

# PHYS180 Projects 4 and 5

Project 4: Dynamics of Small Objects with Random Forces and External Potentials: Nr.3-Diffusion in Double Well Potential, Nr.4-Modeling Bacterial Chemotaxis.

Project 5: Dynamics of Complex Biological Situations: Nr.2-Force on Ends of DNA Chain

Owen Morehead, Karoli Clever, Rey Cervantes

May 15, 2021

## 1 Problem 3: Diffusion in Double Well Potential

### Equilibrium

In this problem, we explore diffusion in a “double well” potential, with particle motion occurring in a potential with more than one equilibrium point. When working on updating the code for nonlinear\_hints.py, we must revisit the solution to the Langevin equation with inertia being neglected (mass = 0) provided in the Brownian motion in an optical trap problem, Equation 1:

$$\gamma * dx/dt = -kx + n(t) \quad (1)$$

However, instead of working with the spring force, we replace it with a general force derived from the double well potential included in the code. The potential is given as:

$$V(x) = k(x(1-x))^2 \quad (2)$$

Because of the relationship between potential and force in one dimension, we can take the derivative with respect to position to obtain an expression for  $f(x)$ .

$$f(x) = -\frac{dV}{dx} = -2k(1-2x)(x-x^2) \quad (3)$$

Referring to the Langevin equation (1), we substitute  $-kx$  with  $f(x)$  to obtain:

$$\gamma * dx/dt = -2k(1-2x)(x-x^2) + n(t) \quad (4)$$

With  $k = \beta = 1$ , we can simplify the equation above to a form that can be included in the nonlinear\_hints.py code.

$$dx = dt * ((2(1-2x)(x-x^2) + n(t)) \quad (5)$$

This is represented by  $x+ = dt * (f(x) + amplitude * standard_normal())$  in the code, where the amplitude is dependent on the temperature,

$$Amplitude = \sqrt{\frac{2*T}{dt}} \quad (6)$$

and when multiplied by the standard normal distribution, it creates the noise component in the Langevin equation. Running this completed version of nonlinear\_hints.py, we first view the graphic of the ball, spring,

and floor, which provide us a visual. Once the time of 50000 has elapsed, we then receive the figures of the simulation results compared to the Gibbs distribution, which is:

$$P(x) \propto e^{\frac{-V(x)}{k_B T}} \quad (7)$$

as well as the noise output. For the default  $T = 0.1$ , the outputs are shown in Figure 1.

Comparing the analytical and the numerical outputs in this case, shown in Figure 1 (a), we see that the simulation results displayed in the histogram match up well to the Gibbs distribution. Now, we change the temperature to be 1/10 of the original, so  $T = 0.01$ . The resulting output is shown in Figure 2. Then we run the simulation again with 7/100 of the original temperature,  $T = 0.007$ . The resulting output is shown in Figure 3.

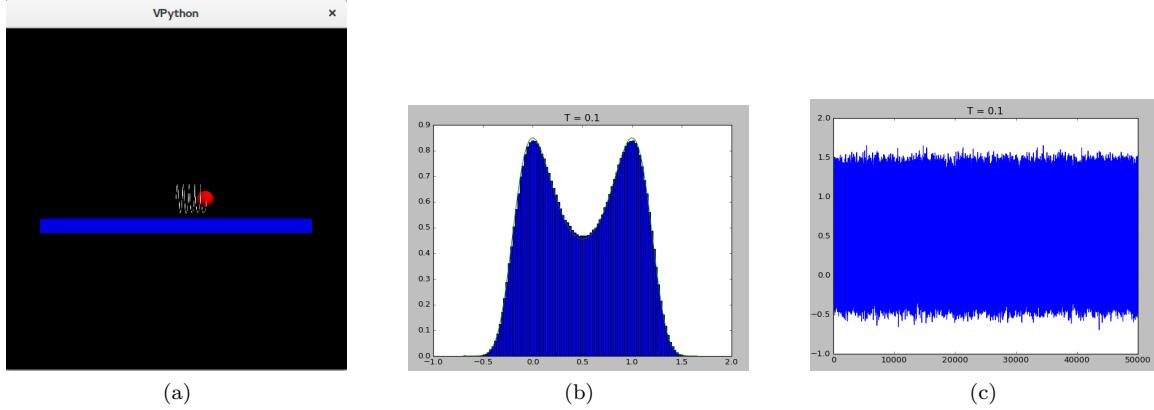


Figure 1: (a) Here the ball is randomly fluctuating— it is not oscillating because it has no inertia. This is due to it being a small object, so we can ignore the mass completely. This was taken in the case of  $T = 0.1$ , but there were negligible differences in how it appeared for the other  $T$  values. (b) The simulation results as a histogram along with the accompanying Gibbs distribution function for  $T = 0.1$ . (c) The noise output for  $T = 0.1$ .

