

In the sections that follow, we describe the results we obtained via simulations and experiments and describe how these compare to the theoretical predictions.

3.4.2 High Concentration Data Sets

The results that correspond to this section were obtained from experiments with high particle concentration (approximately 300 particles were in the microfluidics chamber). For these experiments, the mean particle distance was smaller than 8 μm and, therefore, interactions cannot be neglected (see Sections 3.3.6 and 3.3.8). Also, for such a high concentration it is expected that the error in

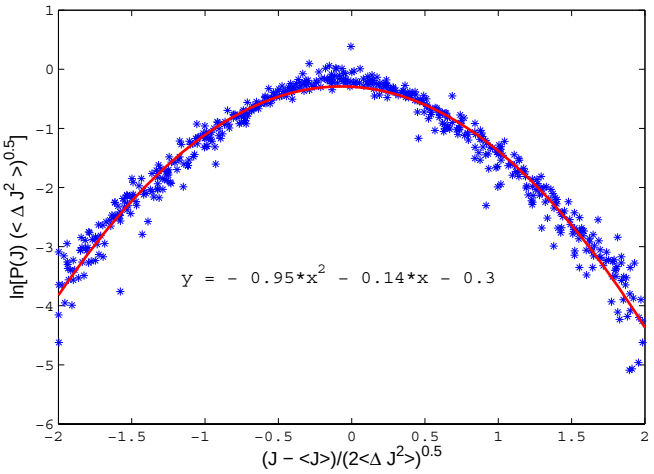


Figure 3.19: The flux distribution function for the high concentration data sets. The graph shows $\ln(\sqrt{\langle(\Delta J)^2\rangle}P(J))$ plotted against $(J-\langle J\rangle)/\sqrt{2\langle(\Delta J)^2\rangle}$. Experimental data is shown in dots, while the solid line represents the quadratic fit to the data. The coefficients of the quadratic equation we get from the fit do not match the values predicted by the theory (see text for details).

the particle number will be greater, since the probability that two particles will overlap increases. In what follows we show just how these results are affected by the high concentration by comparing them with the theoretical expectations. For the figures presented here, neither the experimental correction of multiple images (Section 3.3.14.3) nor the theoretical subtraction (Section 3.3.14.4) have been applied.

The first prediction made is that the particle flux should follow a Gaussian distribution shown in Eq. 3.2. As described in Eq. 3.129, we plot $\ln(\sqrt{\langle(\Delta J)^2\rangle}P(J))$ vs. $(J-\langle J\rangle)/\sqrt{2\langle(\Delta J)^2\rangle}$ in Figure 3.19. As was described in Section 3.2.6.3, we obtained this type of plot by using the different flux distributions corresponding to different values of N_1 and N_2. From Figure 3.19 we see that they all fall on a single master curve, where $\langle J\rangle$ and $\langle(\Delta J)^2\rangle$ have been calculated separately from each combination of N_1 and N_2. On this plot, as predicted, the curve has a quadratic form. The best fit curve to the experimental data yields the coefficient to the square term to be 0.95 (close to the predicted value of 1), while the prefactor for the linear term is -0.14 and the constant term is -0.3. For these last two values the theory predicts 0 and -0.22, respectively, a discrepancy of 14% for the former and 36% for the latter case. Thus, the distribution function does not agree with the one expected from the theory. This discrepancy, however, is to be expected since, as was described above, in this case of the high particle concentration, both the error in the particle number and the interactions between the beads are significant.

Similar discrepancies are observed for the other quantities. The fraction of bad actors should be linearly proportional to $\Delta N/\sqrt{N}$. Figure 3.20 shows that the experiments come very close to verifying this prediction of linearity. Also as predicted, the intercept of the graph is 0.5. This means

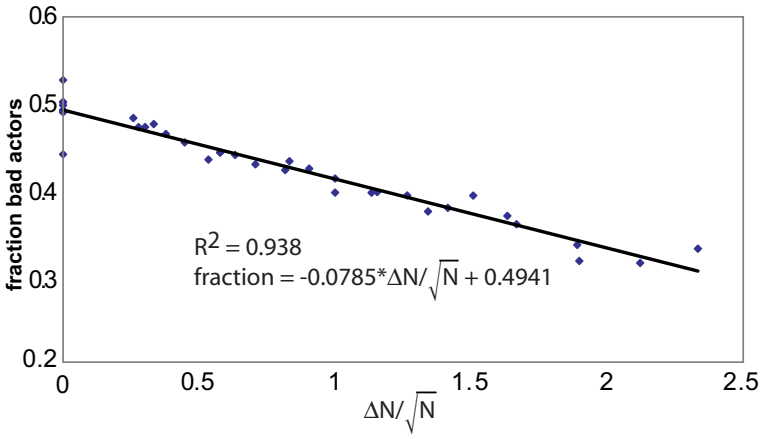


Figure 3.20: Fraction of trajectories that are bad actors vs. the deviation from equilibrium characterized by the quantity $\Delta N/\sqrt{N}$ for the high concentration data sets. Experimental data is shown in dots, while the solid line represents the linear fit to the data. The graph shows the linear dependence predicted by the theory but the slope does not match the expected value of 0.27.

that as the system approaches equilibrium (i.e., $N_1 \approx N_2$), about half the trajectories involved flow down the vanishingly small gradient and half the trajectories involved flow up the same small gradient. However, the slope of 0.078 is an order of magnitude away from the expected value of 0.27. Again, as will be shown in the later sections (see Section 3.4.5), these differences between the experimentally obtained and theoretically expected values are attributed to the high particle density and the fact that the error in the particle number has not been corrected for.

In the inset of Figure 3.21 we plot the fluctuational quantity $\ln(P(J)/P(-J))$, introduced via Eq. 3.11, for the example where $N_1 = 8$ and $N_2 = 3$, vs. the microscopic flux $J \times \Delta t$. Our data shows the same linear behavior for this prediction of the fluctuation theorem as shown in Eq. 3.11. According to the theoretical model, the slope of this curve is expected to be proportional to the quantity $\Delta N/N$. When we plot the different values of slopes obtained for various combinations of N_1 and N_2 in Figure 3.21, we see the linear dependence on the driving force $\Delta N/N$ that is expected from the model (see Section 3.4.1). However, the slope is not equal to 3, as was expected. In fact, the value of 0.42 is again an order of magnitude away from the theoretically predicted value. As was explained before, we will see that as soon as we move to data sets with lower particle concentrations, all these discrepancies disappear, mainly because the error in the particle number is reduced and corrected for.

After comparing the experimentally realized dynamical distribution function (see Figure 3.19) to the distribution predicted by the theoretical model (see Eq. 3.129), we now turn to the different moments for the flux J and compare them to the predictions in Eqs. 3.3, 3.4, and 3.5 (see Section 3.4.1). We show a few of the lowest moments here because of the insights they give. We first consider the odd moments. Figure 3.22 shows the experimentally obtained mean flux, $\langle J \rangle$, as a function of ΔN.

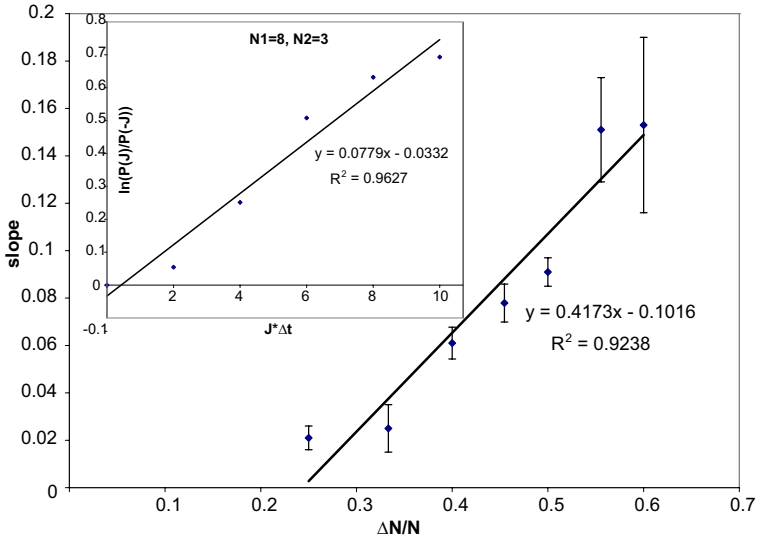


Figure 3.21: The flux fluctuation theorem for the high concentration data sets. The inset shows the plot of $\ln(P(J)/P(-J))$ vs. $J \times \Delta t$ for $N_1 = 8$ and $N_2 = 3$. A straight line fits through the experimental points (shown in dots), which suggests an experimental verification of the fluctuation theorem. We then plot all the slopes of the $\ln(P(J)/P(-J))$ vs. $J \times \Delta t$ graphs for different values of N_1 and N_2, vs. $\Delta N/N$. Again, experimental data is shown in dots, while the solid line represents the fit to the data. The plot is linear but the value of this slope does not match the theoretical expectation.

In this and all the equivalent graphs presented in the following sections, we find the mean flux $\langle J \rangle$ for a given N_1 and N_2. We average over those values of $\langle J \rangle$ that correspond to a given ΔN, since there are many possible values of N_1 and N_2 that give a particular $\Delta N = N_1 - N_2$. We find that $\langle J \rangle$ depends linearly on the particle number gradient ΔN, even down to differences as small as a single particle, indicating that Fick's Law holds in the small numbers limit. The slope of this graph is also, as expected, equal to 0.032 and the intercept is essentially 0. Eq. 3.5 predicts that $\langle J^3 \rangle$ should equal $3\langle \Delta J^2 \rangle \langle J \rangle + \langle J \rangle^3$. Figure 3.23 shows that this is indeed the case, as the relation between the two variables is linear and the slope is 1. From this and the previous Figure 3.22 we see that the odd moments of the high concentration data sets seem to be unaffected by the fact that the error in the particle number is greater, as we explained in Section 3.3.14.

It may be noted that till now we discussed only the odd moments of flux. In what follows, we look at the lowest even moment of flux, i.e., the second moment, and see its compatibility with the theoretical prediction (Section 3.4.1). The second moment should be proportional to N. Figure 3.24 qualitatively confirms that prediction of the linear dependence of $\langle \Delta J^2 \rangle = \langle (J - \langle J \rangle)^2 \rangle$ on the total number of particles N. In this and the equivalent figures presented in the following sections, for a given N_1 and N_2 we find $\langle (J - \langle J \rangle)^2 \rangle$. We then average over all these values that correspond to the same N, since, for a given N, there are many possible values of N_1 and N_2 such that $N_1 + N_2 = N$.

It is worth mentioning that the last two data points do not have as good statistics when compared

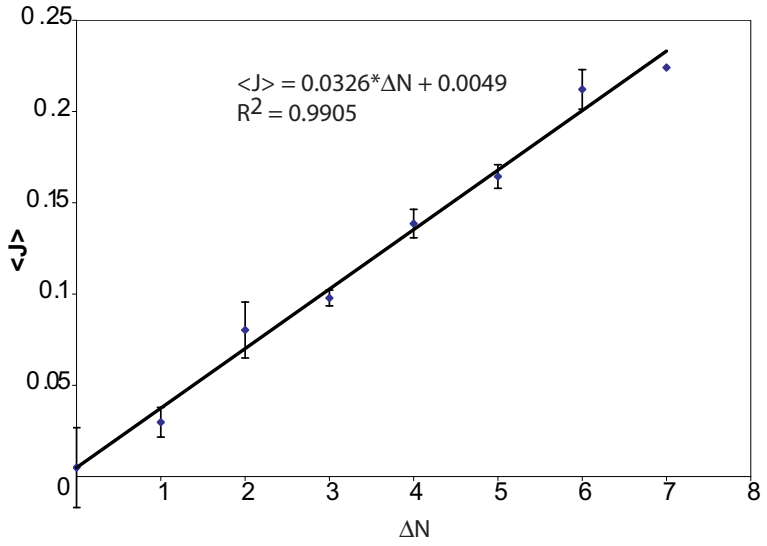


Figure 3.22: Experimental support for Fick's law for the high concentration data sets. The average flux, $\langle J \rangle$, is plotted as a function of ΔN, the gradient in the particle number in the two neighboring bins. Experimental data is shown in dots, while the solid line represents the fit to the data. Even though the total number of particles is small, the plot is clearly linear, thus verifying Fick's law for the mean flux. The error bars shown are the variances due to the different combinations of N_1 and N_2 resulting in the same ΔN. The value of the slope is equal to the expected value of 0.032. The intercept is almost 0.

-

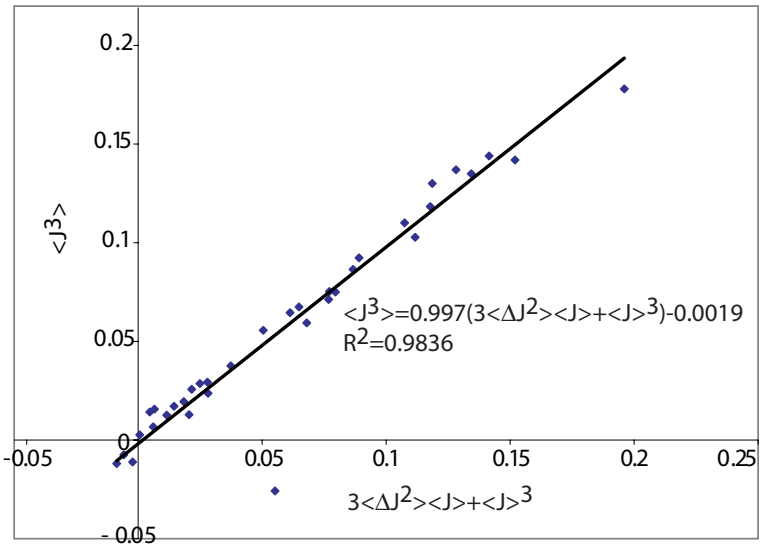


Figure 3.23: The third moment of the flux $\langle J^3 \rangle$ with $\langle 3\Delta J^2 \rangle \langle J \rangle + \langle J \rangle^3$ for the high concentration data sets. Experimental data is shown in dots, while the solid line represents the fit to the data. This plot agrees well with the experimental prediction given by Eq. 3.5, since the graph is linear and the slope, within the experimental error, is 1.

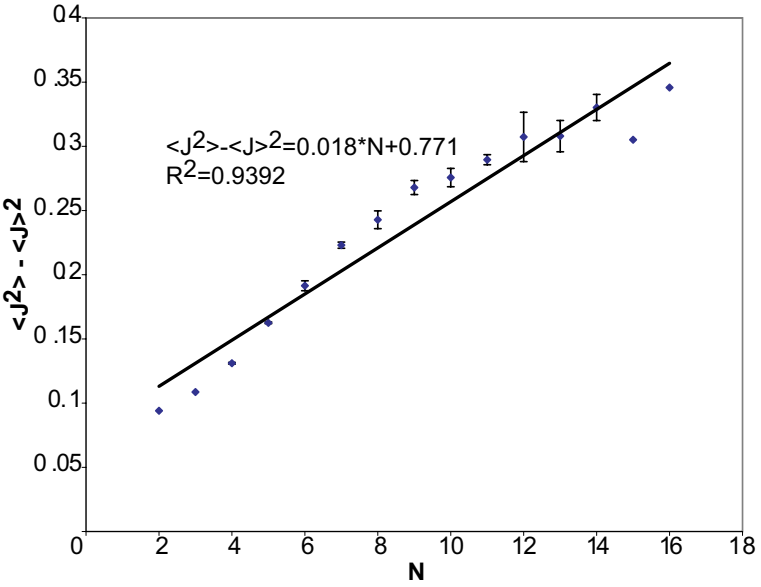


Figure 3.24: The second cumulant, $\langle \Delta J^2 \rangle = \langle J^2 \rangle - \langle J \rangle^2$, vs. the total number of particles, N for the high concentration data sets. Experimental data is shown in dots, while the solid line represents the fit to the data. The second cumulant in flux is almost linearly proportional to N. The error bars shown are the variances due to the different combinations of N_1 and N_2 resulting in the same N. If we remove the last two points with low statistics the linearity will be more pronounced. The slope, however, is an order of magnitude higher than the expected value and the intercept is not 0.

to the remaining points. If we remove those data points in Figure 3.24, we get better least-square fit to a straight line and the intercept is closer to zero. Nevertheless, the curve is not quite a straight line, as it has a "sigmoidal" shape, and the slope is an order of magnitude higher than expected. In addition, the intercept is not close to 0. As has been explained, this huge discrepancy in the value of the slope between the experimental result and the theoretical prediction is due to the error in the particle number for these high concentration data sets. Also, as was explained in Section 3.3.14, this error did not affect the odd moments, as was seen from Figures 3.22 and 3.23, where the slopes agree with the theoretical expectations.

So, we have seen that the high concentration data sets provided results that are in good agreement with the theoretical predictions qualitatively. However, the slopes for the graphs of the bad actors, the fluctuation theorem and the second moment did not agree with the theoretical expectations. It will become clearer later on why the error in the particle number is responsible for this discrepancy.

3.4.3 Low Concentration Data Set

We claimed that the strange shape of the $\langle \Delta J^2 \rangle$ versus N curve is due to the error in the number of particles from frame to frame. When the total number of particles is large the probability that the particles overlap and, hence, remain untracked increases. Thus, it is expected that if we use a low concentration of particles it is very likely that this error will be reduced. In Section 3.4.2, the

total number of particles is around 300. In this section we will analyze the data set where the total number of particles is around 80. In this case the error in the number of particles was about 2%. In what follows it will be seen that the discrepancy between the theory and the experimental results reduces quite a bit for the low concentration data. As was the case before, for the figures presented here, neither the experimental correction of multiple images (Section 3.3.14.3) nor the theoretical subtraction (Section 3.3.14.4) have been applied.

For this particular data set, however, we do a different but simple correction before we begin the calculation of the mean flux $\langle J \rangle$, the fluctuations in the flux, etc. The total number of frames, at a rate of 6 frames per minute, in the data set is approximately 2000. We first evaluate the total number of particles $N_{\rm tot}(i)$ for each frame i and then take the difference $\Delta N(i) = N_{\rm tot}(i) - N_{\rm tot}(i-1)$. If the absolute value of this difference $|\Delta N(i)|$ is larger than the variance $\sigma_{\Delta N}$ over the entire 2000 frames, we discard the flux obtained form such frames. Therefore, as will be seen in this section in the figures that follow, there is a significant improvement in the agreement between the experimentally obtained and the theoretically predicted quantities of interest when compared with the case of the high concentration data sets. However, discrepancies are still there since the error in the particle number has not been completely corrected for.

With this procedure of discarding frames with high $|\Delta N(i)|$, we see in Figure 3.25 the plot of $\langle J \rangle$ versus ΔN, the experimental support of Fick's law. It can be seen that the plot is linear with a slope of 0.027, a discrepancy of only 16% from the expected value of 0.032. In addition, we see that the intercept is almost 0. Thus, just like in the case of the high concentration data sets, the first moment for the flux matches the predictions made by the theoretical model. The plot of the troublesome second cumulant, the fluctuations in the flux $\langle \Delta J^2 \rangle$ vs. N, is shown in Figure 3.26. It can be seen that the data points lie in a straight line that does not have the "sigmoidal" shape found in the high concentration data (Figure 3.24), though the slope (0.0049) is not quite what is expected (0.0022). It is off by a factor of 2, which is, nevertheless, a big improvement over the value obtained in the high concentration case, where the slope was bigger by an order of magnitude (Figure 3.24). In addition, even though the intercept has also improved, it is still not close to 0. Therefore, the low concentration of particles did reduce the error in the particle number, just as expected. However, it is still not completely eliminated (since it has not been corrected for), which is why these discrepancies are observed.

Improvements over the high concentration case are also seen in the remaining quantities of interest. The plot for the "fraction bad-actors" vs. $\Delta N/\sqrt{N}$ is linear, as can be seen from Figure 3.27. The intercept is still 0.5, as expected, but the slope, even though it has improved from the high concentration case (see Figure 3.20), is still not close to the expected value of 0.27. Again, it is off by, approximately, a factor of 2 due to the 2% error in the particle number. As suggested in Eq. 3.12, we also plot the flux fluctuation theorem, the master curve for $\ln(P(J)/P(-J))$ vs.

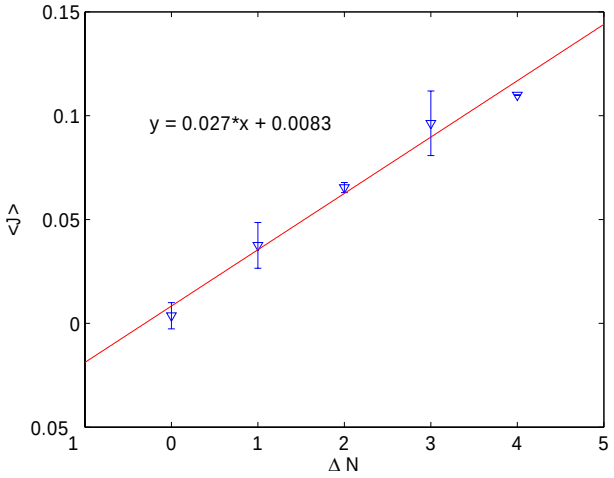


Figure 3.25: Experimental support of Fick's law for the low concentration data sets. The average flux $\langle J \rangle$ is plotted vs. ΔN, the difference in the number of particles between two consecutive bins. Experimental data is shown in triangles, while the solid line represents the fit to the data. The error bars shown are the variances due to the different combinations of N_1 and N_2 resulting in the same ΔN. The slope is around 0.027 and deviates by only 16% from the theoretically predicted value. The intercept is almost 0, as expected from theory. The discrepancies are still due to the fact that even though the error in the particle number has been reduced, it is still not corrected for.

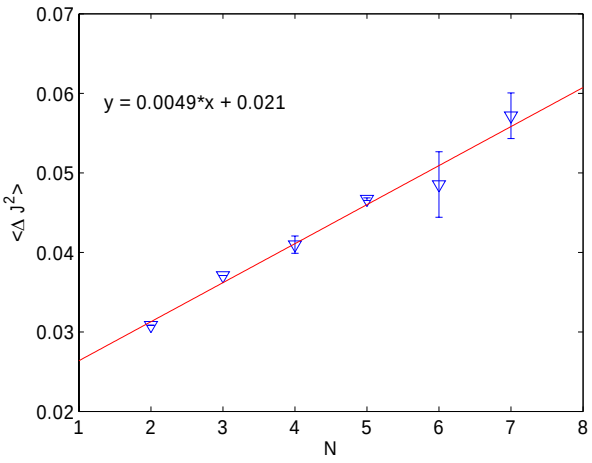


Figure 3.26: The plot of the second cumulant $\langle \Delta J^2 \rangle$ vs. N for the low concentration data set. Experimental data is shown in triangles, while the solid line represents the fit to the data. The error bars shown are the variances due to the different combinations of N_1 and N_2 resulting in the same N. The slope is around 0.0048, the same order of magnitude as the theoretically expected value (0.0022) but the intercept is still not 0. The discrepancy is there due to the error in the particle number. However, most importantly, we recover back the linear nature of the plot.

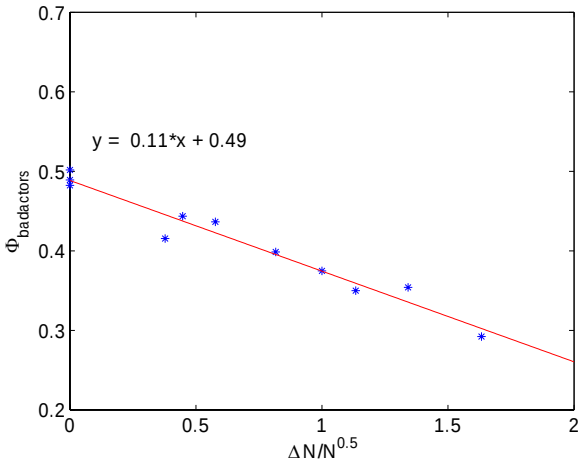


Figure 3.27: Fraction of trajectories that are bad actors vs. the deviation from equilibrium characterized by the quantity $\Delta N/\sqrt{N}$ for the low concentration data set. The data points are shown as asterisks while the solid line is the fit to these data points. The plot is linear and the intercept is 0.5, as expected. However, the slope is off by about a factor of 2 from the value predicted by the theory, the discrepancy again being due to the error in the particle number.

$\langle J\rangle/\langle\Delta J^2\rangle$. It can be seen from Figure 3.28 that the plot shows a linear trend, as is expected from Eq. 3.12. There are quite a few stray points, but the majority of them fall on a straight line. Also, the slope of 1.7 deviates by only 15% from the expected slope of 2. Finally, we plot the overall probability distribution function from Eq. 3.129 as shown in Figure 3.29. We use $\Delta J = 1$ particle, i.e, $\Delta J = 1/\Delta t = 0.1$/s for binning the histogram. It can be seen that a quadratic can be nicely fit to the data set, with the coefficient of the quadratic term almost equal to 1, the coefficient with the linear term is almost 0, while the constant term is almost equal to 0.92 (see Section 3.4.1).

The most troublesome aspect for the high concentration case was the second cumulant, the $\langle\Delta J^2\rangle$ vs. N plot with its funny shape and high slope. It was seen here that the same plot is linear with this low concentration data set (see Figure 3.26). We also saw in Section 3.3.14.4 that the correction in this plot can be obtained by calculating the flux starting once from the right and once from the left of the microfluidics chamber. Using Eq. 3.127, i.e., taking the mean of $\langle\Delta J^2_{\text{right}}\rangle$ and $\langle\Delta J^2_{\text{left}}\rangle$ and subtracting the mean square particle discrepancy $\langle\Delta N^2\rangle$, we obtain a very good result for this plot (as shown in Figure 3.18). When the same procedure is repeated on this low concentration data set, we obtain a very good result as shown in Figure 3.30. The slope (0.0034) is much closer to the expected value (0.0022), and the intercept is essentially 0. So, by comparing Figures 3.26 and 3.30 we see how the theoretical correction has improved the obtained results.

If this low concentration data is exhibiting the expected behavior, especially after the theoretical

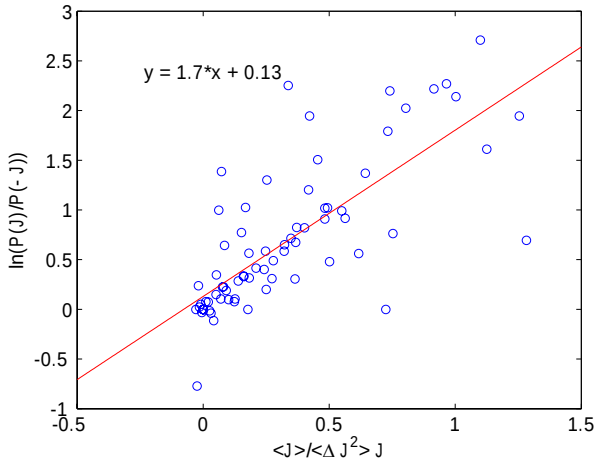


Figure 3.28: The flux fluctuation theorem, the master plot of $\ln(P(J)/P(-J))$ vs. $(\langle J\rangle/\langle\Delta J^2\rangle)\times J$, in the case of the low concentration data. The solid line represents the fit to the data shown as open circles. The plot is linear as expected but there are stray points. The deviation in the slope from the expected value of 2 is due to the error in the particle number.

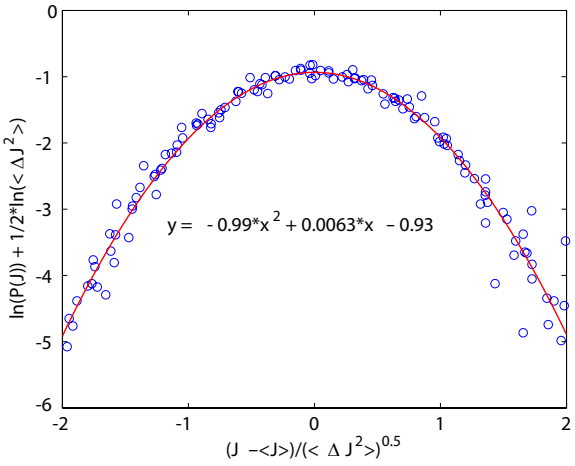


Figure 3.29: The flux probability distribution function given in Eq. 3.129 for the low concentration data. The data points are shown as open circles while the solid line is the fit to the data. It can be seen that the plot is quadratic as expected and the coefficients match almost exactly the predicted values summarized in Section 3.4.1.

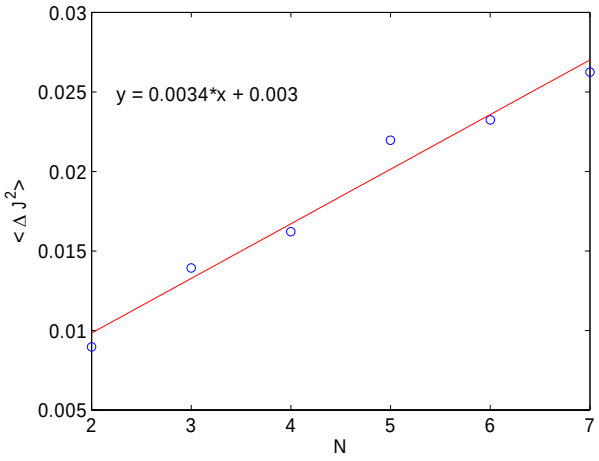


Figure 3.30: The plot for the second cumulant $\langle \Delta J^2 \rangle$ vs. N for the low concentration data after corrections suggested in Section 3.3.14.4. The data points are shown as open circles and the solid line is the fit to those points. It can be seen that the slope (0.0034) has improved when compared to that of Figure 3.26 and is very close to the expected value, and the intercept is almost 0.

correction has been applied, then why not just stop here? The only drawback of using the low concentration data set is that the ranges along the x axis of the figures become very small because of the low concentration. If we compare the x axes of Figures 3.22 and 3.25 and Figures 3.24 and 3.30 we see that for the low concentration data we have small ranges for the values of N and ΔN. Otherwise the low concentration case gives admirable results in pretty much all aspects.

3.4.4 Simulation with Constant and Non-Constant Particle Number

We have observed that the plot of the second cumulant $\langle \Delta J^2 \rangle$ vs. N has a shape that deviates from the expected linear behavior, in the case of the high concentration data, and a slope that is not what is expected based on the theoretical predictions. We attributed this deviation to the change in the total particle number from frame to frame and not to the number of data sets. That is, a greater number of data sets (i.e., statistics) would not result in a more linear behavior. Indeed, when we conducted more experiments for the high concentration case, we saw that, even though the error bars for the individual points decreased (as shown in Figure 3.24), the shape and the slope of the plot remained unchanged. In fact, the only time that the shape and slope changed was when we reduced the bead concentration inside the microfluidic chamber. But why does the concentration affect the shape and slope of the curve?

The answer to this question was given in great detail in Section 3.3.14 so here we will provide a brief summary of the conclusions made. When the concentration changes, so does the mean distance

between the beads. When the concentration is low, then the mean distance increases and, thus, the probability that two beads will get close enough to each other (or overlap) that the IDL algorithm will not be able to distinguish them is low. This was demonstrated by the fact that when we used low bead concentrations, the change in the number of particles from frame to frame was not more than 2-3% (occasionally, this number was higher but that was due to the low frame quality. In that case, those bad frames were disregarded). In contrast, when the bead concentration was high (i.e., higher probability that the beads will get close enough and overlap and, thus, IDL will not be able to tell them apart), this number changed to 5% particle difference from frame to frame. Because of this change in the particle number, we have an error in the value of the calculated flux. Due to the fact that we are using the continuity equation for the evaluation of the flux, this error gets propagated. As a result, whenever a change in the particle number occurs, the flux will be overestimated or underestimated, depending on whether the change in the particle number was positive or negative. However, as was proven by the experiments, this overestimation/underestimation does not affect the values for the odd moments $\langle J \rangle$ and $\langle J^3 \rangle$ and only affects the even moment $\langle \Delta J^2 \rangle$.

In order to further prove that it is this change in the particle number that results in the deviation of the experimental results from the theoretically predicted behavior, we run two simulations. For both simulations we have assumed that the particle concentration is high by considering 295 particles. We assumed that at this concentration the particles are not interacting. Also, in order to be as close to the experimental conditions as possible, we run both simulations so that the total number of "frames" considered is equal to the experimental total number of frames (around 2200, since we had 6 frames per minute for 6 hours of experiment). In addition, in the first simulation we assumed that there is no change in the particle number. In the second simulation, we randomly removed about 5% of the particles from frame to frame. After the simulations were run, we calculated the same quantities as those considered above.

As can be seen in the plots for the odd moments ($\langle J \rangle$ and $\langle J^3 \rangle$, Figures 3.31 and 3.32, respectively) the results of the two simulations were identical and agree with the experimental results, proving that the odd moments are unaffected by this particle change. Also, as expected, the slope and intercept in both simulations exactly match the theoretical predictions. However, it needs to be noted that in the case of the simulation for the third moment, Figure 3.32, the range of values for the ideal simulation (data shown in asterisks, '*'), where the particle number remains constant, is smaller than the range of values for the non-ideal simulation (data shown in crosses, '+'). This is because, as was explained above, the change in the particle number that is assumed in the non-ideal simulation results in the calculation of fluxes that are spurious, which in turn are responsible for the increase of the range of values observed.

The plot of the second cumulant, $\langle \Delta J^2 \rangle$ vs. N, though, is affected by the change in the particle number. From Figure 3.33 we see that the result of the simulation that takes into account the

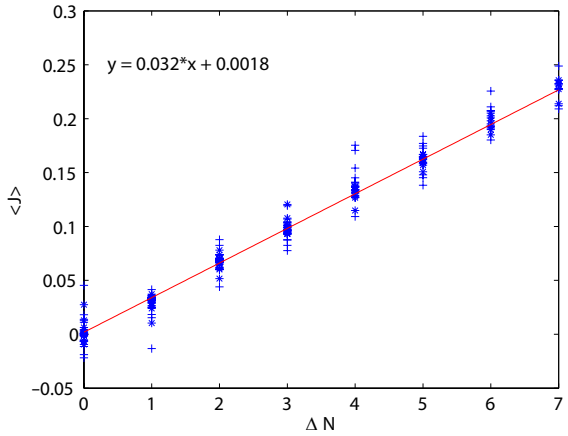


Figure 3.31: Support for Fick's law obtained using simulation. The average flux $\langle J \rangle$ is plotted vs. ΔN, the difference in the particle number between two neighboring bins. The simulation was run with 295 particles both keeping the number of particles constant (data shown in asterisks, '*')and changing them by 5-6% (data shown in crosses '+') from frame to frame. The solid line is the linear fit to the data. It can be seen that the resulting plots are identical and from the linear fit, we can see that the slope is equal to the expected value and the one experimentally obtained in Figure 3.40.

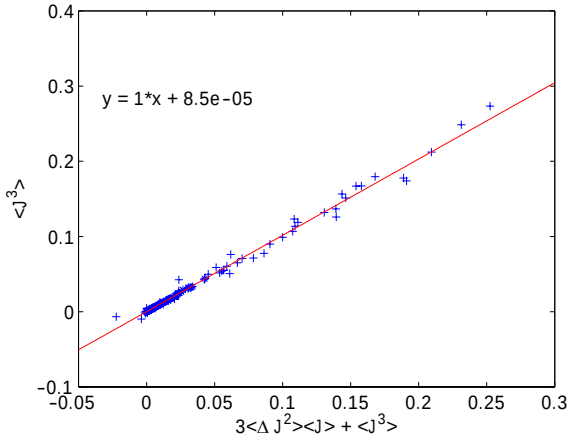


Figure 3.32: The third moment of the flux obtained with simulation. The third moment is shown as a plot of J^3 vs. $3\langle \Delta J^2 \rangle \langle J \rangle + \langle J \rangle^3$. The simulation was run with 295 particles both keeping the number of particles constant (data shown in asterisks '*') and changing them by 5-6% (data shown in crosses '+') from frame to frame. The solid line is the linear fit to the data. The resulting plots have identical slopes (unity). On the other hand, the range for $\langle J^3 \rangle$ from the ideal simulation ('*') is smaller than the non-ideal simulation ('+') (see text for details).

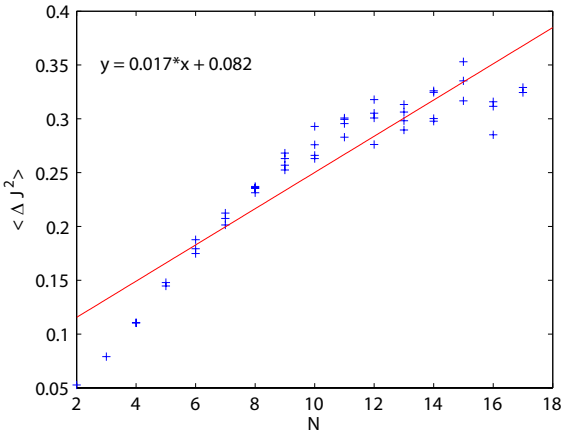


Figure 3.33: The plot of the second cumulant, $\langle \Delta J^2 \rangle$, vs. N, the total number of particles in two neighboring bins, obtained from the simulation, when the total number of particles changes from frame to frame. The total number of particles is 295. We first run the simulation by keeping the particle number constant. After the complete evolution we 'bin' the system and obtain the total number of particles in each bin. Then we randomly start removing the particles from each bin till the total number of particles removed is less than 5-6% of the total number of particles (10-15 out of 295). The flux and the fluctuations in the flux are then obtained by using the continuity equation, Eq. 3.17, described earlier. The resulting data are shown in crosses '+' and the solid line is the linear fit to the data. The plot of $\langle (\Delta J)^2 \rangle$ vs. N shows the same pattern as the experimentally obtained slope in Figure 3.24. Even the slope of 0.017 is almost equal to the one of Figure 3.24, 0.018. Hence, we can conclude that the sigmoidal shape of the curve comes from the fact that the total number of tracked particles changes from frame to frame.

particle change has, just like the experiments, higher slope and a sigmoidal shape. In fact the slope and shape of this figure are identical to those of the equivalent figure for the high concentration case (see Figure 3.24). From Figure 3.34 we see that the corresponding result of the simulation that assumes no change in the particle number is a nice straight line with a slope (0.0028) that deviates from the expected value of 0.0022 by 27%, just like Figure 3.30 for the low concentration experimental data. Also, the intercept, as expected, is close to 0. It is, therefore, safe to conclude that it is due to the change in the particle number, and not due to the number of statistics, that we observe this shape/slope for the $\langle \Delta J^2 \rangle$ vs. N graph.

We had also explained in Section 3.3.14 that, because J can be either positive or negative, when we calculate $\langle J \rangle$ and $\langle J^3 \rangle$ the negative, underestimated fluxes cancel the positive, overestimated ones and, thus, the averages come out unaffected. However, when we calculate the even moment $\langle \Delta J^2 \rangle$, there is no room for cancellation since these Js get squared and, thus, they all become positive. This is why the change in the particle number only affects the even moments of the flux distribution.

For comparison, we also present the results for the case of the fraction of bad actors obtained by the two simulations. Figure 3.35 shows this quantity plotted vs. $\Delta N/N$ for the simulation where the total number of particles remains constant (data shown in asterisks '*'). In this figure, we see the

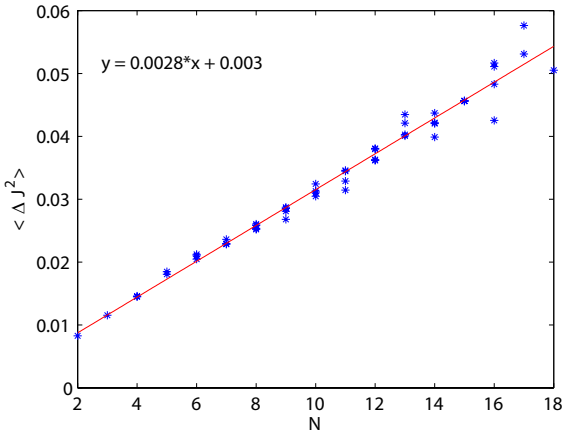


Figure 3.34: The plot of the second cumulant, $\langle \Delta J^2 \rangle$, vs. N, the total number of particles in two neighboring bins, obtained from the simulation when the total number of particles remains constant and equal to 295 from frame to frame. The data are shown in asterisks '*' and the solid line is the linear fit to the data. It can be seen that the curve is a nice straight line. The slope is expected to be around 0.0022. The slope on the other hand, is around 0.0028, a deviation of 27% from the expected value. The intercept is, as expected, close to 0. This plot is very similar to Figure 3.30 for the low concentration data.

expected from theory linear behavior and intercept. However, the slope (0.18) is still smaller than the expected value (0.27), just like in the case of the low concentration data (Figure 3.27, where the slope is 0.11). Figure 3.36, which is the result of the simulation that takes into consideration the change in the particle number, has almost the same slope (0.087) as Figure 3.20 for the high concentration data (0.078). This value of the slope is much different from the expected value of 0.27. Once again, from the two simulations we see that the change in the particle number is responsible for the great deviation between the expected values and the values obtained from the high concentration data sets.

Since the data obtained by the low concentration experiments did not provide big ranges for the values of ΔN and N, we then conducted experiments at the medium concentration of about 200 particles. In this case, as was explained in Section 3.3.14.3, in order to reduce the change in the particle number as much as possible, we took 3 snapshot at each time interval. We also theoretically found what corrections need to be made to the data, as was shown in Section 3.3.14.4. In the section that follows, we show the experimental results obtained with this multiple frame capture method after the theoretical correction has been implemented.

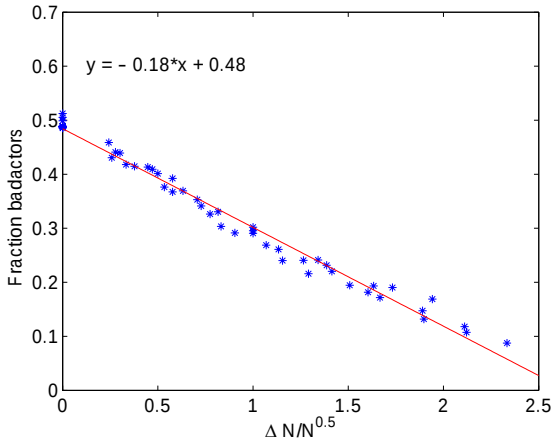


Figure 3.35: Fraction of trajectories that are bad actors vs. the deviation from equilibrium characterized by the quantity $\Delta N/\sqrt{N}$ obtained from the simulation where the particle number is constant and equal to 295 particles. The data are shown in asterisks '*' and the solid line is the linear fit to the data. The slope of this plot (0.18) is different than the one expected from the theoretical result ($\sqrt{p/(2\pi(1-p))} = 0.27$) for the theoretical model. This figure is very similar to the one obtained from the low concentration data, Figure 3.27, where the slope was 0.11.

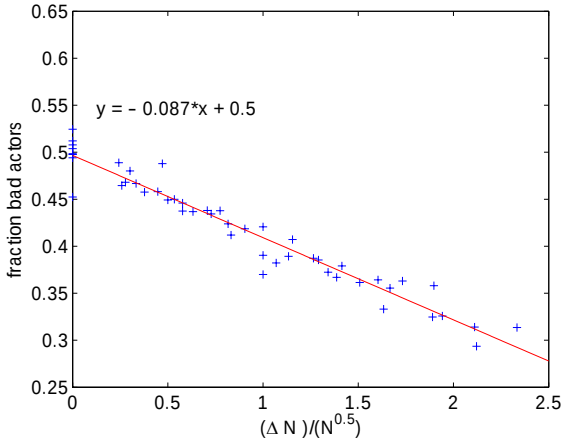


Figure 3.36: Fraction of trajectories that are bad actors vs. the deviation from equilibrium characterized by the quantity $\Delta N/\sqrt{N}$. The plot was obtained using the simulation with 295 particles by changing the number of particles from frame to frame. The data are shown in crosses '+' and the solid line is the linear fit to the data. The slope of this plot (0.087) is almost the same as that obtained from the high concentration experiments (0.078) as was seen in Figure 3.20.

3.4.5 Medium Concentration Data Obtained Using the Method of Multiple Frames

In this section, we are presenting the data that was obtained at a medium particle concentration (approximately 200 particles are trapped in the microfluidics chamber), which is also the optimized particle concentration (see Sections 3.3.6 and 3.3.8) by taking 3 frames at each time interval $\Delta t = 10$ s for a total time of 6 h. As was explained in Section 3.3.14.3, the 3 snapshots were used in order to minimize the error in the particle number to 1-2%, which corresponds to 2-4 particles out of the 200. The snapshots are taken using fluorescent microscopy with a SONY DFW-V500 camera. More details on the data acquisition are given in Section 3.2.5. After the images were obtained and analyzed and the values for the flux J and N and ΔN were found, as was described in detail in Section 3.2.6, the results for $\langle(\Delta J)^2\rangle$ were further corrected by subtracting off the correction term, as was explained in Section 3.3.14.4. As we will see below, the final results agree extremely well with the theoretical predictions summarized above (see Section 3.4.1).

3.4.5.1 The Flux Distribution Function Is Gaussian

Figure 3.37 shows our observed particle flux distribution function at the optimized concentration. All the data falls on a single master curve where $\langle J\rangle$ and $\langle(\Delta J)^2\rangle$ have been calculated separately from each combination of N_1 and N_2. The quadratic form observed on this log plot shows that the distribution function is Gaussian. Just like the theory predicts, (i) the coefficient of the square term is -0.98, very close to -1, (ii) the coefficient of the linear term is -0.0018, close to 0, and (iii) the constant term is -0.94, almost equal to the expected value $\ln(\Delta J_{\text{bin}}/\sqrt{2\pi}) \approx -0.92$, where ΔJ_{bin} is the bin-size used to obtain the histogram and is equal to $0.1\ \text{s}^{-1}$, i.e., 1 particle per unit time. The coefficient of determination for the quadratic fit is $R^2 = 0.98$. Next, we analyze the bad actors—the backward flows—in two different ways.

3.4.5.2 The Bad-Actor Trajectory Counts Are Well Predicted by the Model

From Eq. 3.9 we predicted that, for small values of $\langle J\rangle/\sqrt{\langle(\Delta J)^2\rangle}$, the fraction of bad actors should be linearly proportional to $\langle J\rangle/\sqrt{\langle(\Delta J)^2\rangle}$. In good agreement, Figure 3.38 confirms this linearity and gives the predicted intercept of 0.5. This means that as the system approaches equilibrium (i.e., $\langle J\rangle \approx 0$), about half the trajectories involve flow down the vanishingly small gradient and half the trajectories involve flow up that small gradient. In the linear regime, the best fit line shows the slope to be 0.37, which agrees well with the expected value of $1/\sqrt{2\pi} \approx 0.4$ from Eq. 3.9. The coefficient of determination for the linear fit is $R^2 = 0.99$. Another key feature of this graph is that when the system is farther away from equilibrium (as implied by a larger mean flux), the bad actor fraction reduces. What this means is that more of the microtrajectories available to the system in

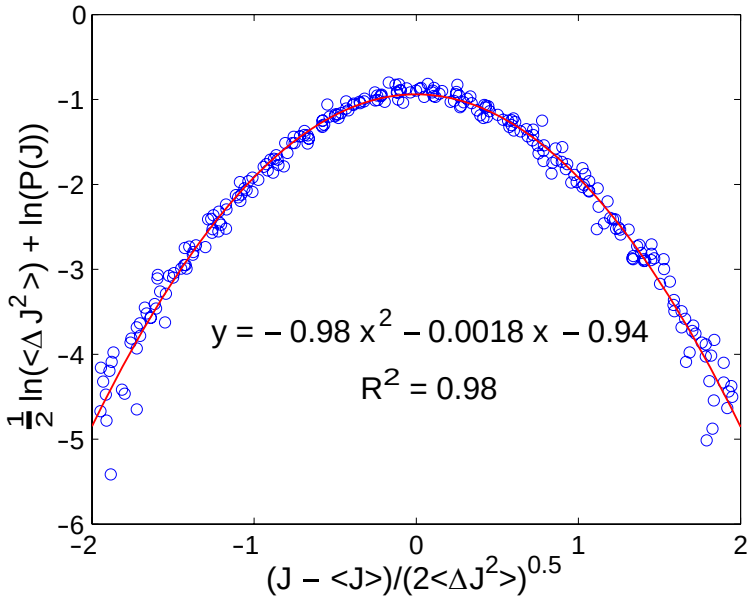


Figure 3.37: The flux distribution function for the case of the optimized particle concentration. $\frac{1}{2}\ln(\langle(\Delta J)^2\rangle) + \ln(P(J))$ is plotted against $(J - \langle J\rangle)/\sqrt{2\langle(\Delta J)^2\rangle}$, based on the form indicated by Eq. 3.129. Circles indicate experimental points, the line shows a quadratic fit to the data. The coefficient of determination (R^2) for the fit is also reported. This demonstrates that the distribution function is Gaussian, and we find that the coefficients are well predicted by Eq. 3.129 (see text for details).

this case are potent to change the current state of the system. Indeed, this idea is quantitatively characterized in the form of the potency [76].

3.4.5.3 The Experiments Confirm the Flux Fluctuation Theorem

Figure 3.39 shows $\ln(P(J)/P(-J))$ vs. the flux, normalized by $\langle J\rangle J/\langle(\Delta J)^2\rangle$, to account for different averages and variances of the flux distribution. This rescaling leads to a linear master curve as predicted by Eq. 3.12. Experiments show the slope to be 2.0, in perfect agreement with the predicted slope of 2 from Eq. 3.12. In this figure, there are a few aberrant data points that we are unable to explain. However, the coefficient of determination is still close to unity, $R^2 = 0.77$. Though ultimately the fluctuation theorem is a reflection of the nature of the flux distribution function, such theorems are a compact way to quantitatively illustrate the significance of bad actor microtrajectories.

3.4.5.4 Fick's Law Holds Even in the Small-Numbers Limit

We compare the average flux between two neighboring bins, $\langle J\rangle$, with the difference in particle numbers, $\Delta N = N_1 - N_2$. This data is compiled from all the values of N_1 and N_2 that provide a given ΔN. Figure 3.40 shows that $\langle J\rangle$ depends linearly on the particle number gradient ΔN, even down to "gradients" of zero or one particle, indicating that Fick's Law holds in the small-numbers

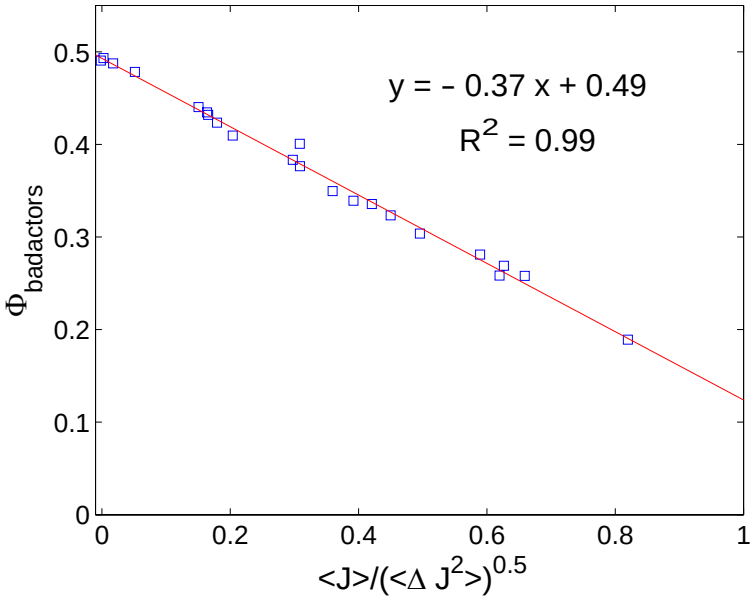


Figure 3.38: The fraction of trajectories that are bad actors vs. the deviation from equilibrium as characterized by the normalized mean flux, $\langle J\rangle/\sqrt{\langle(\Delta J)^2\rangle}$, for the data with the optimized particle concentration. Experimental data is shown in squares, while the solid line represents the fit to the data. The coefficient of determination (R^2) for the fit is also reported. The slope and intercept agree well with the model (see text for details).

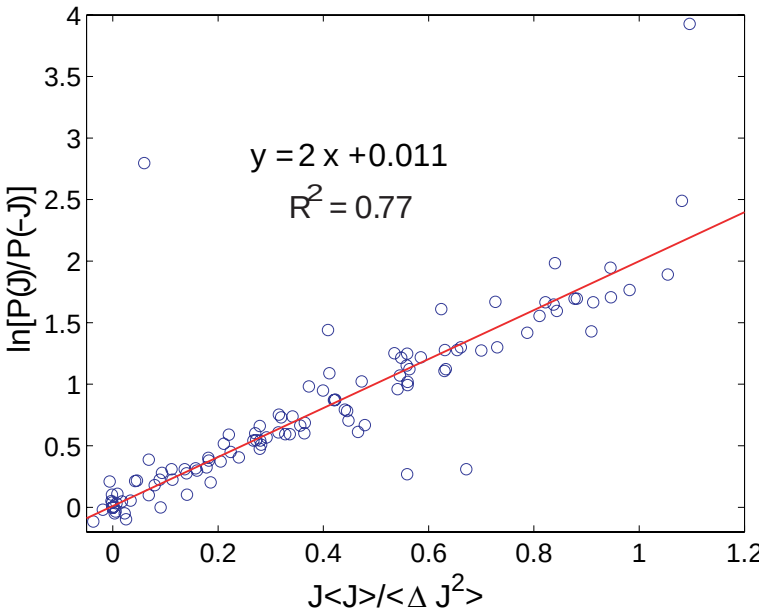


Figure 3.39: The flux fluctuation theorem for the optimized particle concentration data sets. The plot shows $\ln[P(J)/P(-J)]$ vs. $\langle J\rangle J/\langle(\Delta J)^2\rangle$ for different values of $\langle J\rangle$ and $\langle(\Delta J)^2\rangle$ arising due to different combinations of N_1 and N_2. Experimental data is shown in circles, while the solid line represents the fit to the data. The coefficient of determination (R^2) for the fit is also reported. The slope and intercept agree with the prediction of Eq. 3.12 (see text for details).

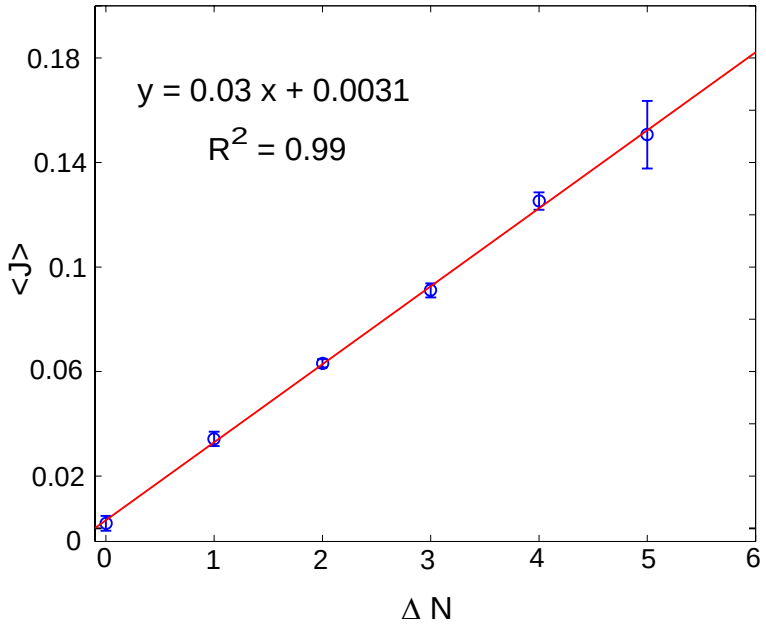


Figure 3.40: Experimental support for Fick's Law, even down to few-particle gradients for the optimized particle concentration data sets. The average flux, $\langle J \rangle$, is shown as a function of ΔN, the gradient in the particle number between two neighboring bins. Experimental data is shown in circles, while the solid line represents the fit to the data. The coefficient of determination (R^2) for the fit is also reported. The error bars shown are the variances due to the different combinations of N_1 and N_2 resulting in the same ΔN. The slope and intercept are in agreement with the expected theoretical values, based on Eq. 3.3 (see text for details).

limit. The slope of the graph (0.03/s) also gives us a value of the jump rate, $p = 0.3$, which is in good agreement with the theoretical estimate of 0.32 made in terms of the bin-size and the diffusion coefficient (see Section 3.3.13). As expected, the intercept is close to 0 (see Eq. 3.3). The linear fit has a coefficient of determination $R^2 = 0.99$.

3.4.5.5 The Second Moment of Particle Flux Is Proportional to the Sum of Particle Numbers

Eq. 3.4 predicts that the second moment of the flux should be proportional to the sum of particle numbers in the two bins, $N = N_1 + N_2$. Figure 3.41 confirms this dependence of $\langle \Delta J^2 \rangle = \langle (J - \langle J \rangle)^2 \rangle$ on N. For the optimized particle concentrations, the slope ($0.0022/\text{s}^2$) is equal to the expected slope of $0.0022/\text{s}^2$ for the value of $p = 0.32$. The coefficient of determination is $R^2 = 0.96$. As has been shown at higher particle concentrations (see Section 3.4.2), however, not surprisingly, systematic errors begin to appear and the slope deviation is quite high compared to the expected value. We performed Brownian dynamics simulations that show the likely cause of these concentration-dependent errors is non-conservation of bin counts, from particles that either overlap or go out of focus in one snapshot and into focus in the next (see Section 3.4.4). Thus, we see that the medium

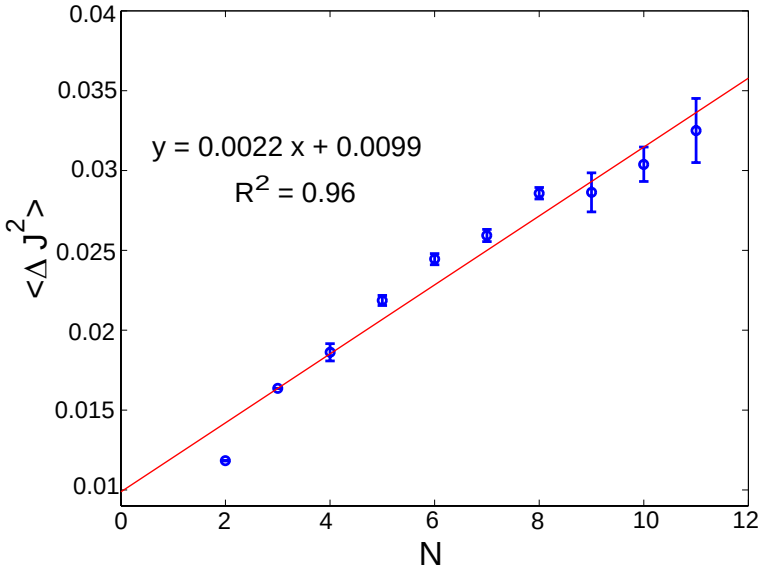


Figure 3.41: The second cumulant, $\langle \Delta J^2 \rangle = \langle J^2 \rangle - \langle J \rangle^2$, vs. the total number of particles, N, for the optimized particle concentration data sets. The second moment of particle flux is proportional to the sum of particle numbers in the two bins. Experimental data is shown in circles, while the solid line represents the fit to the data. The coefficient of determination (R^2) for the fit is also reported. The error bars show the variances due to the different combinations of N_1 and N_2 that result in the same N. The slope and intercept are well predicted by Eq. 3.4.

concentration data provide results that agree extremely well with the theoretical predictions while, at the same time, the particle concentration is such that the range of values for ΔN and N is quite big.

3.5 Conclusions

We have set up a microfluidics experiment in order to study the diffusion of colloidal particles. We have verified that in our set-up the particles are diffusing freely, as there are no hydrodynamic or electrostatic interactions, and the particle concentration evolves in agreement with the diffusion equation. In addition, by running experiments at three different particle concentrations (high, low, and optimized/medium), as well as via simulations, we have established what the experimental error is and how it can be corrected both experimentally and analytically. As a result, we have determined the distribution of particle fluxes in few-particle diffusion. Random-flight and maximum-caliber models predict a gaussian distribution of fluxes. Moreover, with a single parameter p, which is essentially the diffusion constant, we find agreement of the theory with several experimental properties which are usually not examined in diffusion, and are in direct analogy with quantities of recent interest in other nonequilibrium experiments [17, 116, 117, 118]. First, we find that Fick's law—the proportionality of average flux to the gradient of average concentration—holds even down

to concentration gradients as small as a single particle. Experiments also confirm that the variance in the flux is proportional to the total number of particles, $\langle J^2 \rangle \propto N_1 + N_2$, with correct slopes within experimental errors. In addition, we describe a new "flux fluctuation theorem," that is found to be consistent with the data in predicting an exponentially diminishing number of variant trajectories, as a function of the deviation from equilibrium. The model predicts the backwards flows, the bad actors, which are relatively infrequent situations in which particles flow up—rather than down—their concentration gradients and shows that this subset of the overall repertoire of microscopic trajectories can be characterized quantitatively.

Chapter 4

Assembly of the 30S Ribosomal Subunit Using Microfluidics and a Fluorescence Assay

[*Most of the work in this chapter appears in Ridgeway et al. [81]. A big thank you goes to W. Ridgeway.*]

In Chapter 1, we presented the importance of the ribosome, with a historical overview, as well as the importance of studying its self-assembly. We saw that the detailed determination of the ribosomal structure has been a process that has taken up more than 40 years. However, now that the structure has been established, there are still questions about the ribosome that have not yet been answered. Understanding how the ribosome assembles is one of them. In particular, as will be explained later in this chapter, it has been found that the assembly process occurs via several parallel pathways [125] and we aim to attribute probability distributions, so that the probability of each assembly pathway (i.e., a statistically weighted map of assembly) is obtained. In essence, we are again interested in the dynamic process of RNA folding, RNA-protein, and protein-protein interactions.

We start this chapter by reviewing the ribosome, the importance of self-assembly as well as the challenges associated with studying ribosomal self-assembly. We then move on to explore the literature of a well-studied system, that of the binding of the primary protein S15 to the central domain of the ribosomal RNA. We then present an experimental setup that addresses the challenges of studying self-assembly and we test the setup by using the well-studied system of the S15 binding.

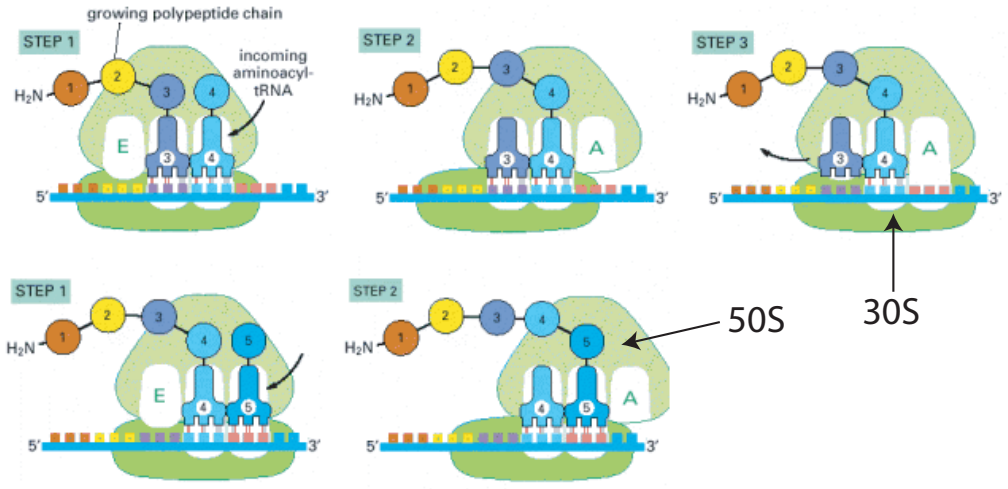


Figure 4.1: The functions of the 30S subunit. During protein synthesis each of the 3 sites, the acceptor (A), the peptidyl (middle), and the exit (E), is occupied in succession by a tRNA. The 3-step cycle shown is repeated as the ribosome moves down the mRNA molecule. The 30S subunit is responsible for proofreading the tRNA, making sure that its anticodon matches the mRNA codon, and for keeping the mRNA in frame, by shifting down the mRNA by only 3 nucleotides at a time. Figure taken from Alberts et al. [26].