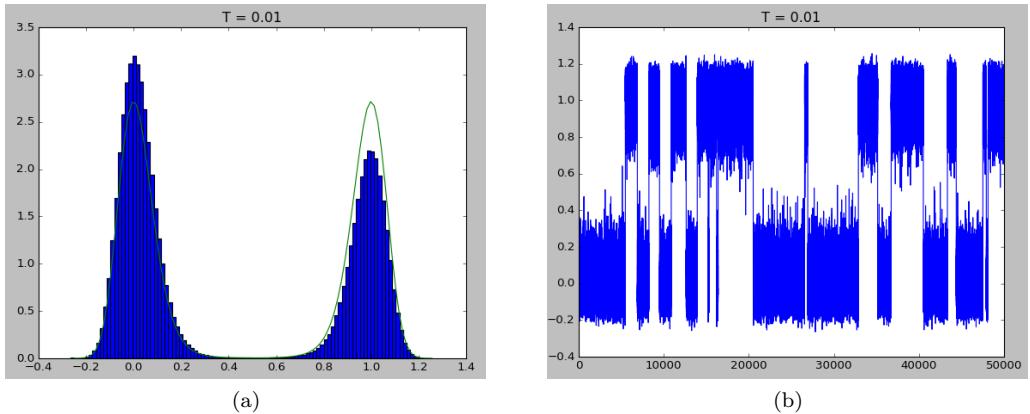


Figure 2: (a) The simulation results as a histogram along with the accompanying Gibbs distribution function for  $T = 0.01$ . (b) The noise output for  $T = 0.01$ .

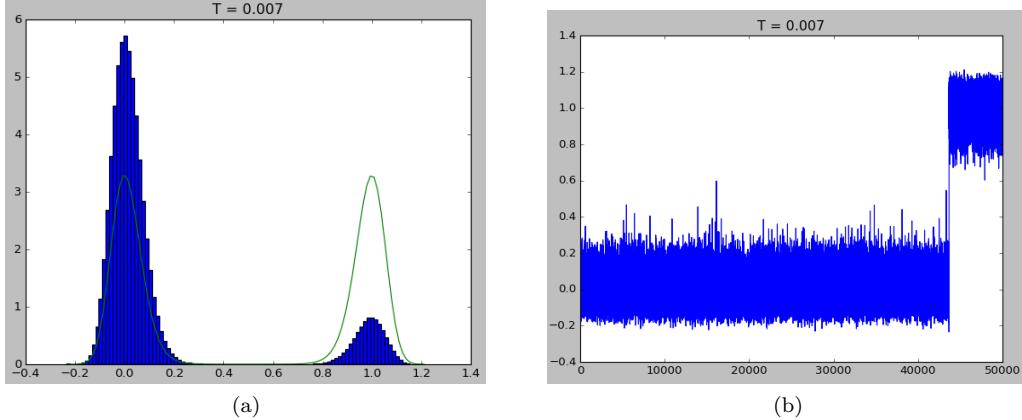


Figure 3: (a) The simulation results as a histogram along with the accompanying Gibbs distribution function for  $T = 0.007$ . (b) The noise output for  $T = 0.007$ .

In the cases when  $T$  was changed to 0.01 and to 0.007, we see a vast difference in how the theoretical Gibbs distribution and what results the simulation gave. This is due to the fact that as  $T$  decreased, the hopping time did not scale accordingly to remain large enough to the averaging time, which remained at 50000. Due to the particle starting at the well located at 0, we expect there to be more of a presence at that well compared to the right well at 1. The system is not in equilibrium, and thus the simulation does not yield the same result as it would if the averaging time would be larger. As can be seen especially in Figure 3(b), the object resides in the "0" well for a majority of the time, and once it had switched to the "1" well, it only spent a small fraction of time there, which reflected a poor match in the histogram plot with the respective Gibbs distribution. As  $T \rightarrow 0$ , and  $t$  remains at 50000, then we would see the histogram bins completely migrate to the left well of "0". It would need an infinite amount of time to see a match to that modeled by the Gibbs distribution function.