4.1 Importance of the Ribosome and Macromolecular Assemblies

The ribosome is the macromolecular assembly inside the cell that is responsible for protein synthesis. In a bacterial cell there are approximately $20,000$ ribosomes [28]. An *E.coli* ribosome is about 20 nm in diameter and has a mass of around 2700 kDa [29]. Ribosomes of both prokaryotes and eukaryotes consist of two subunits, the large and the small. For bacterial and archaeal ribosomes the large and small subunits are called 50S and 30S, respectively [30]. Together these subunits are responsible for translating the messenger RNA (mRNA) and synthesizing the proteins. As is shown in Figure 4.1, protein synthesis in *E.coli* cells is done as follows: the 30S subunit contains the binding site for the mRNA and together with the 50S subunit they form 3 sites, the acceptor, the peptidyl and the exit sites. Each of these sites is occupied in succession by a particular transfer RNA (tRNA) during protein synthesis. The tRNA is "charged" with an amino acid, and delivers the appropriate amino acid specified by the triplet of nucleotides on the mRNA, the codon. The tRNAs bridge the large and small subunits, with the anticodon arm of the tRNA pointing towards the 30S subunit for decoding, and the arm with the amino acid pointing into the 50S subunit. Therefore, the function of the 30S subunit is two-fold. First, it is responsible for proofreading by making sure that it accepts only the tRNAs whose anticodons match the mRNA codons. The other function is to keep the mRNA in frame by shifting down the mRNA only by 3 nucleotides [26].

As was explained in Chapter 1, the importance of the study of the ribosomal self-assembly lies on several different facts. Firstly, ribosomes are similar across different organisms [69]. Thus, understanding their structure, function, and assembly process will provide insights into the similarities between the various organisms and, consequently, the evolution process. In addition, exactly because this structure is responsible for protein synthesis, knowing the mechanism by which the ribosome assembles can be useful for the design of antibiotics: a lot of antibiotics used in the treatment of bacterial infections target protein synthesis via the ribosome [70, 30, 71, 72, 73]. Furthermore, the assembly process is based on RNA folding, protein-protein, and RNA-protein interactions, which, in general, play a central role for the determination of biological function and behavior. As a result, ribosomal self-assembly can serve as a specific case study of these more general phenomena.

Macromolecular assembly is in itself a process of great interest, because it is one of the simplest possible mechanisms for morphogenesis [126]. There are fundamental processes, such as transcription, DNA replication, cellular transport, viral capsid formation, etc., that are based on the self-assembly of macromolecular machines. This process is defined by the fact that the individual components contain all the information for determining the resulting structure in their shapes and binding characteristics [126]. What is even more interesting is the fact that self-assembling structures are able to perform error correction and quality control, resulting in the reliable assembly of the macromolecular assemblies [126]. In particular, the structures manage to perform this error correction via several mechanisms. First, if during the self-assembly process there is a misfit subunit that has bound to the wrong location, then the binding energy for that will be low (i.e., non-specific binding). As a result, it is very likely that this subunit will dissociate again and, thus, the error will be corrected [127]. In addition, in the case when the self-assembly is an ordered process (i.e., certain components have to bind first before new components are allowed to bind), reliable self-assembly is ensured by having the new binding sites exposed only when the first components are already bound correctly. This exposure can happen via, for example, a conformational change induced in the structure when the first components are binding [128, 129]. This method has two very important consequences. The first is that it forces a strict order in the binding of the components and the second is that it prevents the completion of the self-assembly process when a wrong or misaligned component is bound [126], since the new binding sites will not be available. As will be described in detail in Section 4.2, ribosomal self-assembly is also a hierarchical process. In this case it is known that error correction is done via protein chaperones and helicase activities that unfold the bound proteins or the RNA, allowing in this way for the misfit component to dissociate, and then help it refold again in order for the assembly process to continue [66]. Despite all this information, the main question of how the ribosome assembles still remains. Therefore, due to the facts that ribosomal self-assembly is (a) an example of reliable, hierarchical self-assembly, (b) based on interactions that are found to be dictating various functions inside the cell, (c) a potential target for antibiotics,

and (d) a process that presents similarities across different organisms, ribosomal self-assembly is a process of current and intrinsic interest.

4.1.1 The 30S and 50S Subunits

It has already been mentioned that the ribosome consists of two subunits, the 30S and 50S. Each of these subunits assembles separately, before they both come together to form the ribosome. Both subunits consist of ribosomal RNAs (rRNAs) and proteins. The proteins bind to the rRNA as it folds, creating in this way the subunit.

In Chapter 1 we gave a detailed historical overview for the structural determination of the ribosome. In this section we summarize the structures of the individual subunits. The 30S subunit consists of a 16S ribosomal RNA (about 1500 nucleotides long and 500 kDa) and 20 proteins, referred to as S2 to S21, of sizes ∼8-20 kDa. The 50S subunit consists of two RNA molecules, the 23S and the 5S, and 34 different proteins, L1 through L34 [29]. The fact that there are 21 components in the 30S subunit means that in order to study its assembly, we need to study a 21-dimensional reaction. Experimentally, this reaction amounts to many separate, smaller reactions where different combinations of proteins interact with each other and with the 16S rRNA (for an estimate of the number of reactions/interactions, see Section 4.3). Similarly, we would have to study a 34-dimensional reaction for the case of the 50S subunit. Due to this difference in size and in the number of proteins, as will be described in Section 4.2, a lot of studies on the ribosome have concentrated on the 30S subunit, part of which will also be the object of our experiments.

When 30S is completely assembled, it is divided into 3 parts, the head or 3' domain, the central domain and the body or 5' domain. Figure 4.2 shows the structure of the subunit and all its domains. For our experiments, we will be focusing on the central domain of the 30S subunit.

4.2 Background on Ribosomal Assembly

The $30S$ subunit of the bacterial ribosome is a self-assembling macromolecular machine, but our understanding of its assembly is incomplete, owing to the fact that it requires monitoring the association of the 20 proteins it consists of [125].

As was described in Chapter 1, in 1974 the 30S subunit was first reconstituted *in vitro* from its individual components in the lab of M. Nomura via partial reconstitution experiments [48]. These experiments were based on the principle that, if one protein is left out from the mixture of proteins and rRNA then, as a result, certain proteins will not bind to the partially assembled subunit because of their interaction with the missing protein. Nomura's work was the first to provide some insight into the ribosomal self-assembly process and gave rise to what is known as the Nomura map of assembly [48]. What this map depicts is the ribosomal RNA and all 20 proteins are divided into 3

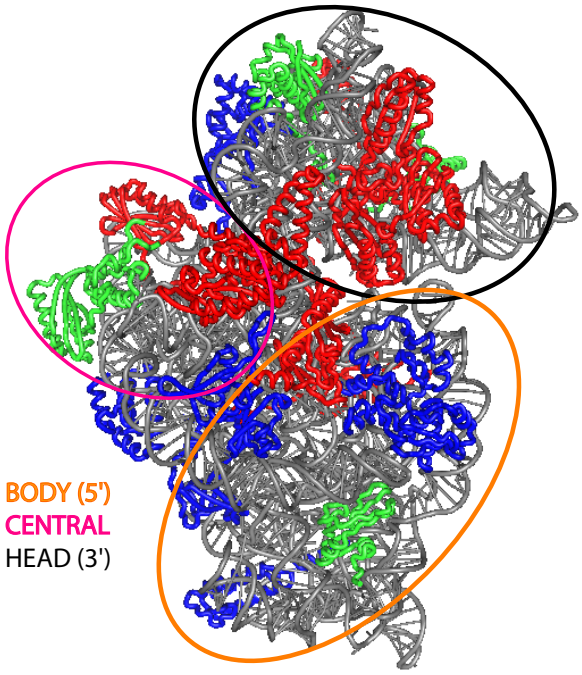


Figure 4.2: The structure of the 30S subunit. When assembled, the subunit is divided into 3 parts, the head domain (circled in black), the central domain (circled in pink), and the body domain (circled in orange). Figure taken from [63]

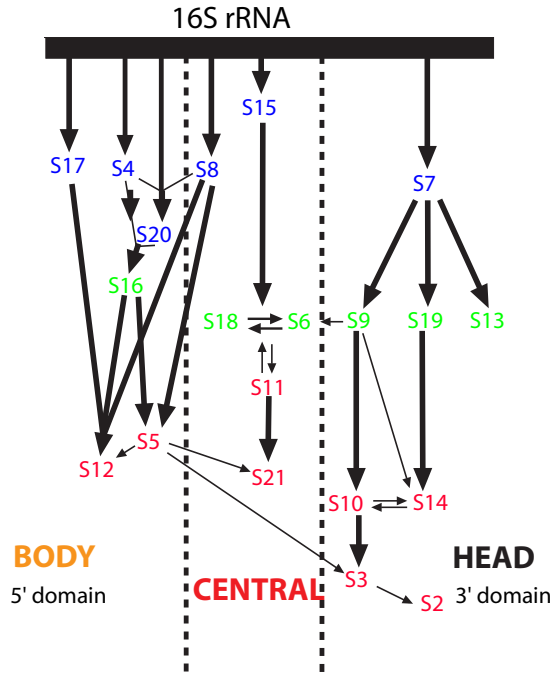


Figure 4.3: The 30S subunit *in vitro* experimental assembly map for *E.coli*. The primary, secondary, and tertiary binding proteins are shown in blue, green, and red, respectively. The domain that each of these proteins corresponds to is also shown. Arrows indicate facilitating effect of binding of one protein on another. From the work of the Nomura Laboratory [48].

categories: the primary, secondary, and tertiary binding proteins. A representation of this map that was based on Nomura's original work is given in [74] and is shown in Figure 4.3. In this figure, the primary, secondary, and tertiary proteins are depicted in blue, green, and red, respectively. In order for the secondary proteins to bind, the primary proteins have to already be present, and in order for the tertiary proteins to bind, the secondary proteins have to be already bound. This is because, as the different proteins bind, the rRNA undergoes conformational changes that reveal new binding sites for the remaining proteins [130]. The domain that each of these proteins corresponds to is also shown in the figure. However, what this map does not tell us is what assembly pathways are more likely to be traversed.

Up until recently it was believed that the assembly process went through an intermediate step where certain proteins would bind to the rRNA at low temperatures (0-15°C) but the assembly would not be completed. These incomplete particles, called the reconstitution intermediate (RI) would have to be heat shocked at 42°C, overcoming in this way a high activation energy barrier, that would allow the assembly to complete at low temperatures [131]. However, using the method of Pulse-Chase monitored by Quantitative Mass Spectroscopy (PC/QMS) Talkington et al. [125] found the protein activation energies. If the RI theory were valid, then the proteins that bound after the heat shock would be expected to have higher activation energies than the proteins that were already bound before the heat shock. However, when plotted, these activation energies did not form any clusters and there were no outliers. In fact, all proteins presented similar activation energies that were within the narrow range of ∼24-44 kcal/mol. This experiment gave rise to the landscape representation of the assembly process as a free-energy surface which is flat when there are no proteins bound to the rRNA. Local foldings are indicated as small minima in the surface, since the folding exposes protein binding sites. When a protein binds, an even greater minimum is created. Thus, the surface is altered every time a protein binds to the rRNA and, since there are several foldings and bindings that can happen at the same time, there are several parallel pathways of assembly, as opposed to the global intermediate (RI) state [125].

Even though we know that there is hierarchical binding, as suggested by the Nomura map of assembly, and that several pathways can be traversed in parallel, as suggested by the landscape representation of assembly, we still do not know which of these pathways are more likely to be traversed. In other words, we are not sure what the statistically weighted map of assembly looks like. A pictorial representation of this question is given in Figure 4.4. We assume that this 21-dimensional process is described by a cube where each vertex represents an intermediate state in the assembly process after each protein has bound and each edge represents a pathway of assembly. Then, the black vertex is the state where none of the proteins are bound to the rRNA and the blue vertex is the completely assembled 30S subunit. As was described, there are several pathways (i.e, edge combinations) that can lead to the completely assembled state from the unbound state.

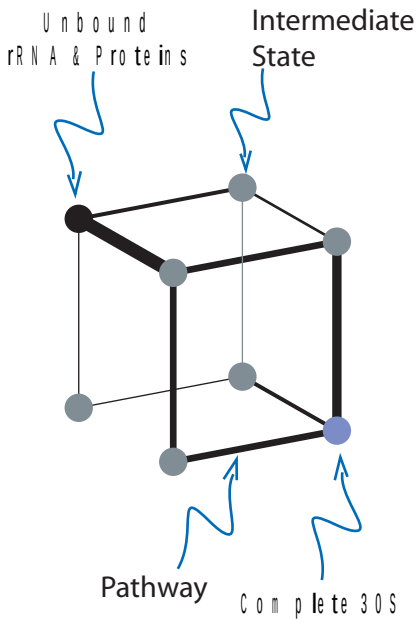


Figure 4.4: Pictorial representation of the statistically weighted assembly map. Each vertex of the cube represents an intermediate state in the assembly process while each edge represents a pathway. Thicker edges are pathways that are traversed often and can be described by the reaction rate k of the binding of each protein to the intermediate state.

However, some pathways will be followed more frequently than others, as is represented by the thickness of the edge in the cube. In order to be able to assign a probability to each edge, we need to determine what the reaction rate k is of the binding of each protein to the intermediate state (i.e, vertex) associated with that pathway.

Since there are 20 proteins that need to bind for the 30S subunit to be formed, it is clear that the determination of each k needs to be an automated process. For this reason, as will be described in the following sections, we are combining microfluidics with fluorescence resonance energy transfer (FRET) and two-photon spectroscopy. However, before we set out to determine all these reaction rates, we will test our system by focusing on the binding of the S15 protein to the central domain of the rRNA.

S15, an 11.9 kDa protein, has been of particular interest because it is a primary protein and, therefore, its interaction with the 16S rRNA is independent of the other proteins [132]. The binding site for S15 at the central domain is a three-way junction (3WJ) formed by the intersection of helices 20, 21, and 22 of the rRNA. Figure 4.5 shows the structure diagram of the 16S rRNA and the position of the 3WJ. Biochemical and biophysical characterization has shown that when neither the S15 protein nor polyvalent cations like Mg^{2+} are present, the angle between these helices is $\sim 120°$ [133]. However, when either S15 or Mg^{2+} or both are present, then a conformational change occurs (folding) whereby helices 21 and 22 are coaxially stacked and helix 20 forms an acute angle

with helix 22 [130, 134]. In addition, it has been found that the folding of the 3WJ is determined entirely by RNA elements and the roles of either the S15 or the Mg^{2+} is to stabilize the intrinsically unstable structure [134]. This data is summarized in the thermodynamic model for the binding of S15 to the 16S rRNA that was proposed in [134] and is shown here in Figure 4.6. Further experiments using isothermal titration calorimetry and gel mobility shift assays have demonstrated the cooperative and ordered assembly of the whole central domain [135, 136].

In addition to the above methods, fluorescence resonance energy transfer (FRET) has also been used in order to report on the conformational change that S15 or Mg^{2+} induce in the 3WJ [78, 79, 80]. The basic idea behind these experiments, demonstrated in Figure 4.7 (the dyes shown in this figure are not the ones used in the work presented in this paragraph), was that the rotation of helix 22 and its alignment with helix 21 at folding also results in the decrease in the end-to-end distance of helices 20 and 22 from 7 nm to 4 nm [80]. If helices 20 and 22 are labeled such that helix 22 has the donor dye and helix 20 has the acceptor dye (acceptor 1), then the FRET efficiency and, thus, the signal from the acceptor dye, would increase markedly once the 3WJ is in the folded configuration [78, 79]. This is because the FRET efficiency, E, depends strongly on the donor-to-acceptor separation distance, r, via the equation:

$$E = \frac{1}{1 + (\frac{r}{R_0})^6}, \tag{4.1}$$

where R_0 is the Förster distance for the specific pair of donor and acceptor dyes. In this case, the donor and acceptor dyes are about 8.5 nm apart when the junction is unfolded but only 5.5 nm apart, when the junction is folded [79, 80]. Similarly, if a second acceptor dye is used to label S15 (acceptor 2), then the binding of S15 stabilizes the folded junction and an increase in the signal coming from acceptor 2 would be expected [80]. Fluorescent correlation spectroscopy (FCS) has also been used in conjunction with FRET to determine the transition rates between the unfolded and folded conformations as a function of Mg^{2+} and Na^{+} concentration, but it was only done to immobilized molecules in order to look at the fluctuations in the fluorescence by FRET [78].

The experimental findings of the above studies concentrated on the dissociation constants, K_D, for the S15 and Mg^{2+} binding on the RNA that result in the folding of the 3WJ. It was found that in the case of S15, K_D ranges from, approximately, $K_D = 3.5$ nM to $K_D = 50$ nM [134, 78, 80] while for Mg^{2+}, $K_D = 77$ μM [80]. In addition, a model that describes the dependence of the RNA folding on Mg^{2+} was also developed [79].

From the studies presented thus far it is clear that there are many aspects of the ribosomal assembly that are well known while other aspects are still a mystery. Even though there is a detailed knowledge of the structure of the 30S subunit and the order at which all the proteins bind to the 16S rRNA, we are still missing all the details about the dynamic interactions that lead to ribosomal assembly [66]. In addition, it has been established that there are several parallel binding

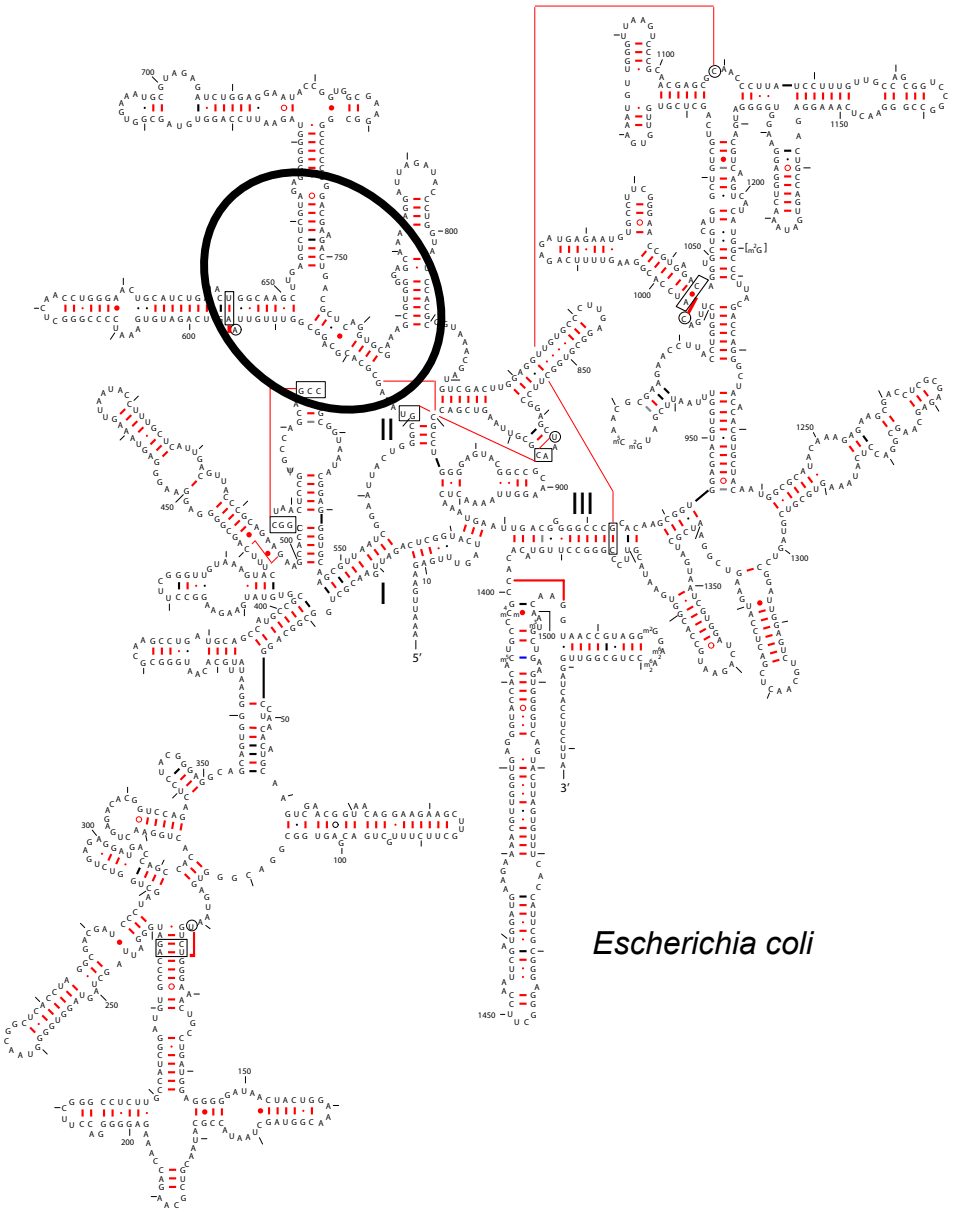


Figure 4.5: Secondary structure diagram of the 16S rRNA of *E.coli*. The part of the central domain where the 3WJ is located is circled. Figure taken from [137].

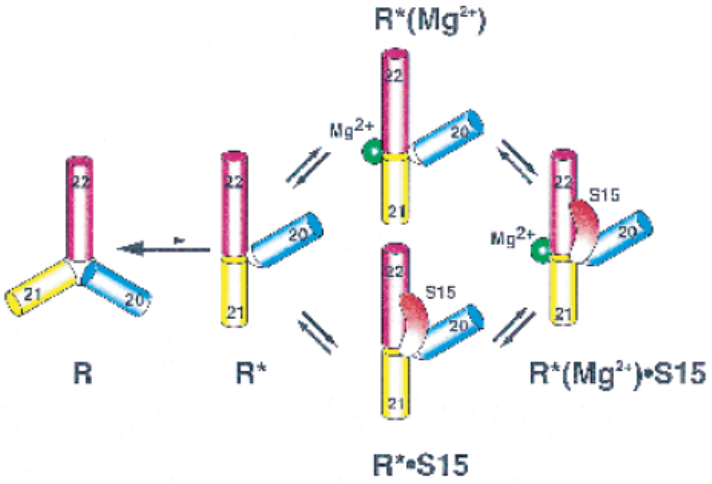


Figure 4.6: Thermodynamic model for S15 binding to 16S rRNA. The RNA is in an equilibrium between unfolded (R) and folded (R*) conformations. The R* conformation is stabilized by either S15 (R* · S15) or Mg^{2+} (R* · Mg^{2+}) binding or both (R* · Mg^{2+} · S15). Figure taken from Batey and Williamson [134].

pathways, but the probability of occurrence for each pathway is still unknown and a systematic way of examining all these different pathways needs to be found. The fact that the binding of S15 to the 3WJ has been so well characterized provides a frame of reference for testing a possible new systematic way. In the sections that follow, we describe such a systematic way that can be used to study the complete assembly of the 30S subunit and we test it by looking at the binding of S15 and Mg^{2+} on the 3WJ of the 16S rRNA.

4.3 Experimental Goal

As a summary to the previous section, here we repeat the question and objectives of our experiments. The assembly of the 30S ribosomal subunit is known to observe a hierarchical binding of the 20 proteins to the 16S rRNA. There are several pathways that can be traversed in parallel during this assembly. However, what has not yet been established is what the statistically weighted assembly map is: Which pathways have a higher probability of being traversed? In order to answer this question, the various reaction rates ks need to be determined.

Since there are 20 proteins that need to be considered, there are many reactions that need to take place in order to obtain all the ks. As an estimate, let's say that we have the 20 proteins and we want to use each one at 3 different concentrations in order to observe the dependence of the protein-protein interactions on protein concentration. Each of the reactions will consist of a specific combination of concentrations of the various proteins. This problem is the same as that of rolling

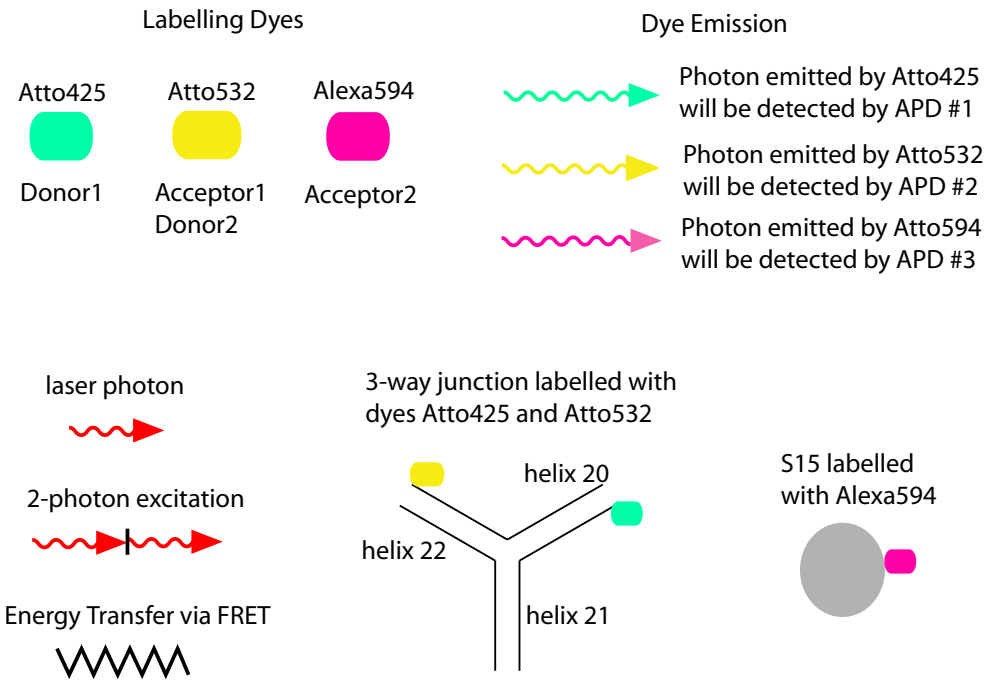


Figure 4.7: Labeling dyes and detection scenarios for the 3-way junction folding and S15 binding. There are two dyes (Atto425, Atto532) that label the 3-way junction (3WJ) of the 16S rRNA. Atto425 is on helix 20 and Atto532 on helix 22. There is also one dye (Alexa594) that labels the S15 protein. The dyes Atto425, Atto532, and Alexa594 are shown in green, yellow, and red, respectively. The dyes have been chosen so that the emission spectrum of Atto425 overlaps with the excitation spectrum of Atto532 and the emission spectrum of Atto532 overlaps with the excitation spectrum of Alexa594. The peaks of the emission spectra of the dyes Atto425, Atto532 and Alexa594 are at 485 nm, 553 nm, and 617 nm, respectively. In this FRET construct, Atto425 is always a donor dye and Alexa594 is always an acceptor dye. Atto532 can be both, depending on whether Alexa594 is present at close proximity. If not, and Atto425 is close enough that FRET can occur, then Atto532 is an acceptor dye. If S15 is bound, then Atto532 is an acceptor dye because it accepts the photon emitted by Atto425 but also a donor dye because its emitted photon will be absorbed by Alexa594. The laser wavelength is 880 nm, approximately twice the excitation wavelength of Atto425 since we are using 2-photon spectroscopy (see Section 4.7). The 2-photon excitation is represented by the two red arrows while the energy transfer that occurs during FRET is shown as a black zig-zag line. These representations are used in Figure 4.8. In the experiments conducted, both the donor and acceptor signals are detected via Avalanche PhotoDiodes (APDs) (see Figure 4.25).

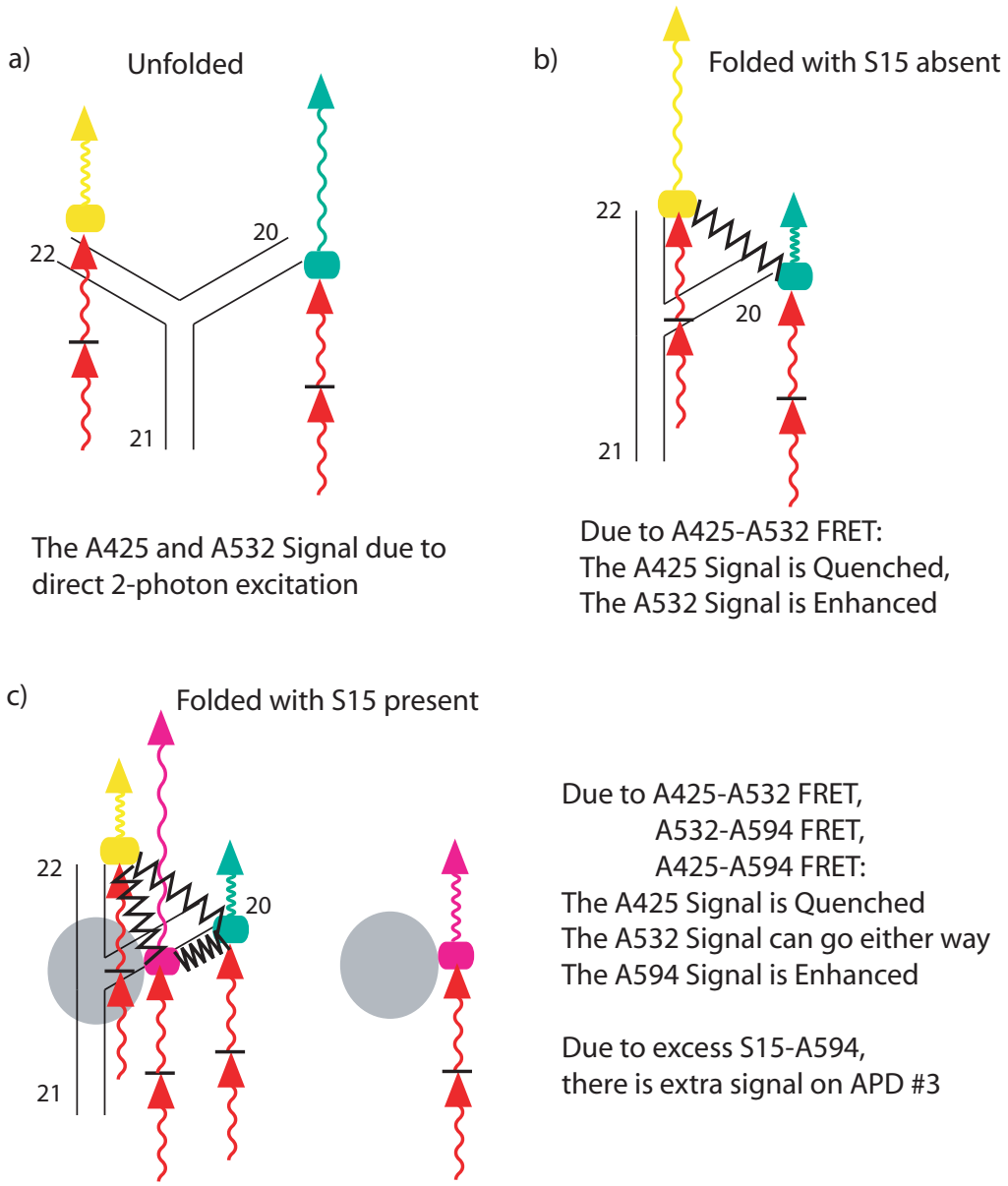


Figure 4.8: Conformations of the 3WJ and dye emissions corresponding to each conformation. (a) When the 3WJ is in the unfolded state, then the 2-photon laser will excite both Atto425 (donor) and Atto532 (acceptor). Since the laser wavelength (880 nm) is set to target Atto425, emission from Atto532 will be low. This is the background emission for Atto532. The 3WJ is in the unfolded state and, thus, the two dyes are not close enough for FRET to occur. Hence, the ratio of acceptor / (donor + acceptor) will be low. (b) The 3WJ is flipping back and forth between the unfolded and folded states, but the presence of Mg^{2+} favors the folded conformation. When the 3WJ is folded (S15 absent), the signals detected are the same as in the unfolded state (a). However, now the Atto532 signal is enhanced while the Atto425 signal is suppressed because of FRET: the photon emitted by Atto425 will be absorbed by Atto532. Thus, in the presence of Mg^{2+}, in addition to its direct excitation from the 2-photon laser, Atto532 will emit light because of FRET. Consequently, when compared with the unfolded state, the ratio acceptor / (donor + acceptor) will be high. (c) The binding of S15 to the 3WJ brings Atto532 and Alexa594 in close enough proximity that FRET can occur between them. Therefore, in addition to the signals observed in the folded state (b), there is emitted signal from Alexa594. The Atto425 (donor) signal will be highly suppressed because the emitted photons can be absorbed by both Atto532 and Alexa594 (FRET). The emitted Atto532 photons can be absorbed, via FRET, by Alexa594 (acceptor). Here the ratio acceptor / (donor + acceptor) will be very high. Just like with Atto532, some emission from Alexa594 due to direct 2-photon excitation is expected. This background emission can be quantified with measurements of the Alexa594 emission signal when 3WJ is absent.

n 6-sided dice, where the total number of outcomes is 6^n. Then, in our case, the total number of reactions, N, that will take place is:

$$N = 3^{20}! \tag{4.2}$$

In the case of the Mg^{2+} and S15 titrations to the 3WJ that we present in Section 4.10, where we have about 14 different concentrations for each, then, if we want to replicate each experiment 5 times, the number of reactions quickly becomes (we assume a constant concentration for the 3WJ and the reactions considered are Mg^{2+} titrations to the 3WJ and S15 titrations to the 3WJ):

$$N = 5 * 14 + 5 * 14 = 140. \tag{4.3}$$

Therefore, it becomes clear that it is important that we automate the process. Consequently, the use of microfluidics is essential. A detailed description of the microfluidics chip designed for this purpose will be given in Section 4.6. Furthermore, as was described in Section 4.2, the distinction between the different state intermediates in the assembly process has been shown to be facilitated with the use of fluorescent dyes, whereby each intermediate is distinguished from the others because it emits light at a different wavelength [78, 79, 80]. Then, our goal is to determine reaction rates based on the emission properties of the intermediates. Therefore, the use of FRET will be necessary, as it has shown to be a powerful technique for such purposes [78, 79, 80]. The observation of reaction rates at various reactant concentrations inside microfluidics, where the volumes are in the order of nanoliters (See Section 4.6) results in a small number of molecules being present in the detection volume. Detection from a small number of molecules implies that the background fluorescence signal needs to be eliminated. In addition, since it is likely that the molecules that are in the vicinity of the detection volume will soon provide the signal when they enter the focal volume, it is important that these molecules are not already photobleached. Both of these conditions are satisfied when two-photon spectroscopy (TPS) is used [138]. These techniques will be described in detail in Section 4.7. In order to test this technique of combining microfluidics with TPS and FRET, we will use a system that has been extensively studied, namely that of S15 and Mg^{2+} binding on the 3WJ of the rRNA. The labeling dyes that we will use for this system and the nature of the donor and acceptor signals that will be observed for each configuration of the 3WJ are shown in detail in Figure 4.7. Recovery of the well-known dissociation rates K_D in this case will establish that our 3WJ labeling and detection systems are functioning properly and will allow us to continue to the more complicated system of the complete 30S assembly.