### Escape Time Over a Barrier

Here we understand and simulate particle motion in a potential field with more than one equilibrium point. We ask how long it takes on average for a particle to get from one stable state to the other in such a metastable or bistable system. The particle starts in one 'well' or stable equilibrium at  $x = 0$  and moves via Brownian motion until it reaches a threshold value that is  $x = 1$ , meaning the particle has traversed over the barrier and reached the other stable well. We run this simulation 20 times (or however many really) for a specific temperature,  $T$ , and can plot a histogram of the escape times and record the mean escape time. The potential function used to describe the field is:

$$V(x) = k(x(1-x))^2. \quad (8)$$

For a temperature  $T = 0.1$  the histogram of 20 simulated escape times is shown below in figure 1. In figure 2 the same histogram is shown but for 1000 simulations. With a larger sample size we can see more easily see that our escape time data follows a folded normal distribution as expected. As the sample size increases, the variability of sampling distribution decreases and we can get a more accurate mean escape time value.

We run this code for a range of temperature values so that we can describe the relationship between temperature and the mean escape time for the particle. We note that if the temperature gets too low, the particle will never escape. For such low temperatures the time scale for many processes is controlled by degrees of freedom in this regime. Our results agree in that as the temperature decreases, the escape time becomes exponentially larger. At different observational time scales, different degrees of freedom will have observable dynamics. Only the degrees of freedom with relaxation times on the same scale as the observation will have observable dynamics (<https://www.degruyter.com/document/doi/10.1515/jnet-2018-0008/html>). As temperature decreases,  $T \rightarrow 0$ , degrees of freedom freeze out. The probability of finding the system in the exited

states of energy,  $E$ , goes like  $\exp(-E/T)$ , which says that at low temperatures, the probability that we will find the system in such an excited energy state exponentially decreases. With a decrease in temperature, the particle loses kinetic energy and there is a decrease in the rate of diffusion. Hence at low temperatures it takes exponentially longer for the particle to go from one well to the other.

From this we can come up with a formula relating the escape time to the temperature of the system. We know that the probability of finding a particle around a point  $x$ ,  $P(x) \propto e^{-V(x)/k_B T}$ , where  $V(x)$  is the potential field and  $k_B$  is the Boltzmann's constant. We see that the probability of finding a particle decreases exponentially with increased potential,  $V(x)$ , and increases exponentially with temperature,  $T$ . From this we can also say that the time it takes for a particle to escape over a potential barrier is exponentially larger for larger  $V(x)$  and exponentially smaller for larger  $T$ . This writes out as

$$t \propto \exp(\Delta V/k_B T) \quad (9)$$

We can see that this analysis agrees with that of Kramers' time shown below.

As analytically described in the slideshow, Biological Physics, Diffusion (Canvendish Lab), utilizing the Fokker-Planck equation tells us that the time it takes to escape over a potential barrier with potential  $V(x)$  that has a difference of  $V_{\max}(x) - V_{\min}(x) = \Delta V$  is described by Kramers' time:

$$t \approx \tau \sqrt{\frac{\pi kT}{4\Delta V}} \exp\left(\frac{\Delta V}{kT}\right). \quad (10)$$

And what's important here is the dominating exponential factor in this equation which lets us write,  $t \propto \exp(\Delta V/kT)$ . We also know that Kramer's formula is only asymptotically correct when  $\Delta V \gg kT$  otherwise stated  $\frac{\Delta V}{T}$  is very large. This means that to correctly simulate Kramer's escape time, we don't want to choose temperature values that are too large.

A plot of the average escape time vs. temperature looks like a decreasing exponential as expected (figure 7). In terms of fitting this data it makes the most sense to take the log of both sides of equation 5 such that we are left with a linear relationship between average escape time and temperature. This will also weigh large  $1/T$  more strongly when doing a linear curve fit. We rewrite equation 5 as:

$$\log(t) = \frac{\Delta V}{k} \frac{1}{T}. \quad (11)$$

We let  $k = 1$  for simplicity. We see that if we plot  $\log(t)$  vs.  $1/T$ , our data should show a linear relationship with a slope of  $\Delta V(x)$ . Because we know the form of the potential as described by equation 4, we can analytically calculate  $\Delta V(x)$ . We can calculate the max and min of  $V(x)$  by taking the first derivative and setting equal to zero:

$$\frac{dV}{dx} = 2(2x - 1)(x^2 - x) = 0 \rightarrow x = \frac{1}{2}, 0, 1. \quad (12)$$

Plugging back these values of  $x$  into  $V(x)$ , we see that  $V(x = 0, 1) = 0$ , and  $V(x = 1/2) = 1/16 = 0.0625$ . Now let us confirm this result with our numerical data and a linear regression fit. The mean escape time was calculated for a range of temperature values. Figure 8 shows a plot of  $\log(t)$  vs.  $1/T$  for values of  $T$  ranging from 0.01 to 0.1. Choosing values of  $T$  smaller than 0.01 is also suitable, but makes for long simulation times so I chose to stick with values of  $T$  that made the simulation run decently fast.