Therefore, the experiments that are considered are the following:

- Folding of the 3WJ RNA as a function of Mg^{2+} concentration. As is described in Figures 4.7 and 4.8, we will use FRET. From this experiment we will obtain the K_D for the Mg^{2+}-induced

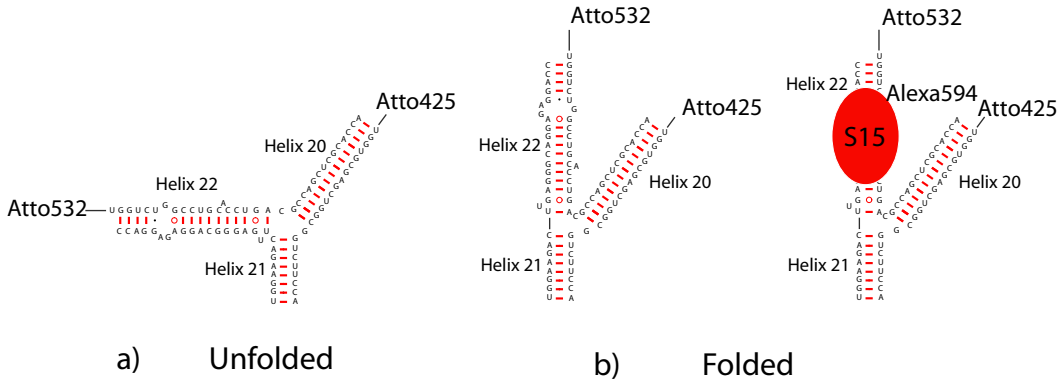


Figure 4.9: The 3WJ construct used for the FRET experiments of Mg^{2+}-induced folding and S15 binding. In this figure we present the exact 3WJ construct used, the base-pairs of each helix, and the positions where the labeling has occurred.

3WJ folding. The results are shown in Figure 4.29.

- Binding of $S15$ on the 3WJ RNA as a function of the $S15$ protein concentration when Mg^{2+} is absent. Again, we will use FRET and, from this experiment, the goal is to obtain the k_{on}, k_{off} and K_D. The results are shown in Figures 4.30 and 4.31. This process is expected to give us a bi-exponential behavior: since S15 cannot bind on the unfolded 3WJ but has to wait until the 3WJ folds, the whole binding process consists of the slow process of the folding of the 3WJ (the 3WJ can still fold in the absence of Mg^{2+} but it is energetically less favorable) and the fast process of the binding of the S15 on the folded 3WJ.

4.4 3WJ Constructs

The three-way junction (3WJ) system is an artificial fragment of the 30S which has been extensively studied [132, 139, 134, 80, 133, 78]. The triple-stranded structure is known to perform two tasks central to ribosome assembly (i) undergo tertiary structure folding and (ii) bind a ribosomal protein, protein S15. FRET constructs have been developed to monitor both the RNA folding and RNA protein binding [80, 133]. We are using this 1 RNA, 1 Protein construct to help us design and test our microfluidic reactor simply because it is much better characterized and behaved than our ultimate target, the 1 RNA, 20 protein 30S.

We have adapted the 3-fluorophore system of Klostermeir et al. [80] to use dyes compatible with the relatively harsh optical conditions associated with 2-photon excitation. We have sought to create a construct which could tell us both the folding state of the RNA and whether or not S15 had bound the RNA (see Figure 4.9). The final system employs one general FRET donor (Atto425),

and two FRET Acceptors, Atto532 (on the RNA) and Alexa594 (on the protein). In the absence of S15-Alexa594, FRET between 425/532 is readily detectable. In the presence of S15-Alexa594, quenching of both Atto425 and Atto532 occurs, and energy is transferred ultimately to Alexa594.

3WJs were assembled from three synthetic strands of RNA [80]:

N6.20-21 N6. ug gug cga gcu ggc ggu cuu cca

N5.21-22 N5. ug gaa gac uug agg gca gga gag gac ca

N5.22-20 N5. ug guc ugg ccu gca ccu gac gcc agc ucg cac ca

where N5 and N6 are amino groups attached to the RNA by 5 and 6 carbon linkers, respectively (Dharmacon Research, La Fayette, CO).

Two strands, 22-20 and 20-21, were labeled with NHS Ester functionalized fluorophores Atto 425, Atto 532, and Atto 633 (Atto-Tec GmbH), using the protocol outlined by Atto-Tec as follows:

RNA was first extracted three times with phenol/chloroform/isoamyl alcohol to remove any traces of primary amines which might be present (e.g., Tris). It was then ethanol precipitated (1/20th volume 5M NaCl, 4 Volumes EtOH, chill at $-20°$C overnight), spun down, and the pellet was washed twice with 20 μl cold 70% EtOH. After partial drying, the RNA was resuspended in ddH_2O.

The coupling reactions were performed on the 40–50 nMole (RNA) scale, with 20-fold excess dye-NHS Ester. RNA was first diluted in 75 mM Sodium Tetraborate, pH = 8.5, while the dye-NHS Ester was resuspended in anhydrous DMSO. The reaction was initiated by mixing the two solutions in a 5 ml glass round-bottomed flask with a Teflon stirbar. The reaction was performed in the dark, at room temperature, for 4 hours. The final RNA concentration was 125 μM, and the DMSO concentration was 15%. While the dye appeared largely insoluble at this low DMSO concentration, the reaction nevertheless progressed to greater than 60% efficiency, as judged by HPLC. The reaction was stopped by ethanol precipitation as before, and 3 phenol/chloroform/isoamy alcohol extractions removed the majority of the unreacted dye.

Labeled RNA was purified from the unlabeled RNA and the unreacted dye by analytical-scale HPLC (Waters Delta-Pak C18, 300 mm x 3.9 mm, H_2O/ACN/0.1 M Tri-Ethyl Amine Acetate pH = 7.0). In all cases, free dye was purified away from the RNA. Labeled RNA eluted at a higher percentage B than unlabeled, and baseline resolution was achieved for Atto425 and Atto633, while Atto532-labeled RNAs were less clean in general. To assess the degree of labeling for Atto532, we used UV/Visible absorption spectroscopy and found that the amount of dye and the amount of RNA agreed within 20%. However, since extinction coefficients are known to about the same precision, it is impossible in principle to get a good quantification of the ratio of unlabeled to labeled RNA, and we will only claim that our labeling yields of the Atto532 are better than 60%. For our purposes, the worst-case scenario of 60% labeling will not adversely affect our experiments, as while low acceptor (Atto532) labeling will decrease the apparent FRET Efficiency (E) in the 3WJ 425/532

FRET construct it will not change the way in which FRET reports on the Mg^{2+}-dependency of 3WJ folding and the S15 binding.

3WJs were created by mixing the three strands at 5 μM : 5 μM : 10 μM in 100 mM $NaPO_4$ pH = 7.0, 100 mM KCl (the strand in excess is unlabeled) and annealed by heating to 95°C for 2 minutes, followed by a slow cool to 35°C at 2°C/min. The annealing reactions were conducted in a PCR machine for consistency. To check for junction formation, all 6 possible junctions were run on a native 6% PolyAcrylamide gel at 4°C. This gel is shown in Figure 4.10. Not only are all junctions formed, but there is no obvious excess of one strand over another, nor do we find any un-conjugated dye in our reactions.

To verify that we have functional 3WJs we tested the ability of the junction to fold in the presence of Mg^{2+} ions (see Figure 4.11). We see that FRET is occurring as anticipated, and the junction binds Mg^{2+} with the same affinity as previously reported [80]. Further testing with labeled S15 (see below, Section 4.5) also confirms that we have functional junctions.

The final construct that we picked is the 22-20-Atto532/20-21-Atto425/21-22, because it allows us to do both the 3WJ folding experiment and the S15 binding experiment by labeling S15 with a redder dye (like Alexa594). In principle, and for the same reason, the construct 22-20-Atto425/20-21-Atto532/21-22 could have also worked, but we had a very limited supply of that. The final construct is shown in Figure 4.9.

4.5 Preparation and Fluorescent Labeling of Protein S15-R78C

S15 was originally cloned from *Bacillus stearothermophilus* [134]. Later an N-terminal His- tag was added for ease of purification, and for purposes of site-specific labeling, Arginine 78 was site-specifically mutated to a cysteine which reacts with maleimide-functionallized dyes [80]. In the end, the protein has the following form:

(HisTag) (Factor Xa Cleavage Site)
ALA LEU THR GLN GLU ARG LYS ARG GLU ILE
ILE GLU GLN PHE LYS VAL HIS GLU ASN ASP
THR GLY SER PRO GLU VAL GLN ILE ALA ILE
LEU THR GLU GLN ILE ASN ASN LEU ASN GLU
HIS LEU ARG VAL HIS LYS LYS ASP HIS HIS
SER ARG ARG GLY LEU LEU LYS MET VAL GLY
LYS ARG ARG ARG LEU LEU ALA TYR LEU ARG
ASN LYS ASP VAL ALA ARG TYR*CYS*GLU ILE
VAL GLU LYS LEU GLY LEU ARG ARG

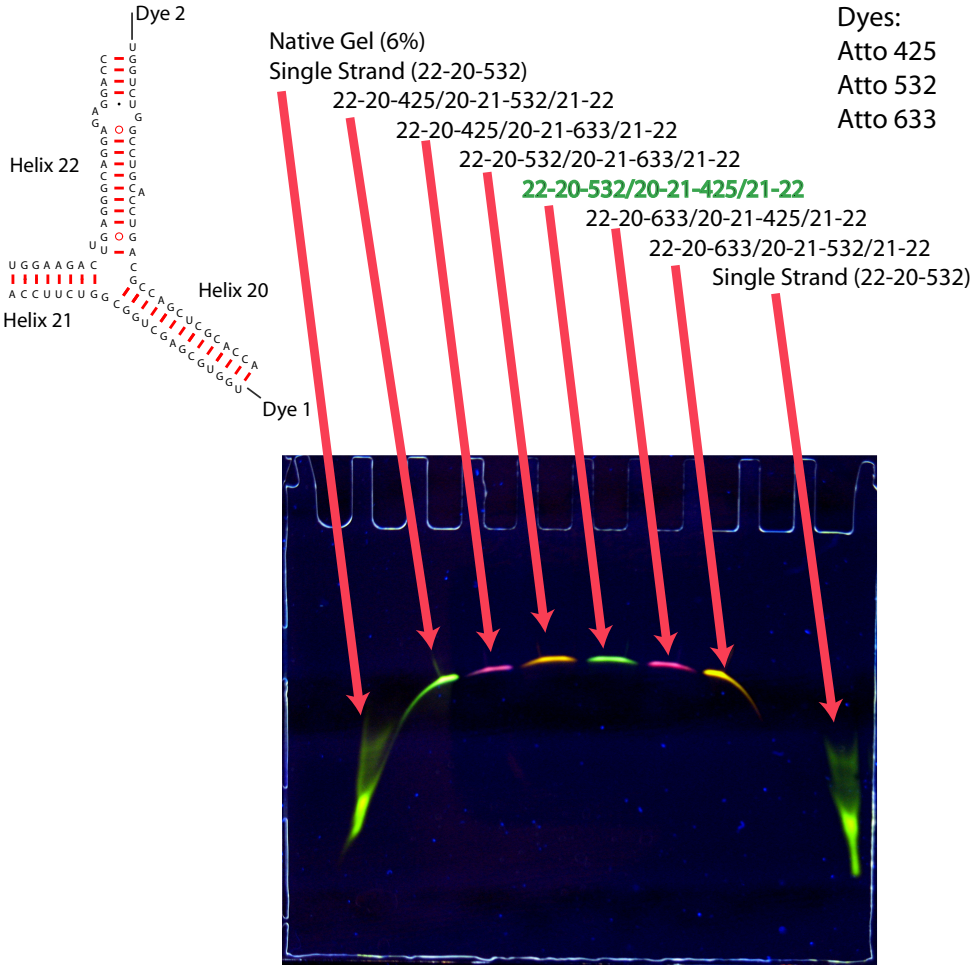


Figure 4.10: Native gel showing all 6 3WJ constructs after the annealing process and the 22-20-Atto532 single strand. The wells of the gel are labeled to show which 3WJ (RNA strand and labeling dye) each well corresponds to. The single strand 22-20-Atto532 is also shown for reference. There are two main observations: i) all 6 3WJ constructs have annealed. That is, the dyes labeling the strands had no effect on the annealing of the 3WJ and ii) the annealing is complete. In the wells containing the 3WJ constructs there is no component corresponding to the single strand, whose position in the gel is shown at the outer wells. Also, the absence of any bands at the bottom of the gel corresponding to free dye is noticeable. The final construct that we picked, shown in green font, is the 22-20-Atto532/20-21-Atto425/21-22. We picked this construct because it allows us to do both the 3WJ folding experiment as well as the S15 binding experiment by labeling S15 with a redder dye (like Alexa594). In principle, and for the same reason, the construct 22-20-Atto425/20-21-Atto532/21-22 could have also worked, but we had a very limited supply of that.

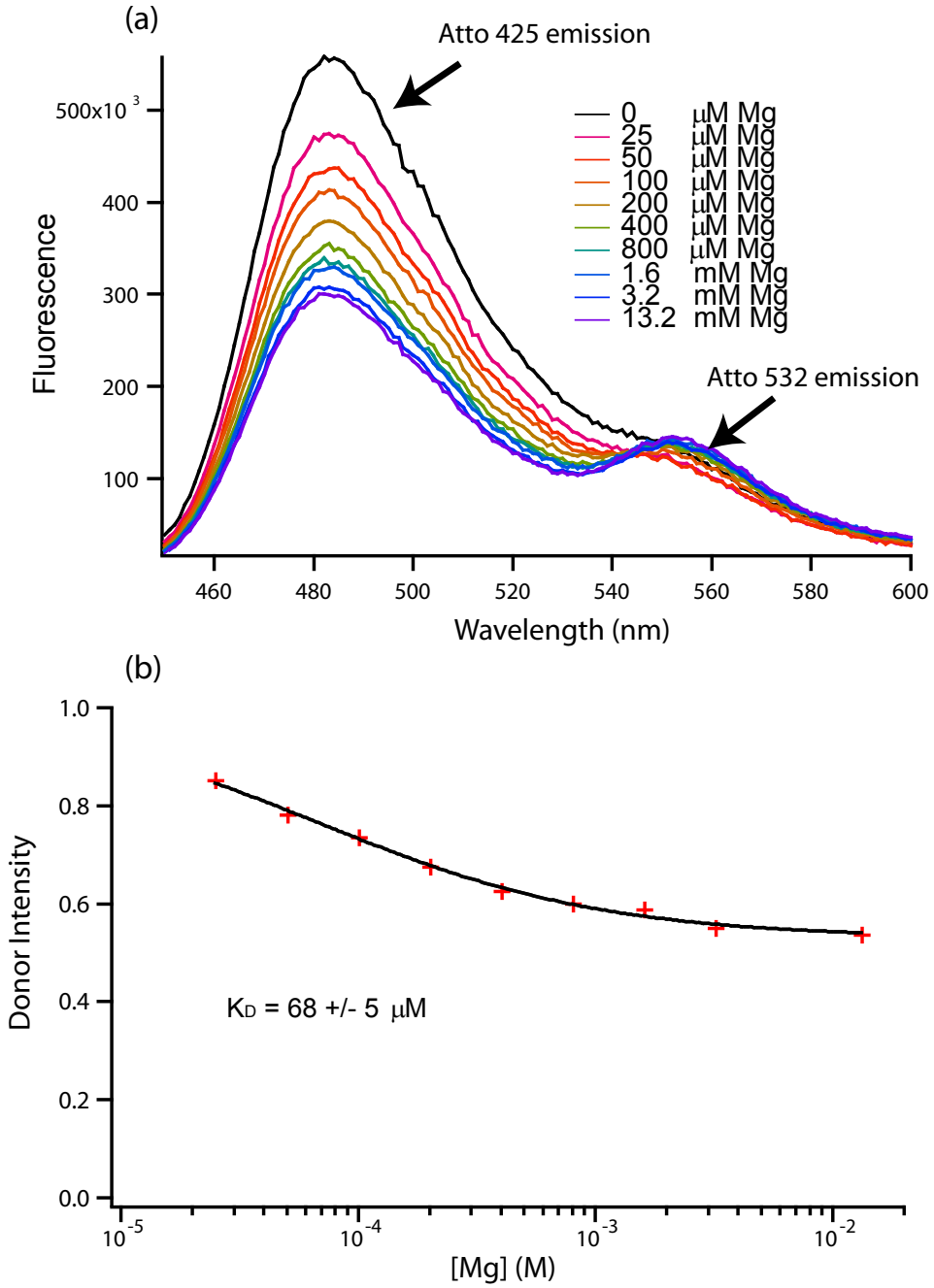


Figure 4.11: Emission spectra of the 3WJ 22-20-atto532/20-21-atto425/21-22 construct as a function of Mg^{2+} concentration. We used 50 nM 3WJ in 50 mM Tris pH = 7.6, 0.5% Tween-20 [80] and the data was obtained using a fluorimeter. (a) Here we report the fluorescence of the sample as a function of wavelength for different Mg^{2+} concentrations. We can see that as the concentration of Mg^{2+} increases, the donor signal (Atto425) from the 3WJ decreases while the acceptor signal (Atto532) increases due to FRET coming from the folding of the 3WJ. (b) This graph represents the folding of the 3WJ (as indicated by the donor intensity and normalized for $[Mg^{2+}] = 0$, which is the black curve in figure (a)) vs. $[Mg^{2+}]$. Each of the data points of this graph is the integral of the donor fluorescence in (a). Fitting the curve to the Hill equation 4.13 gives us the affinity K_D. From the fitting, we have also found that $n = 1$ (see Section 4.9.1).

Our preparation and purification protocols follow both references [134, 80], and proceed as follows:

1. Transform S15 overexpression vector with Amp resistance gene into JM109 cells.
2. Pick single colonies, grow up in LB / AMP (100 μg/ml) overnight at 37°C and 280 rpm.
3. Innoculate 2x1 L cultures, grow at 37°C and 280 rpm. Induce overexpression at O.D. $\sim$ 0.6 with 1 mM IPTG, and grow for 2.5 hours further.
4. Harvest.

Cell pellets were resuspended in 200 ml Lysis buffer (100 mM NaH_2PO_4, pH = 8.0, 10 mM Tris-HCL, 8 M Urea) and lysed by sonication. Supernatant was cleared by centrifugation (15,000 x g, 30 min) before incubation with 50 ml Nickel-NTA slurry (Qiagen) at room temperature for approximately 1 hr. The resin was poured into a column and rinsed with 100 ml wash buffer (same as lysis, but pH = 6.3), before eluting with 50 ml low pH (same as lysis buffer, but pH = 4.5). The eluant was dialyzed 3 times against 4 L 2% acetic acid over 48 hours in a 3.5 molecular weight cutoff (MWCO) membrane (Spectrapore) to remove the lysis buffer, then lyophilized to a powder. Samples from all stages were checked for purity on a 4–20% tris-glycine SDS-PAGE gel run in MES buffer.

To prepare the cysteine for reaction with maleimide dye, the protein was resuspended to $\sim$ 1 mM in 6 M Urea, 50 mM Tris-HCl pH = 7.6 and was reduced with 50 mM DTT for 2 hours at 37°C. The reduction was stopped by the addition of TFA (trifluoroacetic acid) to 0.5% final concentration, and the protein was purified away from DTT by C18 Reverse-Phase chromatography in a H_2O/Acetonitrile/0.1% TFA system (Vydac 218TP column, 5–40% B/25 min to elute DTT, 40–60% B/60 min to elute S15). DTT was separated from S15 with baseline resolution. The S15-containing fractions were pooled and lyophilized. The purity of the final S15 product after C18 is shown in Figure 4.12a.

4.5.1 Protein Labeling with Alexa Fluor 594 C5 Maleimide (Invitrogen)

100 nMoles lyophilize protein was resuspended in 0.94 ml degassed 7 M Guanidinium HCl, 100 mM $NaPO_4$ pH = 7.0 and immediately mixed with 300 nMoles Alexa Fluor 594 C5 maleimide (Alexa594) suspended in 60 μl anhydrous DMSO (Sigma). To avoid photochemical byproducts or photobleaching during the reaction, it was started in very low light and run in the dark (16 hrs, room temperature).

Unreacted dye was separated from protein in two stages. First, it was dialyzed twice against 1 L 6 M Urea, 50 mM Tris pH = 7.6 for 24 hours total. It was then purified using C18 high-performance liquid chromatography (HPLC) with the aim of resolving dye-modified protein from unmodified protein. (Waters, Delta-Pak C18, 300 mm x 3.9 mm 15 micron, H_2O/ACN/0.1% TFA. Inject at 5% and wash for 5 min, then run the gradient of 5–40% for 10 min, then increase to the gradient of

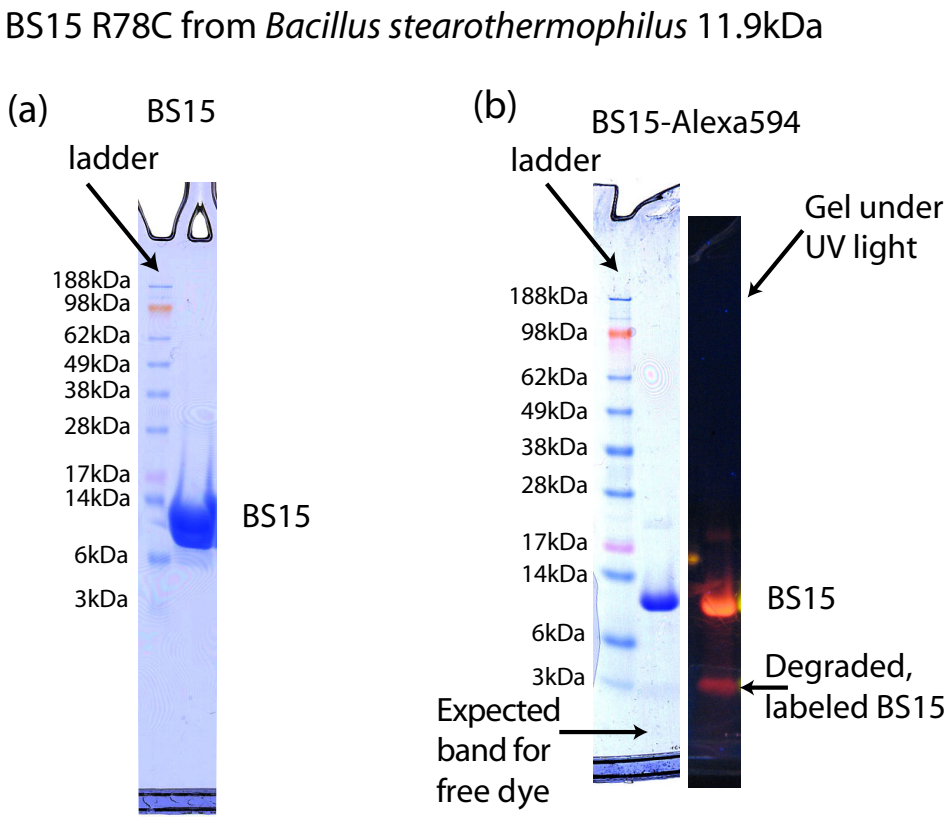


Figure 4.12: S15 purification and labeling gels. Both gels were run in MES buffer. (a) This gel was run by taking the final product after the C18 Reverse-Phase Chromatography. The well on the left is the ladder (Seeblue2), marked to show the size in kDa that each band corresponds to, while on the right we have the S15. We see that all the final product is S15, which explains the smearing of the band. At this point the S15 protein is not labeled. (b) The gel run with S15 after labeling with Alexa594. The Coomassie gel on the left shows two wells, the left being the ladder (Seeblue2), marked to show the size in kDa that each band corresponds to, and the right being where we ran the S15 protein. There is only one band, indicating the presence of labeled S15 and there is no second band, indicating that there is no free dye left. That is, all the Alexa594 dye has been used for labeling. Therefore, the all the signal obtained when the dye is excited comes from the dye on the S15. The fluorescent gel we see on the right shows the labeled S15 band under UV excitation. The lower second band represents degraded but labeled S15, which is only 15% (data not shown) of the labeled (usable) S15 seen on the upper band.

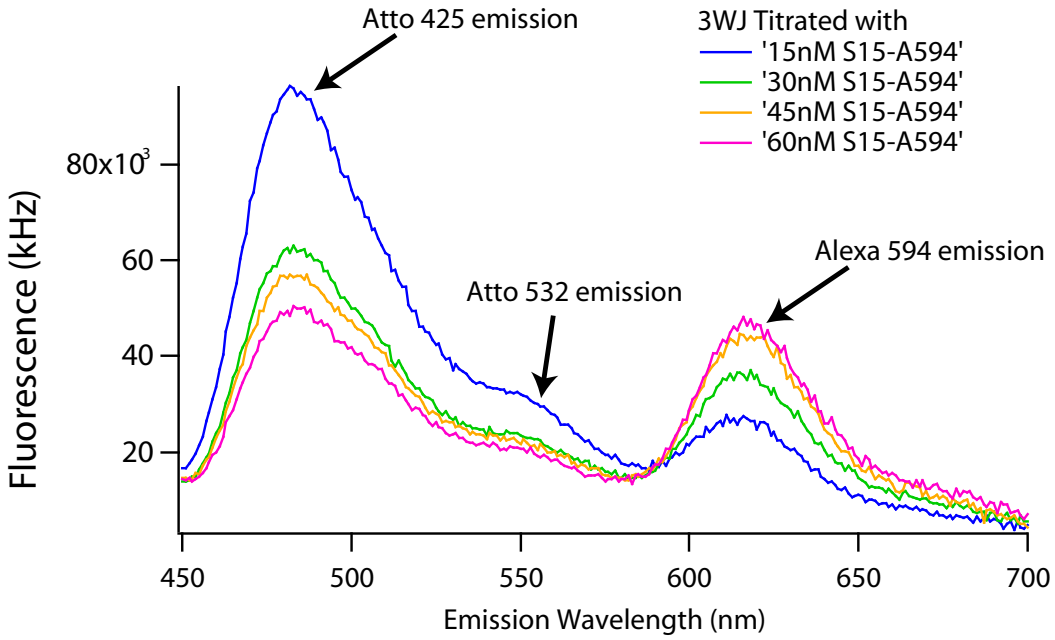


Figure 4.13: Emission spectra of S15-Alexa594 titrated into 10 nM 3WJ construct in 50 mM Tris pH = 7.6, 0.5% Tween-20 [80] and data was obtained using a fluorimeter. The fluorescence of the 3WJ construct 22-20-atto532/20-21-atto425/21-22 with bound S15-Alexa594 is shown as a function of wavelength and for different S15 concentrations. We can see that as the concentration of S15 increases, the donor signal (Atto425) from the 3WJ decreases while the acceptor signal (Alexa594) of S15 increases. From the spectra it is clear that both Atto425 and Atto532 are quenched by the addition of S15-Alexa594, indicating FRET is occurring. From this figure we can also see that the best way to represent FRET from the 3WJ to the S15-Alexa594 is by looking at the donor Atto425 and the acceptor Alexa594. In principle we could have chosen to look at the donor-acceptor pair of Atto532-Alexa594. As can be seen from the figure, however, the pair of Atto425-Alexa594 exhibits better spectral separation between the two dyes. This is why, in the figures reported in Section 4.10.2 the donor is Atto425 and the acceptor, as expected, is Alexa594, the dye labeling S15.

40–55% within 50 min.) The labeled protein eluted at a higher percentage of B than the unlabeled, and although baseline-resolution was not achieved, S15-A594 collected from later fractions showed very little un-labeled contaminating protein (see Figure 4.12b). This was later confirmed by the ability of S15-A594 to quench the 3WJ by FRET to greater than 95% efficiency (see Figures 4.13 and 4.14), placing an upper bound on the amount of unlabeled S15 at 5%.

S15 was lyophilized and stored at −20°C. Stock solutions were created by resuspending to 0.5 mM in 6.8 M Guanidinium HCl, 10 mM Hepes pH = 7.0. S15 was folded immediately before use by rapid dilution into 50 mM Tris pH = 7.6, 0.5% Tween-20. The remaining Guanidinium was in the range of 1–40 mM. Mass spectra confirmed the identity of the protein.

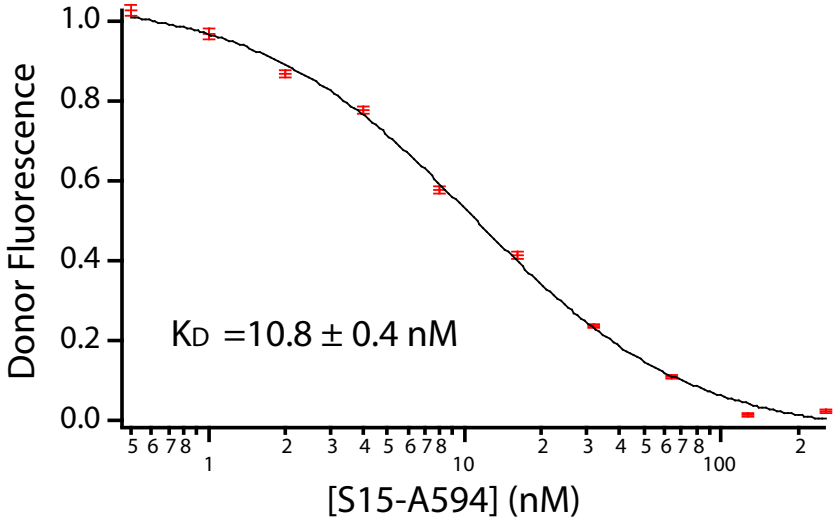


Figure 4.14: S15 titration into 3 nM 3WJ in 50 mM Tris pH = 7.6, 0.5% Tween-20 [80]. The donor fluorescence is reported as a function of S15 concentration. The measurements were taken using the plate reader. Each data point is the average of 10 measurements and the error bars correspond to the standard deviation. The value of $K_D \simeq 10.8$ nM was obtained by fitting the data to Eq. 4.44 (see Section 4.9.2) and agrees with the range of values (3.5–50 nM) reported in the literature [134, 78, 80]. Therefore, labeling S15 with Alexa594 does not perturb the binding reaction to the 3WJ.

4.6 The Microfluidic Chip

As has been explained, the study of the reaction kinetics for the assembly of the $30S$ subunit is a multidimensional process and requires automation. For this reason, we are going to use microfluidics, which allows not only for automation but also for the use of a small amount of reagents (of the order of a 1 nL, as will be shown below). In particular, for our purposes, we need a design that allows for many inputs/reagents to be used at different combinations. From an array of inputs we can select which reagents we want to have as participants in each reaction. The reagents are then loaded one by one into an area where they can be diluted, to obtain them at different concentrations, and then mixed in order to start the reaction. The same area is also where observation can take place. As soon as the reaction/observation is over, the area can be washed clean so that a new reaction can take place.

4.6.1 Chip Design

The microfluidics chip we are using is shown in Figure 4.15 and was designed in Adobe Illustrator. It consists of two layers, the flow layer, shown in blue, and the control layer, shown in red. The flow layer is where all the reagents are loaded at 6 psi, and where reactions are observed. The control

layer controls the flow of the samples that fill the flow layer. When pressure is applied at any given line of the control layer, then that line expands and collapses on top of the flow line it intersects. As a result, the fluid filling the given flow line is unable to flow past that point of intersection. In effect, the control line has closed the flow line underneath it. The control lines are operated at a pressure of 18–21 psi. The chip contains many elements of the microfluidics formulator, first developed and used for the systematic investigation of protein phase behavior [140]. In more detail, the chip consists of the following components:

- A multiplexer. This corresponds to the input lines α-θ of the flow layer and the control valves $A0$-$C1$ of the control layer. Each of the reagents (RNA, proteins, buffer, etc.) is loaded into one of the input lines, with the buffer occupying the last line θ so that, when allowed to enter the mixing ring, it will also wash out the common line of the multiplexer that leads to the ring. The control valves form a binary tree: a logic circuit that minimizes the number of controls needed for a given number of inputs (instead of having 2^N controls for 2^N inputs, we only have $2N$ controls). This tree operates so that an input can be selected to be delivered to the mixing ring at specific amounts. The input lines are loaded as follows:

  α Contains the Alexa488 dye (at an initial concentration of 100 nM) used to do serial dilutions in order to calibrate the chip (see Figure 4.17)

  β Contains the 3WJ at an initial concentration of ~ 100 nM

  γ Contains S15 at an initial concentration of ~ 1 μM

  δ Contains Mg^{2+} at an initial concentration of 10 mM

  ϵ–θ At this point all these lines contain dilution buffer. However, it is important that the last line in the multiplexer (θ) has buffer, so that, when selected to enter the ring, it will rinse the main line of the multiplexer (the one leading to the ring) from any residual S15 and 3WJ rRNA.

- A mixing/reaction ring with the peristaltic pump. This corresponds to the blue circle (ring) and the control valves $P1$–$P3$ (pump). Pumping of the control valves at 33 Hz allows for the precise metering of the reagents in time and volume and results in the delivery of the reagents from the multiplexer into the mixing ring. The ring volume is approximately ~ 3 nl (see below for estimate). When the reagents have entered the mixing ring, the same pump operates again in order to push the reagents around the ring and mix them. The mixing time was determined to be approximately 3 s, as monitored by dye/buffer mixing. Below we will show that this mixing time is about 770 times faster than the mixing time we would expect only via diffusion. The pumping sequence that allows for the precise metering and mixing of the reagents is described in [140].

- The waste line and the switch. This corresponds to the blue zig-zag line labeled $O2$ and the control valves $Z0$–$Z2$. When the purpose of the pump is to deliver the reagents into the ring $Z2$ is open to allow for the reagents coming from the multiplexer to flow into the ring. In order to switch from the delivery of the reagents to the mixing stage, we operate these valves so that only $Z0$ is open. Washing the ring, so that all the reagents that are currently occupying its volume exit through the waste line, requires that both $Z1$ and $Z2$ are open.

Reactions are initiated by combining reagents in the ring and are monitored by two-photon fluorescence. In order to get an estimate about the volume of the reagents required for each reaction, we need to estimate the volume of the mixing ring. In Chapter 3 we had estimated that the area of the cross section of the flow line is approximately 6.7×10^{-10} m^2. We also know that the circumference, C, of the mixing ring is related to the radius $r = 0.75$ mm via the equation:

$$C = 2\pi r = 2\pi \times 0.75 \times 10^{-3} \text{ m} \simeq 0.0047 \text{ m}. \tag{4.4}$$

If we multiply this number with the area of the cross section of the mixing ring, then the approximate volume, V, of the mixing ring is:

$$V = 6.7 \times 10^{-10} \text{ m}^2 \times 0.0047 \text{ m} \simeq 3 \times 10^{-12} \text{ m}^3 = 3 \text{ nl}. \tag{4.5}$$

Thus, we see that the amount of reagents needed for each reaction is of the order of 1 nl.

Our observation point for fluorescence measurements is right at the beginning of the ring, on the right side of the switch, as is shown in Figure 4.15(b). All the control valves are controlled and coordinated with the data acquisition process via software in C nidaqmx library (National Instruments), as will be explained in detail in Section 4.8.

As was explained above, the mixing time was found to be about 3 s. We found this by diluting Alexa488 dye in water and observing the fluctuations in fluorescence intensity while mixing. When the intensity at the observation point did not change with time, we knew that complete mixing had occurred. The time point where the steady-state value was recorded was 3 s after the pump had started mixing. If we had not facilitated mixing with the pumps, then mixing could only occur via diffusion (when the reagents are loaded, due to the small Reynolds number (see [91] and Chapter 3) mixing can only expected to occur via diffusion). If we assume that there are only two reagents in the ring that need to be mixed, each reagent taking up half the volume of the ring, then the maximum distance that a molecule from each of the reagents will need to diffuse will be a quarter of the ring's circumference (see Figure 4.16). Let's assume that this molecule is Rhodamine 6G which has a well-known diffusion coefficient $D = 300$ μm^2/s. Then, in the case of one-dimensional

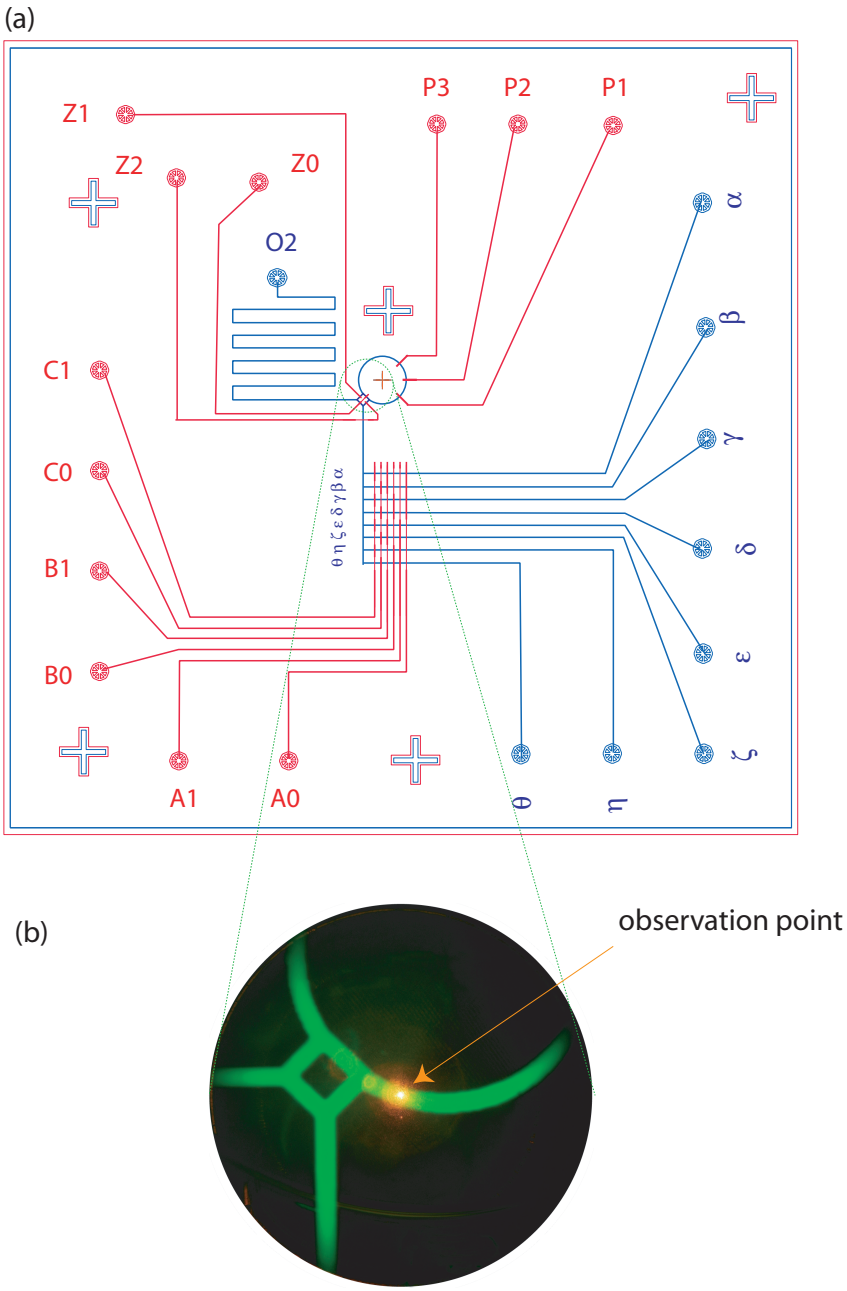


Figure 4.15: The microfluidics chip design. (a) The microfluidics chip was designed in Adobe Illustrator. The flow layer is shown in blue while the control layer is shown in red. The circles at the end of each of the lines indicate where holes need to be punched. These holes are the communication ports between the chip and its user. The input lines (α–θ) contain the different inputs that will be mixed inside the ring. The three control valves labeled $P1$–$P3$ make up the pump, which can accurately pump in specific amounts of the various inputs and mix them at a frequency of 33 Hz. The final product can then be discarded via the exit port labeled $O2$. The valves $Z0$–$Z2$ allow the pump to switch between loading the different inputs into the ring and mixing them. The green dashed circle denotes the region shown in (b). (b) Image of part of the mixing ring as observed under the microscope with a 5x objective. The flow lines are filled with fluorescein and a 1 Watt blue LED is used to excite the fluorescein dye so that it appears green. The control valves cannot be observed since they do not contain the dye. The bright orange point corresponds to the point of observation in the mixing ring for the two-photon microscope (the laser focal point in the plane of the chip/sample). As shown, the observation point is located at the beginning of the ring, on the right side of the switch.

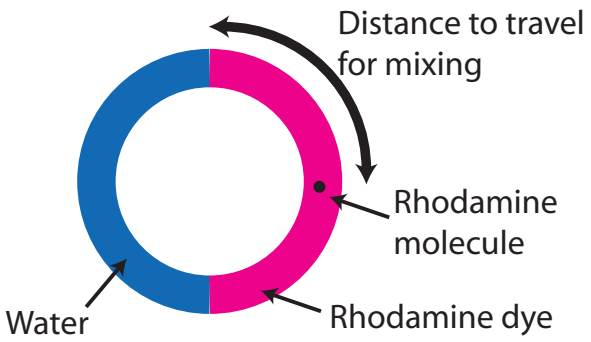


Figure 4.16: Cartoon for estimating diffusion in the ring (see text for the estimate). If the ring contains only two reagents that need to be mixed, then for mixing to occur a molecule will have to diffuse a distance equal to a quarter of the ring's circumference.

diffusion, the time this would take is (we are using the value of the circumference C from above):

$$t = \frac{x^2}{2D} = \frac{(0.25C)^2}{2D} = \frac{(0.25 \times 0.0047 \text{ m})^2}{2 \times 300 \times 10^{-12} \text{ m}^2/\text{s}} = 2300 \text{ s} \approx 40 \text{ min}, \tag{4.6}$$

about 770 times more than the pump's mixing time. This estimate is consistent with the one presented in [91]. In addition, the 3 s mixing time we get with the pump is very short compared to the time scale of the binding reactions. As will be seen, for example, in Figure 4.30, reactions are observed for a time period of 5 min, a time period that allows us to follow the complete reaction.

4.6.2 Chip Fabrication

The chip described was fabricated using the techniques of soft lithography, chip fabrication [119], and the Sylgard 184 Silicon Elastomer kit (Dow Corning Corporation).

Soft Lithography: For the fabrication of the molds, each layer of the chip was treated as follows:

- Flow Layer: The silicon wafer was spun with photoresist SPR $220-7$ at 1500 rpm for 60 s and was soft baked at 115°C for 90 s. The photoresist was then exposed to UV light for 90 s and was developed using developer 2401 at $1:5$ dilution. The wafer was then rinsed with H_2O and dried with N_2. In order to round the channels the wafer was hard baked at 115°C for 35 min. The final wafer (i.e., mold) had channel heights of approximately 12.5 μm, as determined by a profilometer.

- Control Layer: The silicon wafer was spun with photoresist SU8 $-$ 2025 at 2000 rpm for 60 s and was soft baked at 65°C for 2 min and then at 95°C for 5 min. The photoresist was then exposed to UV light for 60 s and was post-baked at 65°C for 60 s and at 95°C for 4 min. The wafer was then developed with SU8 developer, rinsed with Isopropanol and dried with N_2. The final mold had channel heights of approximately 28 μm, as determined by a profilometer.

Chip fabrication: Using the Sylgard 184 Silicon Elastomer kit (PDMS parts A and B from the supplier) we fabricated the push-down configuration chip as follows:

- Flow Layer: We prepared a mixture of $A : B$ in 20 : 1 (15 g part A and 0.75 g part B). The mixture was then poured on the flow layer mold and spun at 2500 rpm for 1 min (acceleration at 15 s). The wafer was then placed in an 80°C oven for 20 min to harden.

- Control Layer: We prepared a mixture of $A : B$ in 5 : 1 (25 g part A and 5 g part B). The mixture was then poured on the control layer mold and was placed in a vacuum chamber for approximately 30 min to degas it and remove any air bubbles from the mixture. The wafer was then placed in an 80°C oven for 20 min to harden.

After removing both molds from the oven, we peeled off the hardened PDMS from the control layer mold. Holes were punched at the input ports of the control layer, which was then aligned on top of the flow layer mold that still had the thin layer of hardened PDMS on it to form the chips. The chips were then incubated at 80°C oven for 1.5 hours to bind. After peeling off the chips from the flow layer mold, we punched the holes at the flow layer input ports, placed the chips on top of No. 1 Gold Seal cover glasses (33 mm x 60 mm, Reorder No. 3328) and incubated them overnight in the 80°C oven in order for the PDMS (i.e., chips) to bond to the cover slip. After removing the chips from the oven the next day, they were ready to be used.

4.6.3 Testing the Chip: Dilutions, Leaking, and Protein Sticking

When we use microfluidic chips, we need to ensure that their operation can be quantified. That is, we need to know the exact amount of dilution (so we can deduce the final concentration of the diluted product) and we also need to know that there is no loss in the amount of reagents, due to, for example, interactions between the sample and the microfluidic walls (sticking of the sample on the PDMS flow channels) or leaky valves (control valves that do not close fully and, as a result, the sample gets directed to the exit valve). Knowing that the sample is not lost is important for determining its concentration: if some of the S15 protein is sticking on the walls, then the amount that is participating in the reaction is less than the known concentration that we put in. In this section we are addressing the various tests and experiments that we have done in order to quantify the chip operation.

4.6.3.1 Finding the Dilution Factor

Every time we use a chip to conduct an experiment, we calibrate it by carrying out dye dilutions. We use Alexa488 at an initial concentration of 100 nM that we dilute in either 50 mM Tris pH = 7.6, 0.5% Tween-20 buffer [80], the same buffer we use when we dilute Mg^{2+}, or the buffer we use for the

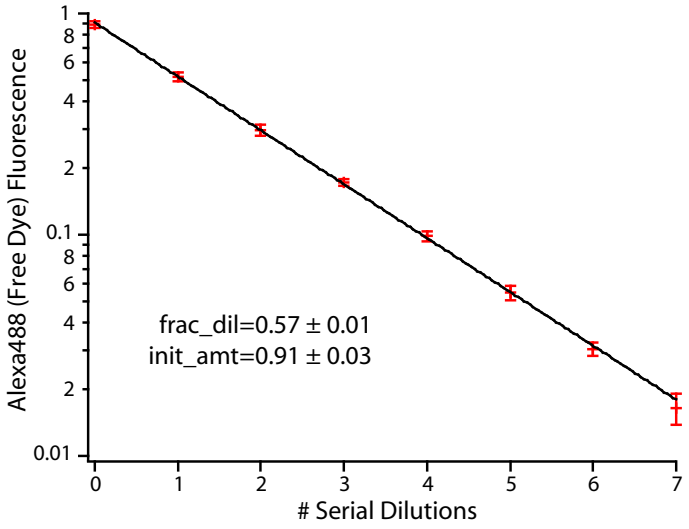


Figure 4.17: Calibration of serial dilutions with free dye (Alexa488) at an initial concentration of 100 nM. The dilution buffer in this case is 50 mM Tris pH = 7.6, 0.5% Tween-20 [80]. The x axis shows the number of dilutions while the y axis shows the fluorescent intensity (random units). Data points are shown in red and the solid line is the fit to the data points. The same dilution sequence was repeated 4 times. Each data point is the average of the corresponding data points from the 4 data sets and the error bars correspond to the standard deviation of each average. We see that all data points fall on the same straight line. The slope of this line gives the dilution factor. As an example, for this chip, we see that every time we do a dilution, we have diluted the dye by a factor of 0.57

S15 titrations (see Section 4.6.3.3). We conduct several dilutions and we run the dilution sequence several times in order to see the reproducibility of this calibration. Figure 4.17 shows one such calibration experiment. On the x axis we have the number of serial dilutions in the sequence while on the y axis we record the fluorescence intensity (the units are random and correspond to photon counts recorded by the photodiode). In this example, we repeated the same dilution sequence 4 times. Each data point is the average of the corresponding data points from the data sets. The error bars correspond to the standard deviation of each average. From the figure we see that we can easily obtain the dilution factor since all the data points fall on a straight line. In addition, from the size of the error bars, we see that this calibration experiment is highly reproducible. This way we can be sure that the chip at hand will always exhibit the same behavior. This type of calibration is done every time we use a new chip, in order to ensure that we have the corresponding dilution factor, as this may vary from chip to chip.

4.6.3.2 Ensuring That the Valves Are Not Leaking

This is perhaps one of the easiest steps in quality control. A leakage occurs when the control line is not fully closing the flow line it intersects. This can be caused by either a low pressure of the control valves that does not allow for complete closing or by damage of the flow or the control valves that

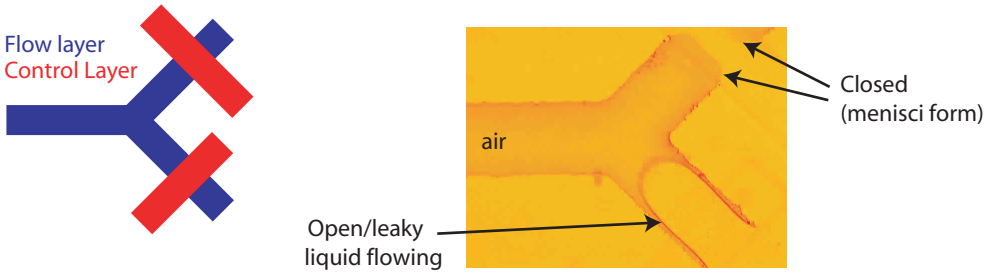


Figure 4.18: Figure showing what a leaky microfluidic valve looks like. The cartoon on the left shows what the design is of the picture on the right. As always, the flow layer is shown in blue and the control is shown in red. If both control valves were closed in the figure on the right, then the bottom valve will also form menisci, just like we see that the one on the top is forming. However, when a valve is open (depressurized) or is leaking because it is not pressurized enough, then the menisci do not form and we see the flow of liquid in the flow line underneath.

prevents closing from taking place. This problem is easily detectable because the chip is located on a microscope (see Section 4.8). When the flow and control lines are pressurized and are filled with the reagents and water, respectively, the chip is observed under the microscope with the $5x$ objective. If a valve is not closing, even though it is pressurized, then the pressure is increased until complete closing occurs. If that is still not the case, then we assume that the chip is damaged and we change the chip for a new one. Figure 4.18 shows what a channel looks like when it is closed (with the formation of the two menisci) and when it is open/leaking.

A leaky valve can also be identified even without looking at the chip under the microscope. Since the reactants that we have in the chip are rRNA and S15 that are labeled with dyes of different colors, if a valve is leaky then we will see a fluorescent signal from the reagent that is leaked through. For example, let's assume that the chip is operating so that only unfolded 3WJ rRNA is supposed to come into the mixing ring and the observation area. Then, we are only expecting to see signal from Atto425, the donor dye. However, if the valve corresponding to S15 is leaky, then we will observe signal coming from Alexa594 (the dye labeling S15) that is not supposed to be there! Thus, if a valve is leaking, it is easily detectable even during the experiment.

4.6.3.3 Addressing the Problem of Protein Sticking

This problem has been the hardest one to solve. The PDMS is hydrophobic and, therefore, the hydrophobic patches of the proteins tend to stick to the PDMS. We have tried several different techniques for resolving this problem, some of which have been more successful than others. In this section we discuss these techniques as well as the results we got.

The sticking of proteins on the PDMS material of the chips was first identified by looking at the signal we obtained from a known concentration of labeled S15-Alexa594 (200 nM) when the sample

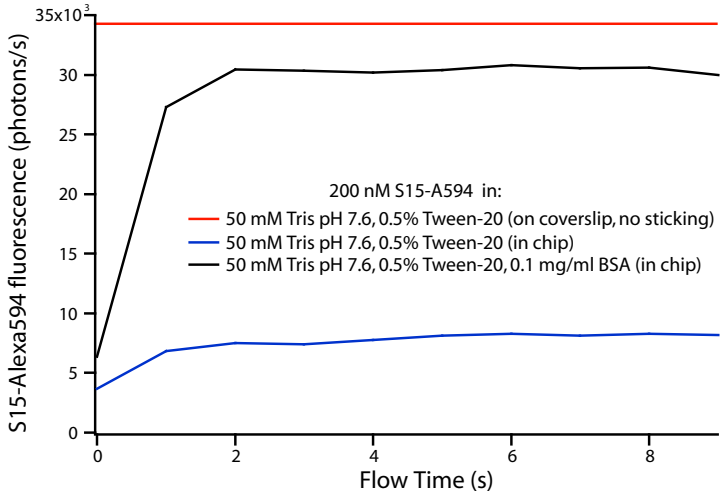


Figure 4.19: S15 flow inside the chip in two different buffers and on a glass coverslip. This figure shows the intensity observed when 200 nM of labeled S15 go inside the chip when compared to the signal we see when the sample is placed on top of a glass coverslip. On top of the coverslip, where there is no sticking, we get the maximum signal for the protein inside a buffer of 50 mM Tris pH = 7.6, 0.5% Tween-20 [80] (red curve). When observed inside the chip, in the case where the buffer is simply Tris-Tween-20 the observed intensity is lower (blue curve) than in the case where the buffer is Tris-Tween-20 and 100 μg/ml BSA (black curve). This shows that if the buffer does not contain BSA, then S15 sticks to the microfluidic channel walls much more than if the buffer does contain BSA. For this reason, the buffer used for the S15 titrations onto the 3WJ contain BSA. As is shown in Figure 4.21, the presence of BSA does not affect the S15 binding reaction onto the 3WJ.

was placed on a coverslip and when it was injected inside the chip. The buffer used was 50 mM Tris pH = 7.6, 0.5% Tween-20 [80] and the laser power was about 7 mWatts. We observed that the signal we obtained when the protein was inside the chip was much less than the signal observed when the protein was on top of the glass coverslip (see Figure 4.19, red and blue curves), indicating that there is a loss in the observed signal from the bulk of the sample due to the protein sticking on the walls of the microfluidic channels.

Quantifying the amount of protein that is stuck on the chip walls: When the problem of protein sticking on the walls was identified, we essentially had two choices. The first was to identify how much protein was sticking on the walls as a function of protein concentration. That is, for every dilution that we performed inside the mixing ring of the chip, there was an expected theoretical concentration, based on the dilution factor. Since the fluorescent signal detected was much less than expected based on the theoretical concentration, the difference between the expected signal and the actual signal would then give us the amount of the lost protein that was stuck on the walls. Therefore, it would provide us with the amount of protein that is left in the ring and can react. However, we soon found that *in situ* calibration based on the intensity signal was not possible. This is because the amount of protein that is sticking on the walls does not only depend on position (how far away the point of observation is from the injection point of the protein inside the

chip) but also on time and dilution. We found that the intensity signal of the S15 protein changed with time for the same amount of dilution. In other words, it was very dynamic (see Figure 4.20 for an example). Therefore, we could not make a reproducible calibration curve for the amount of protein sticking. Considering that our experimental goal was to observe the reaction of S15 binding on the 3WJ in real time, the time-changing concentration of the S15 protein due to sticking imposed a serious problem, since we did not know the initial concentration of S15 that was reacting. Furthermore, we also tried injecting extremely concentrated amounts of S15-Alexa594 protein (of the order of μM) in the hopes of saturating the walls. That is, we hoped that there would be a threshold amount of protein sticking above which no more protein would be adsorbed by the walls. However, exactly due to the dynamic nature of the sticking, we found that a saturation point did not exist. Even after some saturation was observed, a dilution of the protein concentration again changed the dynamics inside the chip and resulted in the S15-Alexa594 signal again changing with time. As an example, Figure 4.20 summarizes the dependence of the S15 signal on time and also the appearance of some saturation of the walls. In this example, the flow channels have been pre-coated via a 45 min incubation with a solution of 50 mM Tris pH = 7.6, 0.5% w/v Pluronic-F68, 0.1 mg/ml BSA (see Table 4.2). After incubation, a concentration of 3 μM S15 was sent through the ring and the protein signal was measured over time in 90 s intervals. We measured for a total of 10 90-s intervals. From the figure we see that the S15 signal changed significantly over time, indicating that the protein is continuously being lost to the chip walls. In addition, we see that in the last couple of 90 s intervals, a saturation point might have been reached, since the signal has stabilized. However, when a dilution of S15 was conducted so that the S15 concentration changed (data not shown), the dynamics of the protein-PDMS reaction changed and the saturation point was not valid any more. Since we are interested in the reaction kinetics of the S15 binding on the 3WJ, it is clear that waiting for a saturation point for each dilution would not work: by the time the saturation point is reached, the reaction will have already taken place.

Attempts to evaluate the concentration of S15 that was reacting via fluorescence correlation spectroscopy (FCS) presented several different problems and limitations. The first is that the use of FCS breaks down for high ($\sim$ 100 nM) S15 concentrations. FCS requires very small sample concentrations (so that only a few molecules are present in the detection volume at a time) in order for the signal fluctuations to be significant. However, how small is small? We know that the observation volume in our experimental setup of two-photon spectroscopy is of the order of $V = 1$ femtoliter (see Section 4.7). Literature suggests that the sample concentration is in the order of 1 nM. For our estimates, let's assume a concentration of $c = 10$ nM. This corresponds to a number of molecules N in the observation volume that is given by:

$$N = V \times c = 10^{-15} \text{ liters} \times 6.023 \times 10^{23} \times 10^{-8} \text{ molecules/liter} \simeq 6 \text{ molecules}. \qquad (4.7)$$

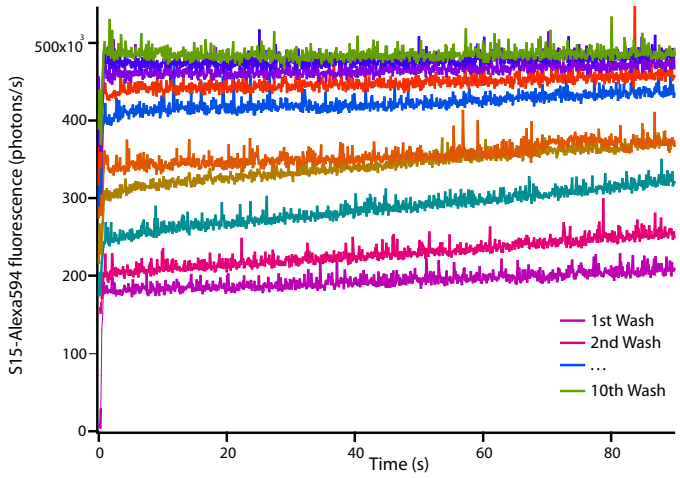


Figure 4.20: Time dependence of the S15 intensity signal. In this experiment, the flow channels have been pre-coated via a 45 min incubation with a solution of 50 mM Tris pH = 7.6, 0.5% w/v Pluronic-F68, 0.1 mg/ml BSA (see Table 4.2). After incubation, a concentration of 3 μM S15 was flown through the ring and the protein signal was measured over time in 10 90-s intervals. It can be seen that the fluorescent signal from S15 is changing over time, as the protein is lost on the walls of the PDMS chip. The signal seems to have stabilized in the last couple of 90 s intervals, however, this "saturation point" is not valid for a different S15 concentration.

If the sample concentration is increased to $c = 50$ nM, we easily see that the number of molecules in the observation volume increases to 30, at which point FCS measurements cannot be taken, since the contribution of each of the individual molecules to the total signal will not be significant. In our experiments, however, we are doing S15 titrations, where protein concentrations exceed 10 nM (for example, see Figure 4.21), and, therefore, the use of FCS will be limited to small S15 concentrations.

In addition, the time scale of the FCS measurement, when compared to those of the 3WJ-S15 binding reaction and the S15 adsorption to the walls, is such that obtaining the concentration of the protein that is actually reacting is a problem. It takes about 1.5 min to complete an FCS measurement, during which time the amount of protein in the sample will have changed significantly (see Figure 4.20) due to protein adsorption, and the 3WJ-S15 binding reaction will have occurred (see Figure 4.30). Within those time scales, and especially since we are interested in observing the binding reaction in real time, it will be difficult to decouple the two reactions and find the S15 concentration that is participating in the 3WJ-S15 binding reaction. Furthermore, as will be seen in Table 4.2, the amount of protein that is in the bulk of the chip (not adsorbed in the walls) is not necessarily equal to the amount of protein that is participating in the 3WJ-S15 binding reaction. Protein that is not stuck on the walls can still be inactive and not bind to the RNA. An FCS measurement, however, cannot distinguish between active and inactive protein. Because of these limitations, we found that the use of FCS for the quantification of the amount of protein that is sticking on the walls would not provide accurate results. Therefore, our next choice was to try and

eliminate the protein sticking completely. We have tried to achieve this in several different ways, by both using several different buffers for the sample and modifying the surface chemistry of the channel walls. We describe these methods in the following sections.

Using Different Buffers: As was explained above and shown in Figure 4.19, the use of a buffer that consists only of 50 mM Tris pH = 7.6, 0.5% Tween-20 [80] results in a lot of protein sticking on the channel walls. For this reason, we have tried to either incorporate in the buffer or substitute for the buffer reagents that have been known to help with this issue.

The use of Superblock as our buffer, by itself or combined with Tris-Tween-20, significantly reduces protein adsorption to the walls, but has also prevented us from seeing much change in FRET taking place. That is, we were unable to see any binding of the labeled S15 on the 3WJ. This behavior was very puzzling and, therefore, we conducted a control experiment where we injected a preformed complex of $[3WJ*S15]$ in the same buffer inside the chip. It was observed that within the time period of 1 min, the donor signal from the atto425 dye of the 3WJ became brighter while the signal from Alexa594 of the S15 got dimmer. In other words, the complex was being weakened, since FRET was decreasing. In conclusion, the use of Superblock weakens the complex and, therefore, a different method should be used for resolving the problem of protein sticking.

Incorporation of BSA in the 50 mM Tris pH = 7.6, 0.5% Tween-20 buffer seemed to be more effective. We tried 2 different BSA concentrations, 10 μg/ml and 100 μg/ml. We found that the signal we were getting from the S15-Alexa594 protein was about 30% less in the case of the lower BSA concentration when compared with the signal we were getting from the S15-Alexa594 protein when we used the higher BSA concentration of 100 μg/ml (results not shown). Therefore, we concluded that the higher BSA concentration was more successful in solving the protein adsorption problem. Furthermore, we found that with this BSA concentration we were able to recover about 82% of the signal that we were gettting when the protein was placed on top of a glass coverslip (compare the red curve with the black curve in Figure 4.19). In order to ensure that the presence of BSA does not affect the binding of S15 on the 3WJ, we performed S15 titrations on the 3WJ and observed the outcome using a plate reader (see Figure 4.21). Even though the value of $K_D = 54$ nM obtained in this case was about 5 times greater than the one that was observed without the presence of BSA (see Figure 4.14 where $K_D = 10.8$ nM), it was still within the range reported in the literature [134, 78, 80], indicating that BSA does not disturb the binding reaction. All these results are summarized in Table 4.1.

So far, we have found that BSA has been the most successful in solving the problem of protein sticking. However, we saw that even with the incorporation of BSA, the protein signal was still changing over time and with S15 dilution. Therefore, our next step was to improve on this result even more by trying chemicals other than Tris-Tween-20 and include in them the same BSA concentration of 100 μg/ml. Our goal was to either achieve 100% signal recovery (i.e., completely eliminate the

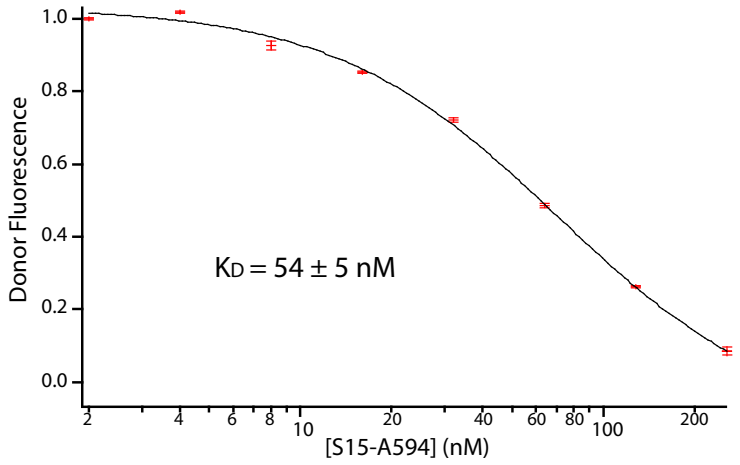


Figure 4.21: S15 titration into 7.5 nM 3WJ in the presence of 100 μg/ml BSA. The donor fluorescence is reported as a function of S15 concentration. The measurements were taken using the plate reader. Each data point is the average of 10 measurements and the error bars correspond to the standard deviation. From the value of $K_D \simeq 54$ nM, which agrees with the range of values (3.5–50 nM) reported in the literature [134, 78, 80], we see that the presence of BSA does not disturb the S15 binding reaction onto the 3WJ. Since BSA prevents the protein from sticking on the chip, we can use BSA in our reactions inside the chip without disturbing the reaction.

Chemical	Complete Buffer Conditions	Observations
Tween-20 (no BSA)	50 mM Tris pH 7.6, 0.5% w/v Tween-20	Continuous flow of the protein sample gave 25% of the expected signal. This value kept changing with time.
Tween-20 (0.01 mg/ml BSA)	50 mM Tris pH 7.6 0.5% w/v Tween-20 0.01 mg/ml BSA	Continuous flow of the protein sample gave 60% of the expected signal. This value kept changing with time.
Tween-20 (0.1 mg/ml BSA)	50 mM Tris pH 7.6 0.5% w/v Tween-20 0.1 mg/ml BSA	Continuous flow of the protein sample gave 80% of the expected signal. This value kept changing with time.
Superblock T20 (no BSA)	100% Superblock T20 in TBS formulation	Protein concentrations were high (not much stickage) and did not change much with time, but no reaction could be observed. Control experiments showed that Superblock broke apart pre-formed complex in the chip.
Superblock T20 with Tween-20 (no BSA)	25mM Tris pH 7.6 0.25% w/v Tween-20, 50% Superblock-T20	Mild protein stickage was observed along with moderate reactivity. The reaction would only go to about 10% completion, based on the observation of the donor's intensity. Also, from the plate reader titrations we observed a 10-fold increase in the K_D, indicating that the complex is being weakened by the presence of Superblock (see reaction kinetics figure at the end of the chapter).

Table 4.1: Summary of results for chemicals used for dynamic coating in order to resolve the protein sticking problem. In this case, the buffer description corresponds not only to the buffer used to do the dilutions for S15 but it is also the buffer in which 3WJ and S15 are delivered into the chip. The initial concentrations of 3WJ and S15 were 50 nM and 200 nM, respectively.

protein loss due to sticking) or stabilize the protein sticking with respect to time. First, we repeated the experiment with the 50 mM Tris pH = 7.6 0.5% w/v Tween-20, 0.1 mg/ml BSA buffer, but we initially pre-coated the chip channels with this buffer by performing a 45 min incubation. The pre-coating helped because the fluorescent signal from the protein did not depend on time. However, it still depended on the dilution factor, since every time we allowed buffer to go through the mixing ring in order to perform dilutions, the protein signal changed. Incorporating casein (1 mg/ml) in this 50 mM Tris pH = 7.6 0.5% w/v Tween-20, 0.1 mg/ml BSA buffer gave us the most signal from the labeled S15, indicating that it prevented the protein from adsorbing to the channel walls, and this signal was time-independent. In fact, the presence of casein in the buffer resulted in obtaining the maximum signal from S15. However, unfortunately, the presence of casein also prevented the reaction (the binding of S15 on the 3WJ) from taking place, as there was no change in the donor signal in the presence of S15 when compared to the donor signal in the absence of S15. Using Tween-80 in lieu of Tween-20 in the absence of casein did not help, as the signal obtained from the same S15 concentration was not high, indicating that there is a loss of protein to the walls. It was also observed that there was some protein that was unreactive.

Based on the work of Huang et al. [141], we used n-dodecyl-β-D-maltoside (DDM) for a final buffer of 50 mM Tris pH = 7.6, 0.5% w/v DDM, 0.1 mg/ml BSA in order to coat the chip channels with the same 45 min incubation. DDM, a very mild nonionic surfactant, adsorbs strongly on hydrophobic surfaces like PDMS and forms a monolayer, which causes the surface to become nonionic and hydrophilic. In this way, the interaction between the surface and the protein is reduced. With DDM, we observed the same signal recovery as with Tween-80. In addition, we found that we could not observe a change in FRET, indicating that a reaction could not take place. Furthermore, we found that in this case there was also instability coming from the signal of the 3WJ, as it kept changing with time. As a result, we concluded that with DDM a reaction could not occur. Finally, we found that washing the channels with this buffer did not clean them, as there was still an amount of protein that was lost due to adsorption.

Based on the work of Essö [142], Boxshall et al. [143], and from personal communication with the S. Quake group at Stanford, we also used poly(ethylene oxide) (PEO) derived surfactants Pluronic F68 and F127 (Sigma), which are known to reduce protein adsorption and cell adhesion by as much as 82% and 99%, respectively. The final buffer concentrations were 50 mM Tris pH = 7.6, 0.5% w/v Pluronic-F68, 0.1 mg/ml BSA and 50 mM Tris pH = 7.6, 0.5% w/v Pluronic-F68, 0.1 mg/ml BSA, respectively, and we used this buffer to pre-coat the channels via a 45 min incubation. We found that Pluronic F-127 resulted in a very low S15 signal (again, indicating that a significant amount of protein was lost to the walls) and behaved in the same way as DDM, since reaction could not be observed and the fluorescent signal from the 3WJ was changing with time. With Pluronic F-68 we obtained the minimum amount of protein signal and, again, we could not observe any FRET.

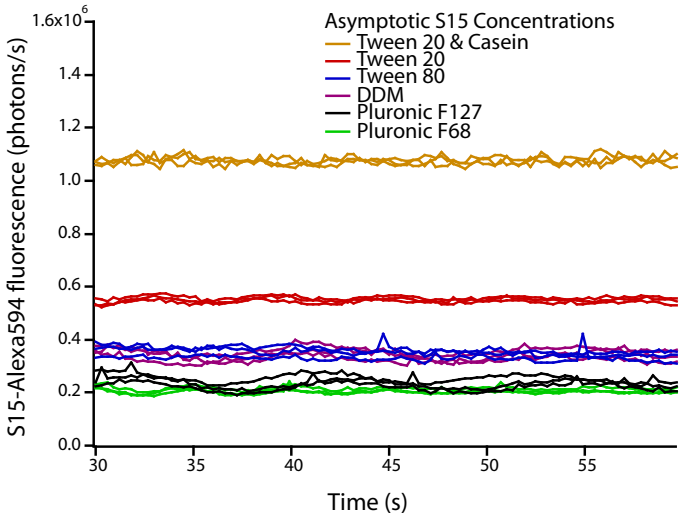


Figure 4.22: S15 fluorescent intensity in various buffers inside the chip. This figure shows the intensity observed when 3 μM of labeled S15 goes inside the chip at different buffers (all containing 0.1 mg/ml BSA) after a 45 min incubation with the respective buffer. Each color corresponds to a different buffer. Each measurement was repeated in triplicate. From this figure we can see that the Tris-Tween-20 buffer with 100 μg/ml BSA and 1 mg/ml casein results in the most S15 signal, suggesting that it has reduced the amount of protein adsorption the most. However, the presence of casein also prevents the S15 binding to the 3WJ from taking place (see text for details). These results and additional observations are summarized in Table 4.2.

Figure 4.22 summarizes and compares the S15 protein fluorescent signals with respect to each other. From this figure we see that the buffer containing Tween-20 with BSA and casein has performed the best as far as reducing the amount of protein that is stuck on the walls. However, as it turned out, prevention of protein adsorption is only half the story since, as we saw, the presence of casein also prevented S15 binding to the 3WJ. All these results and our observations for each of these chemicals are summarized in Table 4.2.

In order to emphasize how destructive the problem of protein adsorption to the walls is for our measurements (when we look at the binding of S15 to the 3WJ), we performed an additional experiment. After continuously flowing S15 protein through the mixing ring in order to obtain a "saturation point" (as was explained above and shown in Figure 4.20), we filled half the ring with 3WJ, mixed, and observed to see if a reaction (binding) would take place between the protein that was adsorbed on the walls and the added 3WJ. After the first observation, we again filled half the ring with 3WJ, mixed, and observed. We repeated this filling of half the ring with 3WJ 20 times and at no point did we add more S15 to the ring. The result was that each addition of 3WJ into the ring resulted in a reaction: the protein that was initially adsorbed to the walls was being removed from the walls and was binding to the RNA, resulting in a change in FRET. We did this experiment for

all the chemicals that have been described above. In Figure 4.23 (a) and (b) we present the results for the case of Tween-20 (without casein) and Tween-80, both of which were shown to recover a large amount of S15 signal (based on Figure 4.22) and also did not prevent the reaction from taking place. From these figures we see the extent of the problem: even though we are adding no protein S15 to the ring, but are only filling half of it at a time with 3WJ, a change in FRET (i.e., a reaction) is taking place because the 3WJ is reacting with whatever S15 protein is stuck on the walls of the mixing ring! Therefore, quantifying the amount of S15 that is actually reacting is getting even more complicated: the 3WJ is reacting not only with the S15 that is in the bulk of the sample, not stuck on the walls, but it can also react with the S15 that is adsorbed on the chip walls.

Modifying the Surface Chemistry: Based on the paper by Sui et al. [144], we modified the surface chemistry of the microfluidic devices. The channels were modified, first, with a solution-phase oxidation reaction by passing a mixture of $H_2O/H_2O_2/HCl$ in a ratio of $5:1:1$ continuously through the channels for 5 min. The channels were then washed with DI water and dried with dry Ar. Later, neat 2-[methoxy(polyethylenoxy)propyl]trimethoxysilane was injected in the channels for 1 h. The unreacted silane was flushed from the channels with DI water to give PEG-grafted channels. The channels were dried using dry Ar and were used after 24 h. In the earlier work, the chips were single (flow) layer chips and did not contain any control valves.

Even though this protocol reduced the protein sticking on the walls, we found this treatment to be very harsh on the chips, as the control valves failed to close the flow lines even when high pressures were applied. As a result, leakage was observed and, thus, the chips were rendered unusable (results not shown). Modifying the above protocol by reducing the amount of PEG-trimethoxysilane in the chip or substituting with Teflon-trimethoxysilane during the reaction still resulted in valves that were not functional. Finally, omitting the last PEG step completely from the process did not help. We concluded that the method described by Sui et al. can have very positive results for chips and designs that consist only of flow layers but are very harmful to chips that also have control valves. Personal communication with members of the S. Quake group at Stanford confirmed this conclusion. We summarize these results in Table 4.3.

In conclusion, the incorporation of blocking agents and detergents in the buffer has greatly decreased protein adsorption to the PDMS walls. However, we have found that this incorporation leads to two regimes: when we manage to decrease protein adsorption and obtain a high S15 signal, we normally observe minimal or no FRET, indicating that there is little or no reaction taking place. On the other hand, when FRET (i.e., reaction) is observed, we are in the regime where a lot of protein is lost to the walls. It seems that the challenge is to be able to find some middle ground.

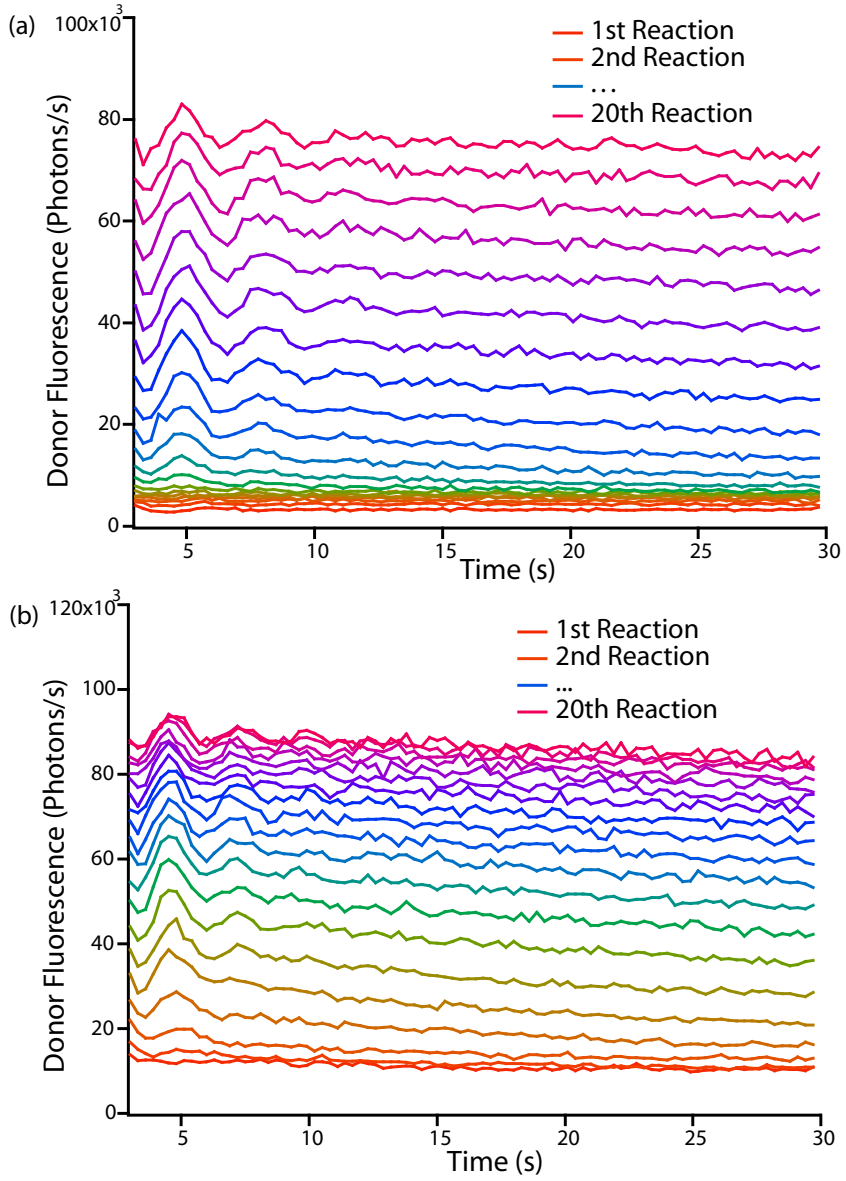


Figure 4.23: FRET induced by the binding of the 3WJ with the S15 that is stuck on the walls in the case of the (a) 50 mM Tris pH = 7.6, 0.5% Tween-20, 100 μg/ml BSA buffer and (b) 50 mM Tris pH = 7.6, 0.5% Tween-80, 100 μg/ml BSA buffer. In both cases the FRET is indicated by the change in the 3WJ donor signal. After the mixing ring has been saturated with S15, we fill half the ring with 3WJ, we mix and observe for 30 s. After this first reaction, we again fill half of the ring with 3WJ and observe for 30 s. We repeated for a total of 20 times (different colored curves). At no point did we add more S15 to the ring. Thus, we were observing the reaction of the binding of the S15 that was initially stuck to the walls on the added 3WJ. We see that initially, when the amount of S15 that was on the walls was maximum, the donor signal was minimum, indicating that FRET was taking place due to the binding of the S15. By the 20th half filling of the ring, the S15 that was stuck on the walls has been depleted (since most of it has reacted with the 3WJ) and, therefore, the donor signal is maximum, indicating that not a lot of reaction (FRET) is occurring. Since 3WJ can also react with S15 that is initially adsorbed on the walls, finding the S15 concentration that is taking place in the reaction is even more complicated, even for the case of the buffer that contains Tween-20 with BSA, which is found to reduce the S15 sticking to the walls significantly.

Coating Solution	Complete Buffer Conditions	Observations
Tween-20	50 mM Tris pH = 7.6 0.5% w/v Tween-20, 0.1 mg/ml BSA	Consistent with experiments in Table 4.1, continuous flow of the protein sample resulted in a significant amount of S15 signal (second best after the Tween-20/BSA/casein buffer). Reactions were observed, as assayed by donor quenching due to FRET (75%). However, there is definite evidence of protein sticking since, every time we flushed the chip with buffer, different amounts of protein were observed to have been removed from the walls, making the behavior very dynamic.
Tween-20 (with casein)	50 mM Tris pH = 7.6 0.5% w/v Tween-20, 0.1 mg/ml BSA, 1 mg/ml casein	This buffer gave the most S15 signal (see Figure 4.22), indicating that it had the best result as far as preventing S15 adsorption to the chip walls. This behavior was also time-independent. However, no reactions could be observed: the 3WJ (donor) signal was the same regardless of whether S15 was present or not.
Tween-80	50 mM Tris pH = 7.6 0.5% w/v Tween-80, 0.1 mg/ml BSA	Continuous flow of the protein sample gave a moderate amount of S15 signal indicating that there is a loss of protein due to sticking. However, flushing with buffer removed S15 protein from the walls making the behavior dynamic. Also, it was observed that some amount of protein was not reacting. Some reaction was observed, since the donor was quenched by roughly 50% due to FRET. These reactions lacked the fast phase of binding ($\sim 1s^{-1}$) possibly reflecting interference of the detergent with the 3WJ folding.
Pluronic-F68	50 mM Tris pH = 7.6 0.5% w/v Pluronic-F68, 0.1 mg/ml BSA	Continuous flow of the protein sample gave poor S15 signal, again indicating loss of protein on the channel walls. No reaction was observed, as assayed by donor quenching (FRET).
Pluronic-F127	50 mM Tris pH = 7.6 0.5% w/v Pluronic-F127, 0.1 mg/ml BSA	Continuous flow of the protein sample gave poor S15 signal, suggesting that a lot of sticking was taking place. A very small amount of FRET (i.e., reaction) was observed by donor quenching. The system was very unstable since the initial signal from the 3WJ RNA value changed a lot with time. Also, the protein could not be washed away from the walls when flushing the chip with buffer.
DDM	50 mM Tris pH = 7.6 0.5% w/v DDM, 0.1 mg/ml BSA	Continuous flow of the protein sample gave a moderate amount of S15 signal. Just like with Pluronic-F127, there was not much donor quenching (FRET) observed and the system was very unstable since the signal from the 3WJ changed a lot with time. The protein could not be washed away from the walls when flushing the chip with buffer.

Table 4.2: Summary of results for non-ionic detergents and blocking agents used for pre-coating in an attempt to resolve the protein adsorption problem. Pre-coating was accomplished by flushing the chip for 45 min with the coating solution. The same buffer description also corresponds to the buffer used to do the dilutions for S15 and to the buffers for the initial concentrations of 3WJ and S15. The initial concentrations of 3WJ and S15 were 50 nM and 3 μM, respectively, for all the buffers described in this table.

Chemical/Reaction	Description	Observations
H_2O : H_2O_2 : HCl oxidation coupled with PEG (2-step reaction)	5 : 1 : 1 (5 min), washed with H_2O, dried with Ar, followed by 1 hour flow of neat (undiluted) PEG-trimethoxysilane (100%)	Control valves would not fully close, resulting in leakage and contamination of measurement (i.e., the chip was non-functional).
H_2O : H_2O_2 : HCl oxidation coupled with PEG (2-step reaction)	5 : 1 : 1 (5 min), washed with H_2O, dried with Ar, followed by 1 hour flow of PEG-trimethoxysilane(95% diluted in H_2O)	Control valves would not fully close, resulting in leakage and contamination of measurement (i.e., the chip was non-functional).
H_2O : H_2O_2 : HCl oxidation coupled with Teflon (2-step reaction)	5 : 1 : 1 (5 min) washed with H_2O, dried with Ar, followed by 1-hour flow of neat (undiluted) Teflon-trimethoxysilane (100%)	Control valves would not fully close, resulting in leakage and contamination of measurement (i.e., the chip was non-functional). In addition, the flow lines would not fill with the samples.
H_2O : H_2O_2 : HCl oxidation	5 : 1 : 1 (5 min), washed with H_2O, dried with Ar	Control valves would not fully close, resulting in leakage and contamination of measurement (i.e., the chip was non-functional). Verified via personal communication with the S. Quake group.

Table 4.3: Summary of surface chemistry reactions used in order to resolve the protein sticking problem. In all cases the buffer used for the 3WJ and S15 samples was 50 mM Tris pH = 7.6 0.5% w/v Tween-20. The initial concentrations of 3WJ and S15 were 50 nM and 500 nM, respectively.

4.7 Two-Photon Spectroscopy (TPS)

The detection of the reaction kinetics and dissociation constants of the rRNA folding and S15 binding that are taking place inside the microfluidics chip is done with FRET and via the use of a microscope that is custom built, as will be described in Section 4.8. For the reasons described above in Section 4.3, we are using two-photon excitation spectroscopy.

The theory for the simultaneous absorption of two photons, what is now called two-photon spectroscopy, was first developed by Maria Goeppert-Mayer in 1931 [145]. The two-photon spectroscopy method is based on the idea that two photons of half the energy (double the wavelength) can be used in lieu of the one photon it would take to excite the fluorophore (one-photon excitation). The fluorophore can then emit a photon that is of higher energy (smaller wavelength) than either of the two photons that were absorbed, resulting in this way in a spectral separation between the exciting laser photons and the emitted one coming from the fluorophore. In order for the fluorophore to be excited, the two photons need to be absorbed within the cross section of the fluorophore (10^{-16} cm^2) in a time interval of 10^{-15} s (1 fs) [138]. The simultaneous absorption of two photons is an event that has a very low probability (quantum forbidden event). As a result, in order to ensure the occurrence of such an absorption, the photon flux must be very high and, based on the above number, of the order of 10^{31} photons/cm^2s . For this reason it is required that we have a high laser power and, also, pulsed excitation. Pulsed excitation is characterized by higher photon densities per pulse relative to the average laser power and, therefore, it increases the probability that two photons will be absorbed at the same time. Such high photon densities are achieved with the use of mode locked lasers (see Section 4.8), which provide short pulses ($\sim$ 100 fs) at high rates ($\sim$ 100 MHz).

Since two photons need to be absorbed in order for the excitation to occur, the probability of emission P_{em} is proportional to the mean square of the intensity I_{ex} of the excitation beam:

$$P_{\text{em}} \propto I_{\text{ex}}^2. \tag{4.8}$$

This ensures that the resulting fluorescence is generated in the immediate vicinity of the objective's focal point (i.e., where the beam is tightly focused), at a volume of about $V = 1$ femtoliter, eliminating excitation of out-of-focus fluorophores. In contrast, in the case of one-photon excitation, the probability of emission is proportional to the mean of the intensity of the excitation beam:

$$P_{\text{em}} \propto I_{\text{ex}}. \tag{4.9}$$

Therefore, fluorescence is generated over a larger sample volume that is extending past the objective's focal point, resulting in background fluorescence. This difference between two-photon and one-photon excitation is shown well in Figure 4.24, where it is clear that in the two-photon case only the

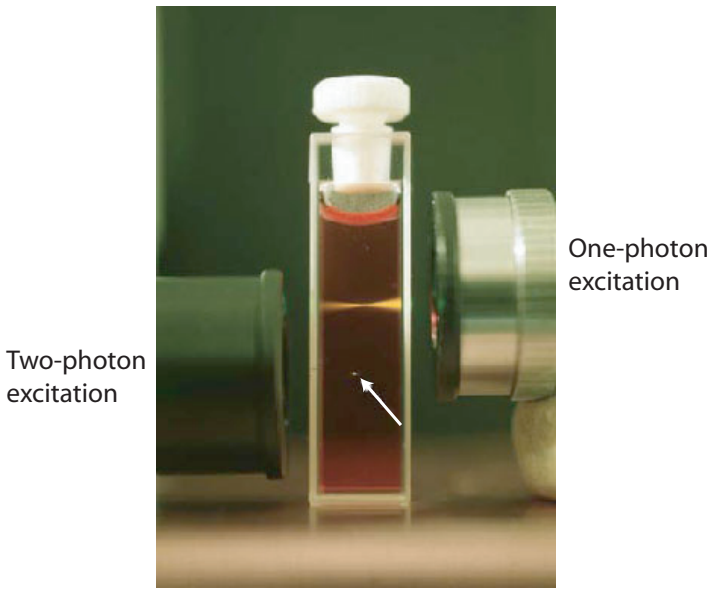


Figure 4.24: Volume excitation in one-photon vs. two-photon spectroscopy. The cuvette contains dye which the upper right objective excites with one-photon excitation while the bottom left objective excites with two-photon excitation. We can see that in the one-photon case essentially all the photon along the beam path are excited, not just those that are in the objective's focal point. In contrast, in the two-photon case only the fluorophores within the tight focal point are excited, as indicated by the white arrow. Figure taken from [146].

fluorophores that are in the focal point are excited. This elimination of background fluorescence, combined with the fact that the longer photon wavelengths used suffer less scattering, result in a significant increase in the signal-to-noise ratio, which is important for sample concentrations that are small, as is our case when we do titrations (nanomolar concentrations).

4.8 The Experimental Setup

Based on the requirements described above, the two-photon excitation fluorescent microscope that we used for the purposes of the experiment is shown in Figure 4.25. For the two-photon excitation, the laser used is a mode locked Coherent Mira 990-F Ti:Saphire (Santa Clara, CA), that is tunable for the wavelength range of 700–980 nm but is typically used in the range of 860–900 nm. The pulse width used is approximately $\sim$ 150 fs. The laser beam is spatially filtered and, in order to overfill the back aperture of the high numerical aperture (N.A.) microscope objective, it is expanded 4.5x. The objective used is an oil-immersion Zeiss, alpha Plan-FLUAR, of magnification equal to 100x with N.A.= 1.45. In order to avoid the saturation of the small focal volume (0.3 fl), the laser excitation power is kept low, within the range of 3–10 mW to avoid exciting the fluorophores that label the S15 protein and the 3WJ of the rRNA to the point of saturation. The excitation radiation is directed into the microscope via a dichroic mirror (730dcspx from Chroma) and is focused on

the sample. The fluorescent light emitted from the sample is then collected by the same objective and passed through the dichroic. Since we are also looking at reactions where both partners are labeled with up to three different colors, we split the three colors of the emitted light using a second dichroic filter (530dcxr from Chroma), and the stray excitation light is also blocked via the use of band-pass filters (HQ580/60m-2p, HQ485/70m-2p from Chroma). In order to block any fluorescent light that does not originate from the focal volume, pinholes are positioned at the image plane. An avalanche photodiode, APD, (SPCM-AQR-14 from PerkinElmer Optoelectronics) is used for the detection of each of the colors. A multifunction data acquisition card (PCLe-6251 from National Instruments) receives the TTL pulses (i.e., the data) from the three photodiodes and also controls the actuation of the solenoid manifold that is connected to the microfluidics device described in Section 4.6. Thus, there is a coordination between the optical detection and the operation of the microfluidics chip. The acquired data is analyzed in real time to produce the autocorrelation and cross-correlation integrals. All functions of the data acquisition card (microfluidics control, data acquisition and analysis) are performed by programs written in C/gtk. In order for the microfluidics chip to be positioned properly with respect to the two-photon microscope (i.e., in order to align the chip so that the laser focal volume is inside the mixing ring) a low (5x) magnification objective is located along the same optical axis. This microscope and the related programs are work done by William Ridgeway from the J. Williamson laboratory at The Scripps Research Institute, La Jolla, CA.

The stage that the microfluidic chips are placed on is temperature controlled. The chip is fine-aligned by detecting two-photon fluorescence from the microfluidic channels and translating the chip so that the focal volume of the microscope is centered in the plane of the chip and halfway up the microfluidic channel of 12.5 μm. At the highest laser powers used ($\sim$ 15 mW) the PDMS does not exhibit any detectable autofluoresence, nor does the PDMS-water interface scatter light. The most significant source of background signal is the fluorescent protein which is stuck to the walls. The selective excitation achieved by 2-photon fluorescence allows us to image within the microfluidic channels even in the presence of lots of protein stuck to the walls. (From FCS calibration experiments, the axial dimension of the 2-photon excitation volume is < 3 μm.)

4.9 Finding the Transition Rates

So far we have established that our experimental observations are based on FRET and that the goal of the experiments is to obtain reaction rates and dissociation constants in an automated way that will allow for our setup to be used in order to study the complete self-assembly of the 30S subunit. In this section we explain how our experimental data, which is in terms of donor and acceptor signals, can be used in order to obtain the quantities of interest, which we proceed to derive.

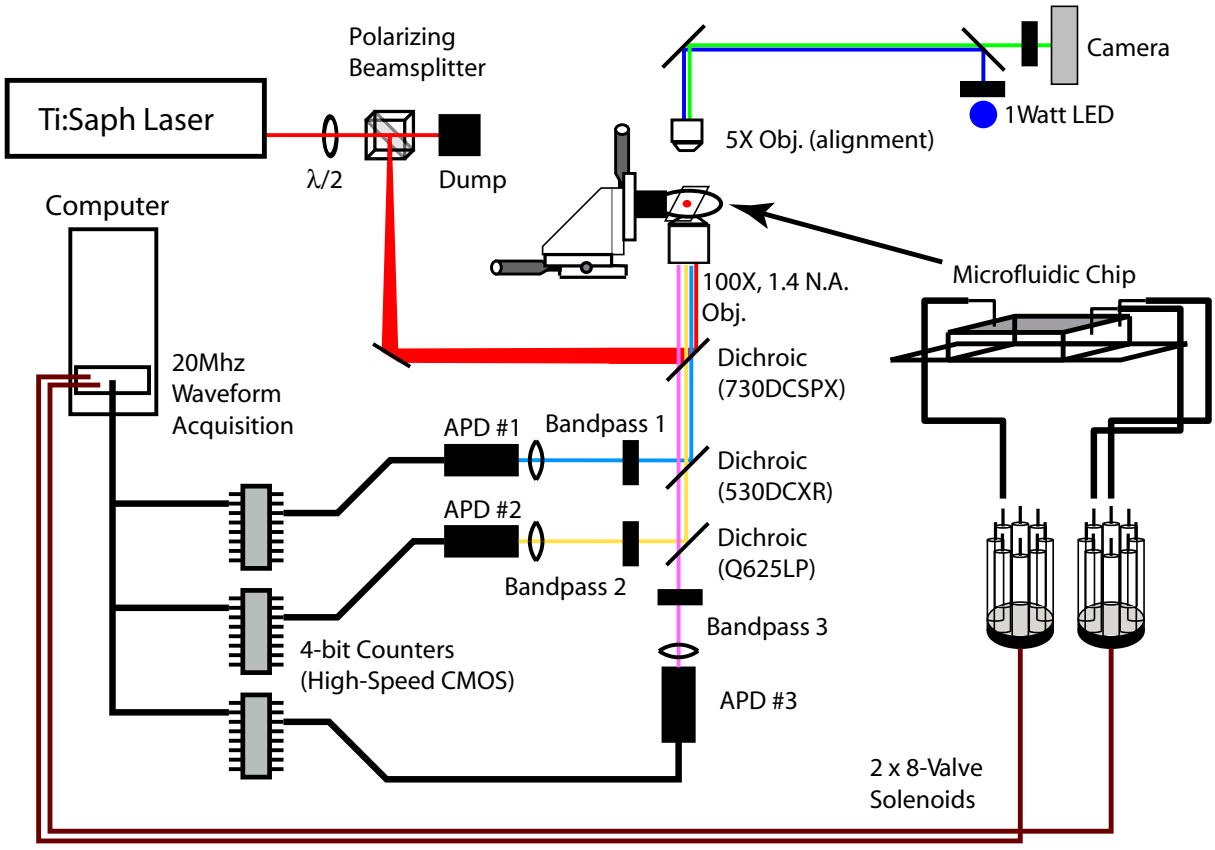


Figure 4.25: The experimental setup. This cartoon shows each of the components of the setup. The laser's incident beam is expanded and directed towards the microscope objective, which focuses it onto the microfluidics chip that contains the sample. The fluorescent dyes of the molecules are excited and then emit light, which is directed and focused onto the APDs via dichroic mirrors and lenses. A data card controls the data acquisition process and also records and analyzes the data.

4.9.1 K_D for Mg^{2+} Titrations

As has already been mentioned, the folded conformation of the 3WJ is favored in the presence of Mg^{2+}. If we assume that n Mg^{2+} ions interact with the 3WJ, then the binding reaction that describes this scheme is given by:

$$[3WJ]_u + n[Mg^{2+}] \overset{K_D}{\rightleftharpoons} [3WJ]_f, \tag{4.10}$$

where $[3WJ]_u$ and $[3WJ]_f$ are the unfolded and folded 3WJ concentrations, respectively, and K_D is the dissociation constant. At equilibrium, the ratio of folded 3WJ to total 3WJ is given by the Hill equation:

$$\frac{[3WJ]_f}{[3WJ]_u + [3WJ]_f}. \tag{4.11}$$

We will use the definition of the dissociation constant in order to solve for the unknown concentration of the folded 3WJ. We have:

$$K_D = \frac{[3WJ]_u[Mg^{2+}]^n}{[3WJ]_f} \Rightarrow [3WJ]_f = \frac{[3WJ]_u[Mg^{2+}]^n}{K_D}. \tag{4.12}$$

Substituting this result to Eq. 4.11 gives:

$$\frac{[3WJ]_f}{[3WJ]_u + [3WJ]_f} = \frac{\frac{[3WJ]_u[Mg^{2+}]^n}{K_D}}{[3WJ]_u + \frac{[3WJ]_u[Mg^{2+}]^n}{K_D}} = \frac{\frac{[[Mg^{2+}]^n}{K_D}}{1 + \frac{[Mg^{2+}]^n}{K_D}} \tag{4.13}$$

In Figure 4.26 we show what this last fraction of folded 3WJ looks like based on equation 4.13. We see that for low $[Mg^{2+}]$ almost none of the 3WJ is in the folded state. In addition, the fraction of folded 3WJ increases with $[Mg^{2+}]$. This ratio approaches 1, indicating that the 3WJ is almost completely folded, for high $[Mg^{2+}]$. Experimentally, we can obtain the quantity $\frac{[3WJ]_f}{[3WJ]_u+[3WJ]_f}$ by looking at the signal from the acceptor dye Atto532 and the donor dye Atto425. That is, the signal from the acceptor corresponds to how much of the 3WJ is in the folded state and the signal from the donor corresponds to how much of the 3WJ is in the unfolded state. Therefore, the fraction of the 3WJ that is in the folded state corresponds to the ratio of acceptor532/(donor425 +acceptor532). Our data consists of this ratio as a function of $[Mg^{2+}]$. Thus, by fitting Eq. 4.13 to our experimental data using K_D and n as the fitting parameters, we can obtain their values. This is the fit equation that we used in Figures 4.11b and 4.29 in order to obtain K_D. By doing those fits, we have found that $n = 1$.

Thus, in this section we have shown how to obtain the dissociation constant K_D in the case of the 3WJ folding due to the presence of Mg^{2+}.

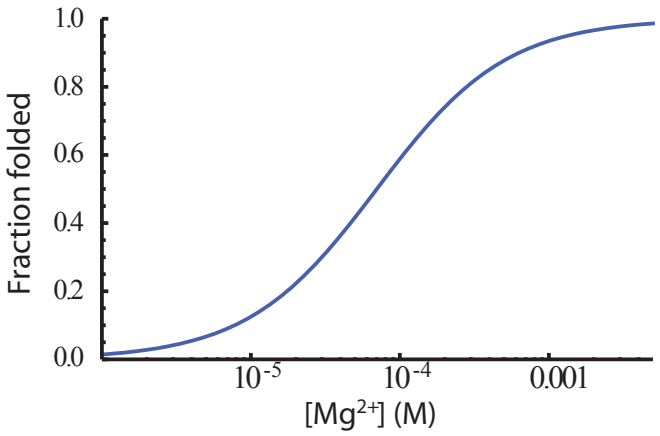


Figure 4.26: Fraction of folded 3WJ as a function of Mg^{2+} concentration from the Hill equation. The fraction of 3WJ that is bound by Mg^{2+} (i.e., the folded 3WJ) is shown as a function of $[Mg^{2+}]$. The plot has the well-known sigmoidal shape, as is given by the Hill equation. For this graph we assume that $n = 1$ and $K_D = 70$ μM.

4.9.2 Transition Rates and K_D for S15 Binding

In the above section we found the dissociation constant for the case of the conformational change of the 3WJ rRNA induced by the presence of Mg^{2+}. In the case of the binding of the S15 protein onto the 3WJ, however, the equation that describes the reaction is:

$$[3WJ] + [S15] \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [3WJ * S15], \tag{4.14}$$

where all the concentrations are dependent on time. For simplicity, we define the following: $a(t) \equiv [3WJ], b(t) \equiv [S15], c(t) \equiv [3WJ * S15]$, and the above equation becomes:

$$a(t) + b(t) \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} c(t). \tag{4.15}$$

Therefore, the differential equation that describes the change in [3WJ*S15] complex as a function of time is:

$$\frac{dc(t)}{dt} = k_{on}a(t)b(t) - k_{off}c(t). \tag{4.16}$$

At equilibrium, the above equation becomes:

$$\frac{dc(t)}{dt} = 0 \Rightarrow k_{on}a_{eq}b_{eq} - k_{off}c_{eq} = 0 \Rightarrow k_{on}a_{eq}b_{eq} = k_{off}c_{eq}, \tag{4.17}$$

where we have defined a_{eq}, b_{eq}, c_{eq} to be the equilibrium concentrations of [3WJ], [S15], and [3WJ*S15], respectively. With this in mind, if we are close to equilibrium, then we can use perturbation theory and define the time-dependent concentrations to be a slight perturbation from the equilibrium

concentrations. That is:

$$a(t) = a_{eq} + \delta a \tag{4.18}$$

$$b(t) = b_{eq} + \delta b \tag{4.19}$$

$$c(t) = c_{eq} + \delta c \tag{4.20}$$

where we have defined δa to be the change in a from its equilibrium concentration a_{eq}, and similarly for δb and δc. However, because of the stoichiometric Eq. 4.15, we have that a change in a has to match the change in b in order to produce the same change in c. In equation form, that translates to:

$$\delta a = \delta b = -\delta c. \tag{4.21}$$

Therefore, now by substituting Eqs. 4.18–4.20 into equation 4.16 we get:

$$\frac{d(c_{eq} + \delta c)}{dt} = k_{on} a(t) b(t) - k_{off} c(t) \Rightarrow \frac{d(\delta c)}{dt} = k_{on}(a_{eq} + \delta a)(b_{eq} + \delta b) - k_{off}(c_{eq} + \delta c), \tag{4.22}$$

where we have used the fact that $\frac{dc_{eq}}{dt} = 0$, since the equilibrium concentration does not change with time. This last equation, with the use of Eq. 4.21, becomes:

$$\frac{d(\delta c)}{dt} = k_{on}(a_{eq} + \delta a)(b_{eq} + \delta b) - k_{off}(c_{eq} + \delta c) \Rightarrow \frac{d(\delta c)}{dt} = k_{on}(a_{eq} - \delta c)(b_{eq} - \delta c) - k_{off}(c_{eq} + \delta c). \tag{4.23}$$

By expanding this last equation and grouping terms based on the order of δc we get:

$$\frac{d(\delta c)}{dt} = k_{on}(\delta c)^2 - [k_{on}(a_{eq} + b_{eq}) + k_{off}]\delta c + k_{on} a_{eq} b_{eq} - k_{off} c_{eq}. \tag{4.24}$$

The first term of this equation is in second order for δc and, therefore, very small. Thus, we can ignore it. Also, based on Eq. 4.17, the last two terms cancel each other out. We are left with:

$$d(\delta c)/dt = -[k_{on}(a_{eq} + b_{eq}) + k_{off}]\delta c. \tag{4.25}$$

If we define an apparent reaction rate k_{app} to be equal to:

$$k_{app} = k_{on}(a_{eq} + b_{eq}) + k_{off}, \tag{4.26}$$

which is linearly dependent on the equilibrium concentration of S15 (i.e., b_{eq}) then the initial differential Eq. 4.16 has now reduced to equation:

$$d(\delta c)/dt = -k_{app}\delta c, \tag{4.27}$$

and the solution is:

$$\delta c = \lambda_0 e^{-k_{app}t} \Rightarrow c(t) - c_{eq} = \lambda_0 e^{-k_{app}t} \Rightarrow c(t) = c_{eq} + \lambda_0 e^{-k_{app}t}, \tag{4.28}$$

where we have made use of Eq. 4.20 and λ_0 is an integration constant. However, at $t = 0$, no complex has formed. Therefore, at $t = 0$:

$$c(0) = 0 \Rightarrow \lambda_0 = -c_{eq}. \tag{4.29}$$

Thus, Eq. 4.28 becomes:

$$c(t) = c_{eq}(1 - e^{-k_{app}t}) \Rightarrow \frac{c(t)}{c_{eq}} = 1 - e^{-k_{app}t}, \tag{4.30}$$

where in the end we have normalized the concentration of the complex by dividing by the equilibrium concentration of the complex. In Figure 4.27 we show what the concentration $[3WJ * S15] \equiv c(t)$ looks like with respect to time, based on Eq. 4.30. In this case, we have randomly picked $k_{app} = 0.1$ s^{-1}. We see that $c(t)$ is increasing with time, indicating that the reaction is progressing and more of the 3WJ*S15 complex is forming. In order to obtain this last Eq. 4.30, we use an approach based on perturbation from the equilibrium concentrations (see Eqs. 4.18–4.20). Obtaining $c(t)$ using a non-perturbative approach would require finding numerical solutions for the set of differential equations:

$$\frac{da(t)}{dt} = k_{off}c(t) - k_{on}a(t)b(t) \tag{4.31}$$

$$\frac{db(t)}{dt} = k_{off}c(t) - k_{on}a(t)b(t) \tag{4.32}$$

$$\frac{dc(t)}{dt} = k_{on}a(t)b(t) - k_{off}c(t). \tag{4.33}$$

Such solution for $c(t)$ was obtained using Mathematica (data not shown). When plotted, this solution agreed with the one obtained using the perturbation approach described in this section and shown in Figure 4.27, indicating that this approach can accurately describe the time evolution of the concentration of the complex.

Experimentally, we obtain the time trace of the variable $c(t) \equiv [3WJ * S15]$ by looking at the FRET signal of the acceptor/(donor+acceptor), as is shown in Figure 4.30. The acceptor signal corresponds to the amount of complex that has formed, which is $c(t)$ normalized over the total signal observed. Therefore, with this data trace of $c(t) \equiv [3WJ * S15]$ in hand, we can fit to that Eq. 4.30, in order to obtain the apparent rate constant k_{app}. Also, we know the value of $[S15]_{eq} \equiv b_{eq}$ because we know how much complex we have (i.e., how much S15 has bound and we also know how much initial concentration $[S15]_0$ we put in. Thus, $[S15]_{eq} = [S15]_0 - [3WJ*S15]_{eq}$. Thus, from that constant k_{app} and since we know the equilibrium concentration $[S15]_{eq} \equiv b_{eq}$ we can use Eq. 4.26 in

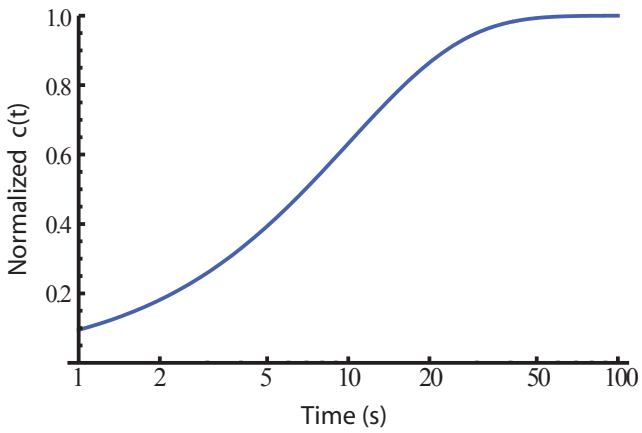


Figure 4.27: Concentration of the complex 3WJ*S15 as a function of time based on Eq. 4.30. The x axis is on a logarithmic scale. For this simple graph, we have picked $k_{app} = 0.1$ s^{-1}. We see that as time progresses, more and more of the complex $[3WJ * S15] \equiv c(t)$ has formed.

order to obtain k_{on} and k_{off} from the slope and the intercept of the k_{app} vs. $[S15]_{eq} \equiv b_{eq}$ curve. Thus, both these rate constants can be obtained at once! The figure obtained via this process is shown as Figure 4.31. Knowing k_{on} and k_{off} can also give us the dissociation constant for S15 binding since, by definition:

$$K_D = \frac{k_{off}}{k_{on}}. \tag{4.34}$$

We can obtain K_D in a different way. In the case where the reaction has reached equilibrium, again we define $[3WJ]_{eq}$, $[S15]_{eq}$, and $[3WJ*S15]_{eq}$ to be the equilibrium concentrations of the 3WJ, S15 and the 3WJ*S15 complex, respectively, and $[3WJ]_0$ and $[S15]_0$ to be the initial concentrations of the 3WJ and the S15, respectively. Of course, initially, $[3WJ * S15]_0 = 0$. We are interested in the fraction of the 3WJ that has been bound by the S15 and has formed the complex:

$$\frac{[3WJ]_{eq}}{[3WJ]_0} \tag{4.35}$$

Experimentally, we get this quantity by looking at the donor (atto425) fluorescence before and after we put S15 into the microfluidic ring. The fluorescence of the donor without the presence of S15 indicates the initial concentration of 3WJ $[3WJ]_0$, since none of the 3WJ is yet bound, while the donor fluorescence with the presence of S15, due to FRET, will be smaller and will indicate how much of the 3WJ has S15 bound. That is, the final donor intensity will correspond to the equilibrium concentration of the 3WJ, $[3WJ]_{eq}$. In other words:
Intensity of the donor without S15: $[3WJ]_0$
Intensity of the donor with S15: $[3WJ]_{eq}$.

Naturally, due to mass conservation, the equilibrium concentration of the 3WJ will equal the

initial concentration minus the concentration of the 3WJ that is now in the complex. Similarly for the equilibrium concentration of S15. That is:

$$[3WJ]_{eq} = [3WJ]_0 - [3WJ * S15]_{eq} \tag{4.36}$$

$$[S15]_{eq} = [S15]_0 - [3WJ * S15]_{eq}. \tag{4.37}$$

By combining Eqs. 4.36 and 4.35, we find:

$$\frac{[3WJ]_{eq}}{[3WJ]_0} = \frac{[3WJ]_0 - [3WJ * S15]_{eq}}{[3WJ]_0} = 1 - \frac{[3WJ * S15]_{eq}}{[3WJ]_0}. \tag{4.38}$$

Our ultimate goal is to obtain the dissociation constant K_D and compare it to the range of values 3.5–50 nM found in the literature [134, 78, 80]. If our K_D is within that range, then this control experiment is successful and we can move on to looking at the rates k_{on} and k_{off} and at reactions with other proteins. K_D, from definition, is given by:

$$K_D = \frac{[3WJ]_{eq}[S15]_{eq}}{[3WJ * S15]_{eq}}. \tag{4.39}$$

Since during the reaction we have control over the initial concentrations that we put in but we do not know what the concentration of the complex is at the end of the reaction, we define that to be x. That is: $[3WJ * S15]_{eq} \equiv x$. Then, based on this and Eqs. 4.36 and 4.37, Eq. 4.39 becomes:

$$K_D = \frac{([3WJ]_0 - x)([S15]_0 - x)}{x}. \tag{4.40}$$

Our goal is to solve for the unknown equilibrium concentration of the complex, x. Therefore, if we expand the above equation we get a quadratic equation:

$$x^2 - ([3WJ]_0 + [S15]_0 + K_D)x + [3WJ]_0[S15]_0 = 0. \tag{4.41}$$

The solutions to this quadratic equation are:

$$x = \frac{([3WJ]_0 + [S15]_0 + K_D) \pm \sqrt{([3WJ]_0 + [S15]_0 + K_D)^2 - 4[3WJ]_0[S15]_0}}{2}. \tag{4.42}$$

We can only accept the negative root because the positive root would imply that the equilibrium concentration of the complex $[3WJ * S15]_{eq} \equiv x$ is more than the initial concentration of the 3WJ, $[3WJ]_0$, which contradicts mass conservation. Therefore, we have that:

$$x = [3WJ * S15]_{eq} = \frac{([3WJ]_0 + [S15]_0 + K_D) - \sqrt{([3WJ]_0 + [S15]_0 + K_D)^2 - 4[3WJ]_0[S15]_0}}{2}, \tag{4.43}$$

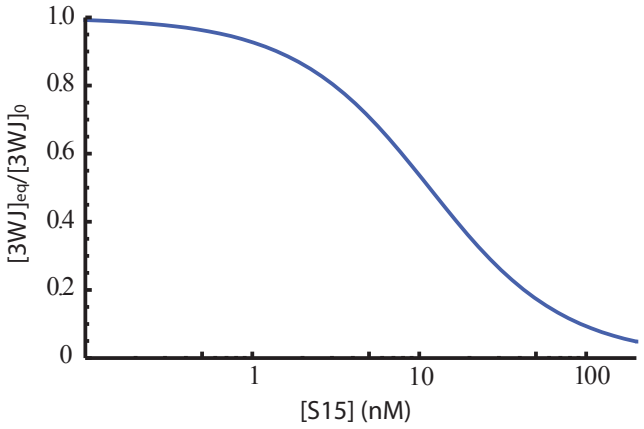


Figure 4.28: Fraction of 3WJ bound as a function of S15 concentration in equilibrium based on Eq. 4.44 for a $K_D = 10$ nM and $[3WJ]_0 = 3$ nM. As expected, at low [S15], essentially none of the 3WJ is bound. Therefore, $[3WJ]_{eq} = [3WJ]_0$ and the fraction $\frac{[3WJ]_{eq}}{[3WJ]_0} = 1$. As $[S15]$ increases, more of the $[3WJ * S15]_{eq}$ will have formed and, thus, $[3WJ]_{eq}$ will decrease. Thus, the ratio $\frac{[3WJ]_{eq}}{[3WJ]_0}$ will get smaller and smaller with increasing $[S15]$.

and, thus, the fraction of the 3WJ that has bound is (from Eq. 4.38)

$$\frac{[3WJ]_{eq}}{[3WJ]_0} = 1 - \frac{[3WJ * S15]_{eq}}{[3WJ]_0} = \frac{([3WJ]_0 - [S15]_0 + K_D) - \sqrt{([3WJ]_0 - [S15]_0 + K_D)^2 - 4[3WJ]_0[S15]_0}}{2[3WJ]_0}. \tag{4.44}$$

A plot of this equation is shown in Figure 4.28. The analysis we have presented in this case, where the dissociation constant can be expressed in terms of the unknown concentration x (see Eq. 4.40), which is found by solving a quadratic equation, is very commonly used in the cases of chemical equilibria [147]. In those cases, an equilibrium table is constructed in order to obtain changes in concentrations of each of the species involved in the reaction using the definition of an equilibrium constant (in our case, this constant was K_D). An example where this treatment is applicable is that of acid-base reactions [147]. The final Eq. 4.44 is essentially the fitting equation to our experimental data. By looking at the donor (atto425) fluorescence, we can experimentally obtain the fraction of the 3WJ that is bound, $\frac{[3WJ]_{eq}}{[3WJ]_0}$. In order to obtain K_D, we can fit the data to the above Eq. 4.44 using K_D as the fitting parameter. This is the fit equation that we used in Figures 4.14 and 4.21 in order to obtain K_D.

Thus, in this section we have shown how to obtain the reaction rates k_{on} and k_{off} as well as the dissociation constant K_D in the case of the S15 binding to the 3WJ. Even though we are interested in the reaction kinetics and, consequently, in the reaction rates k_{on} and k_{off}, obtaining K_D is our control experiment: we would like to ensure that the K_D value obtained using the microfluidics

setup with the FRET assay agrees with the reported literature value. This way, we can place more trust in the obtained values for the reaction rates.

4.10 Results

4.10.1 Mg^{2+} Titrations

As was described in detail in Section 4.2, a FRET assay has been established that reports on the conformational change that takes place at the 3-way junction (3WJ) of the central domain of the $16S$ rRNA, at the position where the $S15$ protein binds. When either $S15$ or Mg^{2+} is present, then folding of the junction is induced so that helices 22 and 21 stack coaxially and helices 20 and 22 form a 60° angle. If helices 20 and 22 are labeled, then in the unfolded state the interchromophore distance is long (8.5 nm) and there is low FRET. Addition of Mg^{2+} or $S15$ causes this distance to decrease to 5.5 nm and, therefore, there is high FRET.

We have used this well-characterized system to show the quantitative mixing in the microfluidics reaction ring by measuring the binding constant K_D of Mg^{2+} to the 3WJ using the dyes Atto425 (donor) and Alexa532 (acceptor). In this experiment 80 mM of $MgCl_2$ was delivered in a supply line to the mixing ring where it was diluted to the desired concentration. The fluorescently labeled RNA was then also delivered to the ring and was mixed with the Mg^{2+}. Figure 4.29 shows the outcome of this experiment. As was described in Section 4.9.1, in order to find the fraction of 3WJ that is folded, we plot FRET (acceptor/(donor+acceptor)) against $[Mg^{2+}]$. In order to obtain K_D we fit the experimental data to Eq. 4.13. The same experiment was repeated five times and each data point in the figure corresponds to an average of the corresponding 5 points from the data sets. The error bars were calculated from the standard deviation of each average. We see that the K_D value is in good agreement with the conventionally measured value of $K_D = 77\ \mu$M reported in the literature [80]. This was one of the first experiments performed with this set-up and it demonstrated that we can quantitatively deliver reagents in the microfluidics reaction ring and extract measurements of thermodynamics quantities.

4.10.2 S15 Kinetic Rates

In addition to the dependence of the folding of the 3WJ on $[Mg^{2+}]$, we also wanted to obtain the kinetic rates of the binding of the S15 on the same 3WJ. In this case a high concentration of S15 was delivered into the mixing ring. In the ring it was diluted to obtain several concentrations in order to examine how the binding depends on the concentration of S15. Just like with the Mg^{2+} titrations, the fluorescently labeled 3WJ rRNA was then also delivered to the ring and was mixed with the S15. The buffer used was 25 mM Tris pH = 7.6 0.25% w/v Tween-20, 50% Superblock-T20

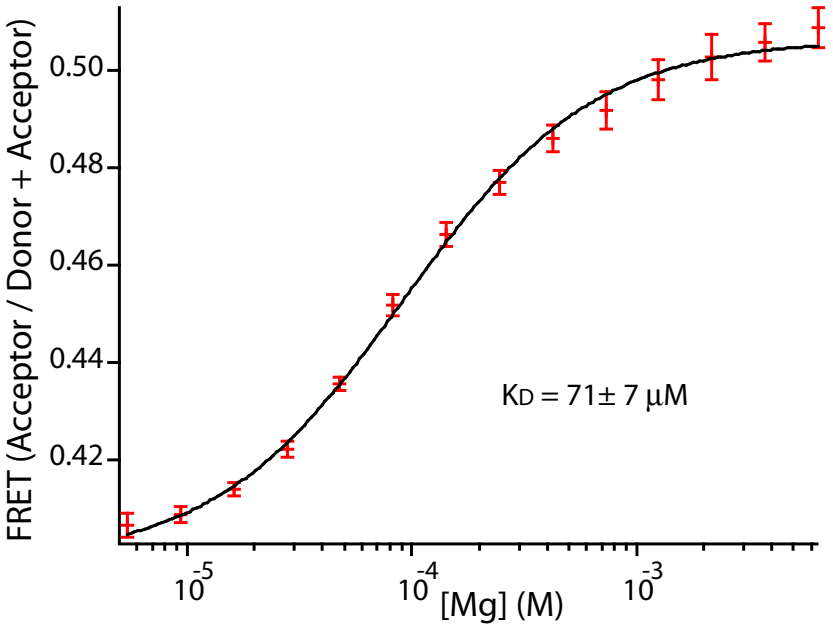


Figure 4.29: Affinity of the 3WJ for Mg^{2+} measured by folding induced FRET change. The plot shows how the folding of the 3WJ of the RNA is affected by the $[Mg^{2+}]$. When the $[Mg^{2+}]$ is low, then the ratio of acceptor/(donor+acceptor) is low, indicating that the signal from the acceptor is low and, thus, there is no FRET (i.e., the 3WJ is in the unfolded state). When the $[Mg^{2+}]$ is high, however, then the acceptor signal is high, indicating that FRET is taking place and, thus, the 3WJ is in the folded state. This experiment was repeated 5 times and each data point is the average of the corresponding 5 points from the data sets. The error bars were calculated from the standard deviation of each average. The experimentally measured value of K_D was obtained by fitting the data to equation 4.13 (solid line) and agrees with the conventionally measured value of $K_D = 77$ μM [80]. The fit also gave us that $n = 1$ (see Section 4.9.1).

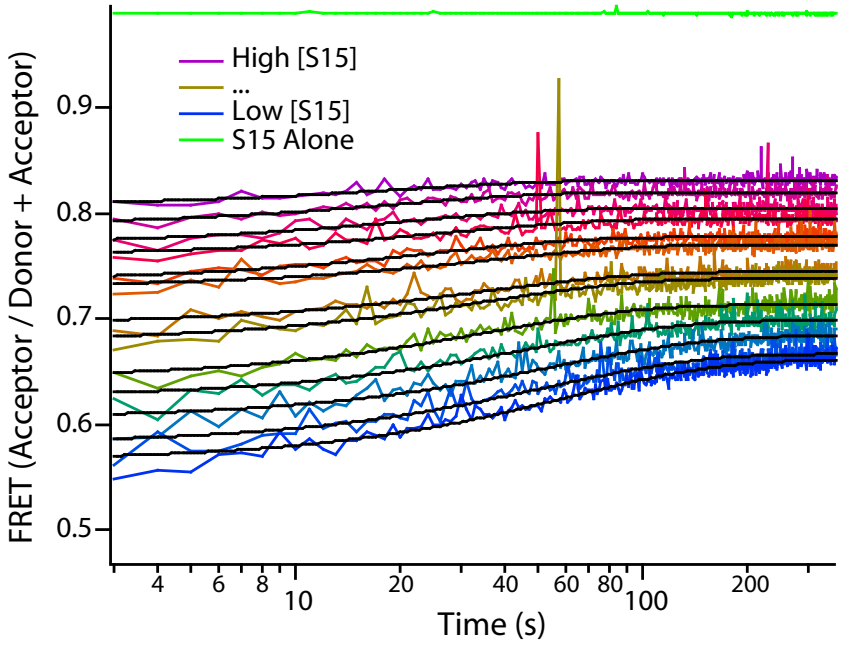


Figure 4.30: FRET induced due to S15 binding as a function of time and [S15]. The graph shows how the ratio of acceptor/(donor+acceptor) changes over time for different S15 concentrations. The buffer used was 25 mM Tris pH = 7.6 0.25% w/v Tween-20, 50% Superblock-T20 (see Table 4.1). The initial concentrations of 3WJ and S15 were 50 nM and 200 nM, respectively. When [S15] is low, the ratio of acceptor/(donor+acceptor) changes slowly over time and, therefore, it takes a long time for the binding to occur. When [S15] is high, then the binding of the protein to the 3WJ takes place faster, since the curve is almost flat, indicating that the FRET has occurred quickly. At the top of the graph we have the control experiment, where the 3WJ is absent. In that case, since there is no donor present, the ratio of acceptor/(donor+acceptor) is equal to 1.

(see Table 4.1). The initial concentrations of 3WJ and S15 were 50 nM and 200 nM, respectively. Figure 4.30 shows the outcome of this experiment. In this case we show how the FRET signal of acceptor/(donor+acceptor) on the y axis changes with time on the x axis as well as $[S15]$ (different colored curves). We see that when [S15] is low, the ratio of acceptor/(donor+acceptor) changes slowly over time and, therefore, it takes a long time for the binding to occur. When [S15] is high, then the binding of the protein to the 3WJ takes place faster, since the curve is almost flat, indicating that the FRET has occurred quickly. At the top of the graph we have the control experiment, where the 3WJ is absent. In that case, since there is no donor present, the ratio of acceptor/(donor+acceptor) is equal to 1.

As was described in Section 4.9.2, we can obtain the $[3WJ{*}S15]$ as a function of time by looking at the FRET signal of (acceptor/(donor+acceptor)) over time. Then, we can fit to that signal Eq. 4.30 using k_{app} as a fitting parameter in order to obtain the apparent reaction rate for each concentration of S15. Based on Eq. 4.26, we can then find k_{on} and k_{off} from the plot of k_{app} vs. $[S15]$. Performing

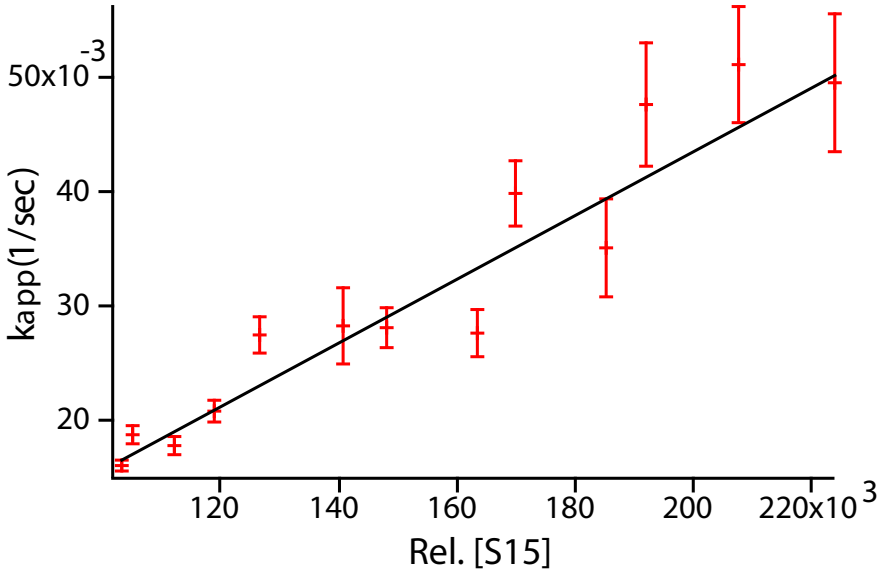


Figure 4.31: Apparent reaction rate constants (k_{app}) as a function of [S15]. We fit Eq. 4.30 to each of the curves shown in Figure 4.30. Each of the obtained k_{app} rates is then plotted as a function of [S15]. As expected, the plot is linear. From the slope and the intercept we can obtain k_{on} and k_{off}, respectively, as was explained in Section 4.9.2.

this fitting of Eq. 4.30 for each of the curves of Figure 4.30, we obtained Figure 4.31, which shows how k_{app} depends on the signal coming from S15 (photon counts), which is directly proportional to the concentration of S15. We see that, as expected based on Eq. 4.26, the plot is linear. However, getting the x axis to be in terms of actual S15 concentration instead of photon counts has proven to be a task. This is because, based on the discussion given in Section 4.6.3.3, the S15 protein interacts strongly with the walls and in a way that changes with time and protein concentration. As a result, even though we see a linear dependence of k_{app} vs. the amount of S15 that is reacting, we cannot infer from the S15 signal what the actual S15 concentration is. In Section 4.6.3.3, we described in detail the methods we used in order to quantify the amount of protein that is sticking or resolve it via surface chemistry techniques. We summarize all these findings in the following section.

4.11 Conclusions

In this chapter we have shown that in order to study the complete self-assembly of the 30S ribosomal subunit, it is important to use automation in the form of a microfluidic device. Studying the assembly of the 30S subunit involves 21 components and, therefore, this high-throughput, automated strategy of using microfluidics will enable the observation of several reactions in a very small amount of time.

We designed and fabricated a chip that can combine reagents at different concentrations and mix them so that a reaction can occur. When combined with an assay that uses Fluorescence Resonance Energy Transfer (FRET), it is possible to observe the reaction in real time and obtain reaction rates and dissociation constants in order to form the statistically weighted assembly map. That is, since there are several parallel pathways in the assembly process, the goal is to obtain the probability of each pathway (i.e., the probability of each individual trajectory in the assembly process).

In order to test the chip and the possibility of obtaining reaction rates, we have tested our setup by studying the Mg^{2+}-dependent folding and subsequent binding by primary protein S15 of a part of the central domain of the rRNA very commonly known as the 3-way junction (3WJ). This system has been studied before and the dissociation constants K_Ds are known. Using this system, we have conducted experiments to show that we can obtain K_D values (that match those reported in the literature) and reaction rates inside the chip by using FRET. The 3WJ is labeled with the dyes Atto425 and Atto532 while S15 is labeled with Alexa594. When the 3WJ is in the unfolded conformation, we expect to see mainly signal coming from dye Atto425. When it is folded, however, due to FRET, the signal coming from dye Atto532 should increase. Similarly, due to FRET, when S15 is bound, then the signal coming from Alexa594 should be dominant.

In order to see how the folding of the 3WJ depends on $[Mg^{2+}]$ we performed Mg^{2+} titrations and obtained the dissociation constant K_D. Similarly, in order to see the dependence of the S15 binding on the 3WJ as a function of $[S15]$ we performed S15 titrations. In this case our goal was to obtain not only the dissociation constant K_D for this reaction but also the reaction rates k_{on} and k_{off}. Based on our experiments, we can draw the following conclusions:

- The chip can indeed provide us with an automated way for obtaining reaction rates and dissociation constants. The chip has been able to reproducibly dilute reagents (see Figure 4.17) in a manner that can be quantified. Therefore, microfluidic devices can be used to automate processes that have been traditionally done with the use of fluorimeters and plate readers, where the user had to perform all the dilutions necessary for the titrations. We believe that, in this respect, the experiments that we have conducted can be easily scaled up to allow for multiple reagents. Thus, it is possible to study the complete assembly of the 30S subunit in this automated manner. However, as will be explained below, there are some technical hurdles that stand in the way when proteins are present inside the microfluidic devices.

- Monitoring FRET inside microfluidic devices is possible. We have found that the folding of the 3WJ due to the presence of Mg^{2+} could be easily detected and quantified in a way that provided us with the dissociation constant K_D (see Figure 4.29). This value of $K_D = 71 \pm 7$ μM agrees with the reported literature values of 77 μM [80]. There is also some evidence (preliminary results, not shown) that we can use fluorescence correlation spectroscopy (FCS)

in order to study reactions at the single-molecule level.

- We have found that, based on fluorescence assays, we can observe reactions in real time (see Figures 4.30 and 4.31) and obtain reaction rates. However, this last conclusion is hindered by the next point.

- Protein sticking on the microfluidic walls is hard to overcome. When observing reactions in real time in order to deduce the reaction rates, it is important to know the initial concentration of the reactants. However, we found that sticking of proteins on the microfluidic walls is a very dynamic process and, therefore, difficult to model and account for or eliminate. The incorporation of blocking agents and detergents in the buffer can greatly decrease protein adsorption to the PDMS walls (see Tables 4.1 and 4.2). However, we have found that this incorporation leads to two regimes: when we manage to decrease protein adsorption and obtain a high S15 signal, we normally observe minimal or no FRET, indicating that there is little or no reaction taking place. On the other hand, when FRET (i.e., reaction) is observed, we are in the regime where a lot of protein is lost to the walls.

Therefore, it seems that the future of the study of ribosomal self-assembly (or other reactions involving proteins) in microfluidic devices is strongly dependent on overcoming the challenge of protein adsorption on the microfluidic walls. More specifically, as was described above, since there are two regimes, one where reaction is observed but a lot of protein is adsorbed and another where there is little reaction but a lot of free protein, the challenge lies on being able to find some middle ground. One way of attaining this middle ground is by optimizing the combination and concentrations of the used detergents and blocking agents. In Tables 4.1 and 4.2 we have presented the detergents and blocking agents we used in order to resolve the problem of protein adsorption, each of which, we found, has its own advantages and disadvantages. It seems possible, however, that a combination of these detergents/blocking agents at optimized concentrations will be able to result in the desired outcome. In addition, we could also add to the buffer molecules that will act as "protein carriers." These molecules will interact weakly with the protein. Thus, when they are present in the protein buffer, the strength of the molecule-protein interaction is such that it will prevent the protein from sticking to the chip walls, but will not stop the protein from interacting with the actual rRNA, since this protein-rRNA interaction is much stronger. In essence, these molecules will be able to "carry" the protein into the chip and keep it away from the walls until the desirable reaction can occur. One example of such a molecule will be a different type of RNA (like a tRNA), a double- or single-stranded DNA, etc., that interacts weakly with the protein. We plan on examining both of these methods (i.e., optimization of blocking agents/detergents and incorporation of "protein carriers") in the near future.

Bibliography

[1] H. B. Callen. *Thermodynamics and an introduction to thermostatistics.* Wiley, second edition, 1985.

[2] R. A. Alberty and R. J. Silbey. *Physical Chemistry.* Wiley, second edition, 1996.

[3] K. Dill and S. Bromberg. *Molecular driving forces: statistical thermodynamics in chemistry and biology.* Garland Science, New York, 2003.

[4] R. Bowley and M. Sanchez. *Introductory Statistical Mechanics.* Oxford Science Publications, second edition, 1999.

[5] E. T. Jaynes. Information theory and statistical mechanics. In R. D. Rosenkrantz, editor, *E. T. Jaynes: Papers on probability, statistics and statistical physics.* Chapter 1. Kluver Academic Publishers, 1957.

[6] L. Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, 1931.

[7] E. T. Jaynes. The minimum entropy production principle. In R. D. Rosenkrantz, editor, *E. T. Jaynes: Papers on probability, statistics and statistical physics.* Chapter 14. Kluver Academic Publishers, 1980.

[8] E. T. Jaynes. Macroscopic predictions. In H. Haken, editor, *Complex systems–operational approaches.* p. 254. Springer-Verlag, 1985.

[9] M. M. Inamdar, W. M. Gelbart, and R. Phillips. Dynamics of DNA ejection in bacteriophage. *Biophys. J.*, 91, 2006.

[10] M. LeBellac, F. Mortessagne, and G. G. Batrouni. *Equilibrium and Non-equilibrium thermodynamics.* University Press, Cambridge, UK, 2004.

[11] M. Doi. *Introduction to polymer physics.* Clarendon Press, February 1996.

[12] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78(14):2690–2691, 1997.

[13] C. Jarzynski. Equilibrium free energy difference from nonequilibrium: A master equation approach. *Phys. Rev. E*, 56:5018–5035, 1997.

[14] C. Jarzynski. Nonequilibrium work theorem for a system strongly coupled to a thermal environment. *J. Stat. Mech.: Theor. Exp.*, 2004.

[15] J. Liphardt, S. Dumont, S. B. Smith, I. Tinocho, and C. Bustamante. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*, 296:1832–1835, 2002.

[16] G. Crooks. *Excursions in statistical dynamics*. PhD thesis, University of California, Berkeley, 1999.

[17] Denis J. Evans and Debra J. Searles. The fluctuation theorem. *Adv. Phys.*, 51(7):1529–1585, 2002.

[18] E. T. Jaynes. Where do we stand on maximum entropy? In R. D. Rosenkrantz, editor, *E. T. Jaynes: Papers on probability, statistics and statistical physics*. Chapter 10. Kluver Academic Publishers, 1978.

[19] B. Robertson. Equations of motion in non-equilibrium statistical mechanics. *Phys. Rev.*, 144:151–161, 1966.

[20] B. Robertson. Projection operators using maximum entropy methods. In R. D. Levine and M. Tribus, editors, *The maximum entropy formalism*. MIT Press, 1978.

[21] W. T. Grandy. Principle of maximum entropy and irreversible processes. *Phys. Rep.*, 62:176–266, 1980.

[22] R. Dewar. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *J. Phys. A*, 36:631–641, 2003.

[23] R. C. Dewar. Maximum entropy production and the fluctuation theorem. *J. Phys. A*, 38:L371–L381, 2005.

[24] R. M. L. Evans. Detailed balance has a counterpart in non-equilibrium steady states. *J. Phys. A*, 38:293–313, 2005.

[25] D. Wu, K. Ghosh, M. M. Inamdar, H. J. Lee, S. Fraser, and K. Dill. A trajectory approach to two-state kinetics of single particles on sculpted energy landscapes. *Phys. Rev. Lett., (submitted)*, 2008.

[26] B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Essential Cell Biology: An Introduction to the Molecular Biology of the Cell.* Garland Publishing, New York, July 1997.

[27] H. Lodish, A. Berk, S. L. Zipurski, P. Matsudaira, D. Baltimore, and J. Darnell. *Molecular Cell Biology.* W. H. Freeman and Company, 2000.

[28] D. Voet and J. G. Voet. *Biochemistry.* Wiley, 2004.

[29] J. M. Berg, J. L. Tymoczko, and L. Stryer. *Biochemistry 2001.* W. H. Freeman and Company, fifth edition, 2001.

[30] J. Poehlsgaard and S. Douthwaite. The bacterial ribosome as a target for antibiotics. *Nature Reviews Microbiology*, 3:870–881, 2005.

[31] University of Texas Medical Branch. `http://cellbio.utmb.edu/`.

[32] National Institute of Health (NIH). Inside the cell, Chapter 2. `http://publications.nigms.nih.gov/`.

[33] G. E. Palade. A small particulate component of the cytoplasm. *J Biophys Biochem Cytol*, 1:59–68, 1955.

[34] H. E. Huxley and G. Zubay. Electron microscope observations on the structure of microsomal particles from *Escherichia coli. J. Mol. Biol.*, 2:10–18, 1960.

[35] R. Bender, S. Hellman, and R. Gordon. Art and the ribosome: a preliminary report on the three-dimensional structure of individual ribosomes determined by an algebraic reconstruction technique. *J. Theor. Biol.*, 29:483–487, 1970.

[36] R. A. Crowther, D. J. DeRosier, and A. Klug. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proc. Royal. Soc. London*, 317:319–340, 1970.

[37] R. A. Crowther. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. *Philos. Trans. R. Soc. Lond. B*, 261:221–230, 1971.

[38] A. Klug and R. A. Crowther. Three-dimensional image reconstruction from the viewpoint of information theory. *Nature*, 238:435–440, 1972.

[39] A. Klug. Image analysis and reconstruction in the electron microscopy of biological macromolecules. *Chemica Scripta*, 14:245–256, 1979.

[40] E. Kaltschmidt and H. G. Wittmann. Ribosomal proteins, xii. number of proteins in small and large ribosomal subunits of *Escherichia coli* as determined by two-dimensional gel electrophoresis. *Proc. Natl. Acad. Sci. USA*, 67:1276–1282, 1970.

[41] G. W. Tischendorf, H. Zeichhardt, and G. Stoffler. Location of proteins S5, S13 and S14 on the surface of the 30S ribosomal subunit from *Escherichia coli* as determined by immune electron microscopy. *Mol. Gen. Genet.*, 134:209223, 1974.

[42] G. W. Tischendorf, H. Zeichhardt, and G. Stoffler. Determination of the location of proteins L14, L17, L18, L19, L22, L23 on the surface of the 50S ribosomal subunit of *Escherichia coli* by immune electron microscopy. *Mol. Gen. Genet.*, 134:187–208, 1974.

[43] G. W. Tischendorf, H. Zeichhardt, and G. Stoffler. Architecture of the *Escherichia coli* ribosome as determined by immune electron microscopy. *Proc. Natl. Acad. Sci. USA*, 72:4820–4824, 1975.

[44] G. W. Tischendorf and G. Stoffler. Localization of *Escherichia coli* ribosomal protein S4 on the surface of the 30S ribosomal subunit by immuno electron microscopy. i. distribution of antibody-binding sites as obtained with immunoglobulins and monovalent antibody fragments from various s4-specific antisera. *Mol. Gen. Genet.*, 142:193–208, 1975.

[45] J. A. Lake and L. Kahan. Ribosomal proteins S5, S11, S13 and S19 localized by electron microscopy of antibody-labeled subunits. *J. Mol. Biol.*, 99:631–644, 1975.

[46] J. A. Lake, W. A. Strycharz, and M. Nomura. Ribosomal proteins L7/L12 localized at a single region of the large subunit by immune electron microscopy. *J. Mol. Biol.*, 126:123–140, 1978.

[47] B. Kastner, M. Stöffler-Meilicke, and G. Stoffler. Arrangement of the subunits in the ribosome of *Escherichia coli*: demonstration by immunoelectron microscopy. *Proc. Natl. Acad. Sci. USA*, 78:6652–6656, 1981.

[48] W. A. Held, B. Ballou, S. Mizushima, and M. Nomura. Assembly mapping of 30S ribosomal proteins from *Escherichia coli*. *J. Biol. Chem.*, 249(10):3103–3111, 1974.

[49] D. M. Engelman, P. B. Moore, and B. P. Schoenborn. Neutron scattering measurements of separation and shape of proteins in 30S ribosomal subunit of *Escherichia coli*: S2-S5, S5-S8, S3-S7. *Proc. Natl. Acad. Sci. USA*, 72:38883892, 1975.

[50] J. A. Langer, D. M. Engelman, and P. B. Moore. Neutron-scattering studies of the ribosome of *Escherichia coli*: a provisional map of the locations of proteins S3, S4, S5, S7, S8 and S9 in the 30S subunit. *J. Mol. Biol.*, 119:463–485, 1978.

[51] P. B. Moore, D. M. Engelman, J. A. Langer, V. R. Ramakrishnan, D. G. Schindler, B. P. Schoenborn, I. Y. Sillers, and S. Yabuki. Neutron scattering and the 30S ribosomal subunit of *E. coli*. *Basic Life Sci.*, 27:73–91, 1984.

[52] J. Frank, P. Penczek, R. Grassucci, and S. Srivastava. Three-dimensional reconstruction of the 70S *Escherichia coli* ribosome in ice: The distribution of ribosomal RNA. *J. Cell Biol.*, 115:597–605, 1991.

[53] J. Frank, J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, R. K. Lata, and R. K. Agrawal. A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome. *Nature*, 376:441–444, 1995.

[54] J. Frank, A. Verschoor, Y. Li, J. Zhu, R. K. Lata, M. Radermacher, P. Penczek, R. Grassucci, R. K. Agrawal, and S. Srivastava. A model of the translational apparatus based on a three-dimensional reconstruction of the *Escherichia coli* ribosome. 73:757–765, 1995.

[55] H. Stark, F. Mueller, E. V. Orlova, M. Schatz, P. Dube, T. Erdemir, F. Zemlin, R. Brimacombe, and M. van Heel. The 70S *Escherichia coli* ribosome at 23 Å resolution: fitting the ribosomal RNA. *Structure*, 3:815–821, 1995.

[56] J. Frank. *Three-dimensional Electron Microscopy of Macromolecular Assemblies*. San Diego: Academic Press, 1996.

[57] J. Frank. The ribosome at higher resolution – the donut takes shape. *Curr. Op. Struct. Biol.*, 7:266–272, 1997.

[58] J. Frank and R. K. Agrawal. The movement of tRNA through the ribosome. *Biophys. J.*, 74:589–594, 1998.

[59] A. Malhotra, P. Penczek, R.K. Agrawal, I.S. Gabashvili, R.A. Grassucci, R. Jnemann, N. Burkhardt, K.H. Nierhaus, and J. Frank. *Escherichia coli* 70S ribosome at 15 Å resolution by cryo-electron microscopy: localization of fMet-tRNAfMet and fitting of L1 protein. *J. Mol. Biol.*, 280:103–115, 1998.

[60] P. Dube, M. Wieske, H. Stark, M. Schatz, J. Stahl, F. Zemlin, G. Lutsch, and M. van Heel. The rat liver ribosome at 25 Å resolution by electron cryomicroscopy and angular reconstitution. *Structure*, 6:389–399, 1998.

[61] N. Ban, B. Freeborn, P. Nissen, P. Penczek, R. A. Grassucci, R. Sweet, J. Frank, P. B. Moore, and T. A. Steitz. A 9-Å resolution x-ray crystallographic map of the large ribosomal subunit. *Cell*, 93:1105–1115, 1998.

[62] N. Ban, P. Nissen, J. Hansen, P. B. Moore, and T. A. Steitz. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, 289:905–920, 2000.

[63] B. T. Wimberly, D. E. Brodersen, W. M. Clemons Jr., R. J. Morgan-Warren, A. P. Carter, C. Vonrhein, T. Hartsch, and V. Ramakrishnan. Structure of the 30S ribosomal subunit. *Nature*, 407:327–339, 2000.

[64] F. Schluenzen, A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, and A. Yonath. Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell*, 102:615–623, 2000.

[65] M. M. Yusupov, G. Zh. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. D. Cate, and H. F. Noller. Crystal structure of the ribosome at 5.5 Å resolution. *Science*, 292:883–896, 2001.

[66] J. R. Williamson. After the ribosome structures: How are the subunits assembled? *RNA*, 9:165–167, 2003.

[67] R. Garrett. Mechanics of the ribosome. *Nature*, 400:811–812, 1999.

[68] J. A. Lake, E. Henderson, M. W. Clark, and A. T. Matheson. Mapping evolution with ribosome structure: intralineage constancy and interlineage variation. *Proc. Natl. Acad. Sci. USA*, 79:5948–5952, 1982.

[69] C. R. Woese and G. E. Fox. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. USA*, 74:5088–5090, 1977.

[70] A. S. Mankin. Ribosomal antibiotics. *Mol. Biol.*, 35:509–520, 2001.

[71] D. Vazquez. Inhibitors of protein biosynthesis. *Mol. Biol. Biochem. Biophys.*, 30:1–312, 1979.

[72] E. F. Gale, E. Cundliffe, P. E. Reynolds, M. H. Richmond, and M. J. Waring. *The molecular basis of antiobiotic action*. John Wiley and Sons, London, 1981.

[73] C. M. Spahn and C. D. Prescott. Throwing a spanner in the works: antibiotics and the translation apparatus. *J. Mol. Med.*, 74:423–439, 1996.

[74] K. Hamacher, J. Trylska, and J. A. McCammon. Dependency map of proteins in the small ribosomal subunit. *PLoS Computational Biology*, 2(2), 2006.

[75] D. K. Treiber and J. R. Williamson. Beyond kinetic traps in RNA folding. *Curr. Op. Struct. Biol.*, 11:309–314, 2001.

[76] K. Ghosh, K. Dill, M. Inamdar, E. Seitaridou, and R. Phillips. Teaching the principles of statistical dynamics. *Am. J. Phys.*, 74(2):123–33, 2006.

[77] E. Seitaridou, M. Inamdar, R. Phillips, K. Ghosh, and K. Dill. Measuring flux distributions for diffusion in the small-numbers limit. *J. Phys. Chem. B*, 111(9):2288–2292, 2007.

[78] T. Ha, X. Zhuang, H. D. Kim, J. W. Orr, J. R. Williamson, and S. Chu. Ligand-induced conformational changes observed in single RNA molecules. *Proc. Natl. Acad. Sci. USA*, 96:9077–9082, 1999.

[79] H. D. Kim, G. U. Nienhaus, T. Ha, J. W. Orr, J. R. Williamson, and S. Chu. Mg^{2+}-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules. *Proc. Natl. Acad. Sci. USA*, 99(7):4284–4289, 2002.

[80] D. Klostermeier, P. Sears, C.-H. Wong, D. P. Millar, and J. R. Williamson. A three-fluorophone FRET assay for high-throughput screening of small-molecule inhibitors of ribosome assembly. *Nucl. Acids Res.*, 32(9):2707–2715, 2004.

[81] W. K. Ridgeway, E. Seitaridou, R. Phillips, and J. R. Williamson. Using a microfluidics reactor for assembling RNA - protein complexes, *(in preparation)*.

[82] A. E. Kamholz, E. A. Schilling, and P. Yager. Optical measurement of transverse molecular diffusion in a microchannel. *Biophys. J.*, 80:1967–1972, 2001.

[83] K. Macounova, C. R. Cabrera, M. R. Holl, and P. Yager. Generation of natural ph gradients in microfluidic channels for use in isoelectric focusing. *Anal. Chem.*, 72:3745–3751, 2000.

[84] A. G. Hadd, D. E. Raymond, J. W. Halliwell, S. C. Jacobson, and J. M. Ramsey. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal. Chem.*, 69:3407–3412, 1997.

[85] D. C. Duffy, J. C. McDonald, O. J. A. Schueller, and G. M. Whitesides. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal. Chem.*, 70:4974–4984, 1998.

[86] A. Hatch, A. E. Kamholz, K. R. Hawkins, M. S. Munson, E. A. Schilling, B. H. Weigl, and P. Yager. A rapid diffusion immunoassay in a T-sensor. *Nature Biotech.*, 19:461–465, 2001.

[87] W. K. Wu, C. K. Liang, and J. Z. Huang. MEMS-based flow cytometry: Microfluidics-based cell identification system by fluorescent imaging. *Conf Proc IEEE Eng Med Biol Soc*, 4:2579–2581, 2004.

[88] M. U. Kopp, A. J. de Mello, and A. Manz. Chemical amplification: Continuous-flow PCR on a chip. *Science*, 280:1046–1048, 1998.

[89] C. L. Hansen, E. Skordalakes, J. M. Berger, and S. R. Quake. A robust and scalable microfluidic metering method that allows protein crystal growth by free interface diffusion. *Proc. Natl. Acad. Sci. USA*.

[90] S. Takayama, E. Ostuni, P. LeDuc, K. Naruse, D. E. Ingber, and G. M. Whitesides. Selective chemical treatment of cellular microdomains using multiple laminar streams. *Chem. Biol.*, 10:123–130, 2003.

[91] T. M. Squires and S. R. Quake. Microfluidics: Fluid physics at the nanoliter scale. *Rev. Mod. Phys.*, 77:977–1026, 2005.

[92] J.-Q. Wu and T. D. Pollard. Counting cytokinesis proteins globally and locally in fission yeast. *Science*, 310(5746):310–314, 2005.

[93] J. Kasianowicz, E. Brandin, D. Branton, and D. Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Nat. Acad. Sci. USA*, 93:13770–13773, 1996.

[94] H. P. Lu, L. Xun, and X. S. Xie. Single molecule enzymatic dynamics. *Science*, 282:1887, 1998.

[95] M. Reif, R. S. Rock, Amit D. Mehta, M. S. Mooseker, R. E. Cheney, and J. A. Spudich. Myosin-V stepping kinetics: A molecular model for processivity. *Proc. Nat. Acad. Sci.*, 97:9482–9486, 2000.

[96] A. Meller, L. Nivon, and D. Branton. Voltage-driven DNA translocations through a nano-pore. *Phys. Rev. Lett.*, 86(15):3435–3438, 2001.

[97] D. E. Smith, S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson, and C. Bustamante. The bacteriophage $\phi 29$ portal motor can package DNA against a large internal force. *Nature*, 413:748, 2001.

[98] H. Li, W. A. Linke, A. F. Oberhauser, M. Carrion-Vazquez, J. G. Kerkvliet, H. Lu, P. E. Marszalek, and J. M. Fernandez. Reverse engineering of the giant muscle protein. *Nature*, 418:998–1002, 2002.

[99] C. Bustamante, Z. Bryant, and S. B. Smith. Ten years of tension: single-molecule DNA mechanics. *Nature*, 421:423–427, 2003.

[100] B. Hille. *Ion channels of excitable membranes.* Sinauer Associates, third edition, 2001.

[101] J. Howard. *Mechanics of motor proteins and the cytoskeleton.* Sinauer Associates Inc., 2001.

[102] K. C. Neuman and S. M. Block. Optical trapping. *Rev. Sci. Inst.*, 75:2787–2809, 2004.

[103] J. C. Crocker and D. G. Grier. Methods of digital video microscopy for colloidal studies. *J. Coll. Int. Sci.*, 179:298–310, 1996.

[104] E. R. Dufresne, D. Altman, and D. G. Grier. Brownian dynamics of sphere in a slit pore. *Europhys. Lett.*, 53:264–270, 2001.

[105] D. Kondepudi and I. Prigogine. *Modern thermodynamics: From heat engines to dissipative structures*. John Wiley and Sons, 1998.

[106] R. P. Feynman. Nobel lecture in physics, 1965. In *Nobel Lectures in Physics: 1901–1995*. World Scientific Publishing Co., 1998.

[107] M. Klein. Entropy and the Ehrenfest urn model. *Physica*, 22:569–575, 1956.

[108] Gerard G. Emch and Chuang Liu. *The logic of thermostatical physics*. Springer-Verlag, NY, 2002.

[109] R. W. Baker. *Controlled release of biologically active agents*. Wiley, New York, 1987.

[110] D. J. Evans, E. G. D. Cohen, and G. P. Morriss. Probability of 2nd law violations in shearing steady-states. *Phys. Rev. Lett.*, 71(15):2401, 1993.

[111] R. Feynman. *Feynman lectures on computation*. Perseus, 1996.

[112] M. Kac. Paper. *Amer. Math. Monthly*, 54:369, 1947.

[113] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans. Experimental verification of the fluctuation theorem. *Phys. Rev. Lett.*, 89:050601, 2002.

[114] F. Reif. *Fundamentals of statistical and thermal physics*. McGraw-Hill, NY, 1965.

[115] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E.*, 60:2721, 1999.

[116] V. Blickle, T. Speck, L. Helden, U. Seifert, and C. Bechinger. Thermodynamics of colloidal particle in a time-dependent nonharmonic potential. *Phys. Rev. Letts.*, 96:070603, 2006.

[117] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante. Verification of the crooks fluctuation theorem and recovery of RNA folding free energies. *Nature*, 437:231, 2005.

[118] K. Feitosa and N. Menon. Fluidized granular medium as an instance of fluctuation theorem. *Phys. Rev. Lett.*, 94(16):164301–1–4, 2004.

[119] M. A. Unger, H. P. Chou, T. Thorsen, A. Scherer, and S. R. Quake. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science*, 288:113–116, April 2000.

[120] L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*. Butterworth-Heinemann, New York, second edition, 1987.

[121] J.-C. Meiners and S. R. Quake. Direct measurement of hydrodynamic cross correlations between two particles in an external potential. *Phys. Rev. Letts.*, 82(10):2211–2214, 1999.

[122] J. C. Crocker and D. G. Grier. Microscopic measurement of the pair interaction potential of charge-stabilized colloid. *Phys. Rev. Letts.*, 73(2):352–355, 1994.

[123] E. R. Dufresne, T. M. Squires, M. P. Brenner, and D. G. Grier. Hydrodynamic coupling of two Brownian spheres to a planar surface. *Phys. Rev. Lett.*, 85(15):3317–3320, 2000.

[124] P. Nelson. *Biological Physics: Energy, Information, Life.* W. H. Freeman & Co., 2003.

[125] M. W. T. Talkington, G. Siuzdak, and J. R. Williamson. An assembly landscape for the 30S ribosomal subunit. *Nature*, 438:628–632, 2005.

[126] J. A. Davies. *Mechanisms of morphogenesis: the creation of biological form.* Elsevier Academic Press, 2005.

[127] L. F. Lindoy and I. M. Atkinson. *Self-assembly in supramolecular systems.* Royal Society of Chemistry, 2000.

[128] W. B. Wood and R. S. Edgar. Building a bacterial virus. *Sci. Am.*, 217:61–66, 1967.

[129] W. B. Wood, R. S. Edgar, J. King, I. Lielausis, and M. Henninger. Bacteriophage assembly. *Fed. Proc.*, 27:1160–1166, 1968.

[130] S. C. Agalarov, G. S. Prasad, P. M. Funke, C. D. Stout, and J. R. Williamson. Structure of the S15, S6, S18-rRNA complex: Assembly of the 30S ribosome central domain. *Science*, 288:107–112, 2000.

[131] P. Traub and M. Nomura. Structure and function of *Escherichia coli* ribosomes VI. mechanism of assembly of 30S ribosomes studied *in vitro*. *J. Mol. Biol.*, 40:391–413, 1969.

[132] R. T. Batey and J. R. Williamson. Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16S rRNA: I. defining the minimal RNA site. *J. Mol. Biol.*, 261(4):536–549, 1996.

[133] J. W. Orr, P. J. Hagerman, and J. R. Williamson. Protein and Mg^{2+}-induced conformational changes in the S15 binding site of 16S ribosomal RNA. *J. Mol. Biol.*, 275(3):453–464, 1998.

[134] R. T. Batey and J. R. Williamson. Effects of polyvalent cations on the folding of an rRNA three-way junction and binding of ribosomal protein S15. *RNA*, 4:984–997, 1998.

[135] M. I. Recht and J. R. Williamson. Central domain assembly: Thermodynamics and kinetics of S6 and S18 binding to an S15-RNA complex. *J. Mol. Biol.*, 313:35–48, 2001.

[136] M. I. Recht and J. R. Williamson. RNA tertiary structure and cooperative assembly of a large ribonucleoprotein complex. *J. Mol. Biol.*, 344:395–407, 2004.

[137] J. J. Cannone, S. Subramanian, M. N. Schnare, J. R. Collett, L. M. D'Souza, Y. Du, B. Feng, N. Lin, L. V. Madabusi, K. M. Mller, N. Pande, Z. Shang, N. Yu, and R. R. Gutell. The comparative RNA web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BioMed Central Bioinformatics*, 3:327–339, 2002.

[138] P. Schwille, U. Haupts, S. Maiti, and W. W. Webb. Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophys. J.*, 77:2251–2265, 1999.

[139] R. T. Batey and J. R. Williamson. Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16S rRNA: II. specificity determinants of RNA-protein recognition. *J. Mol. Biol.*, 261(4):550–567, 1996.

[140] C. L. Hansen, M. O. A. Sommer, and S. R. Quake. Systematic investigation of protein phase behavior with a microfluidic formulator. *Proc. Natl. Acad. Sci. USA*.

[141] B. Huang, H. Wu, S. Kim, and R. N. Zare. Coating of poly(dimethylsiloxane) with n-dodecyl-β-d-maltoside to minimize nonspecific protein adsorption. *Lab on a Chip*, 5:1005–1007, 2005.

[142] C. Essö. *Modifying polydimethylsiloxane (PDMS) surfaces.* Master's thesis, Mälardalen University, 2007.

[143] K. Boxshall, M.-H. Wu, Z. Cui, Z. Cui, J. F. Watts, and M. A. Baker. Simple surface treatments to modify protein adsorption and cell attachment properties within a poly(dimethylsiloxane) micro-bioreactor. *Surf. Interface Anal.*, 38:198–201, 2006.

[144] G. Sui, J. Wang, C.-C. Lee, W. Lu, S. P. Lee, J. V. Leyton, A. M. Wu, and H.-R. Tseng. Solution-phase surface modification in intact poly(dimethylsiloxane) microfluidic channels. *Anal Chem*, 78:5543–5551, 2006.

[145] M. Goeppert-Mayer. Über elementarakte mit zwei quantensprüngen. *Ann. Physik*, 9:273–294, 1931.

[146] Stowers Institute for Medical Research. *Technology: Two-Photon Microscopy, Non-Linear Optics (NLO).* `http://research.stowers-institute.org/`.

[147] P. W. Atkins. *The elements of physical chemistry.* Oxford University Press, 1995.