The slope of the fitted line was calculated to be 0.06251. This value for  $\Delta V$  agrees very nicely with the value derived analytically. Overall we see that the numerically derived relationship between average escape time and temperature agrees well with the theory.

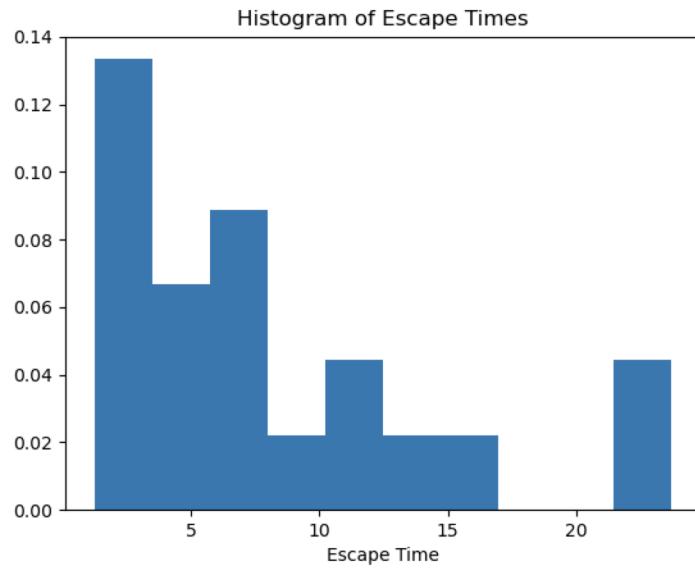


Figure 4: Histogram of escape times for a system temperature  $T = 0.1$ . Sample size is 20 particle simulations. Mean escape time is 8.06 seconds.

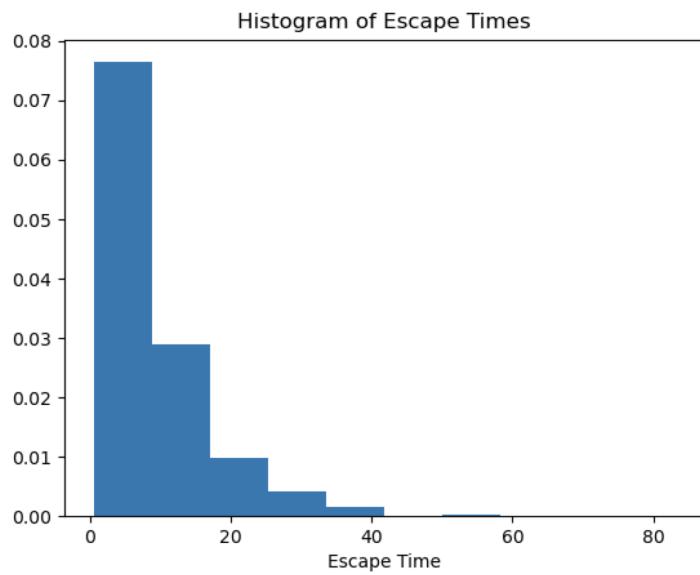


Figure 5: Histogram of escape times for a system temperature  $T = 0.1$ . Sample size is 1000 particle simulations. Mean escape time is 8.95 seconds.

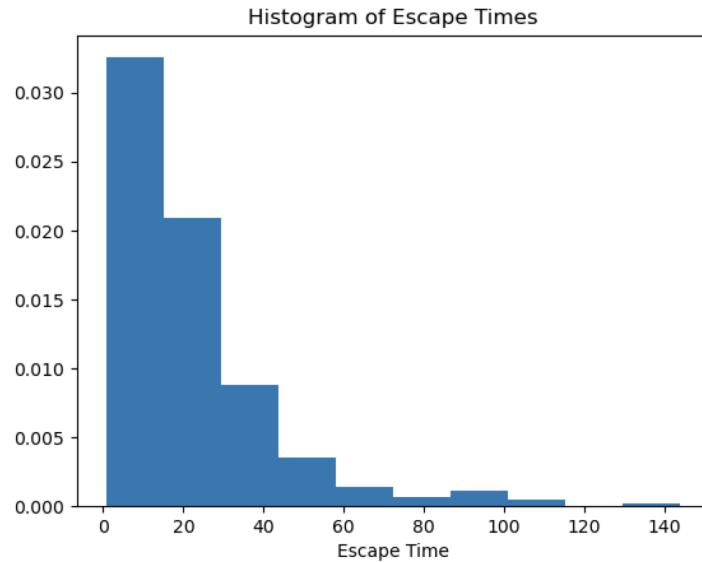


Figure 6: Histogram of escape times for a system temperature  $T = 0.05$ . Sample size is 300 particle simulations. Mean escape time is 21.75 seconds.

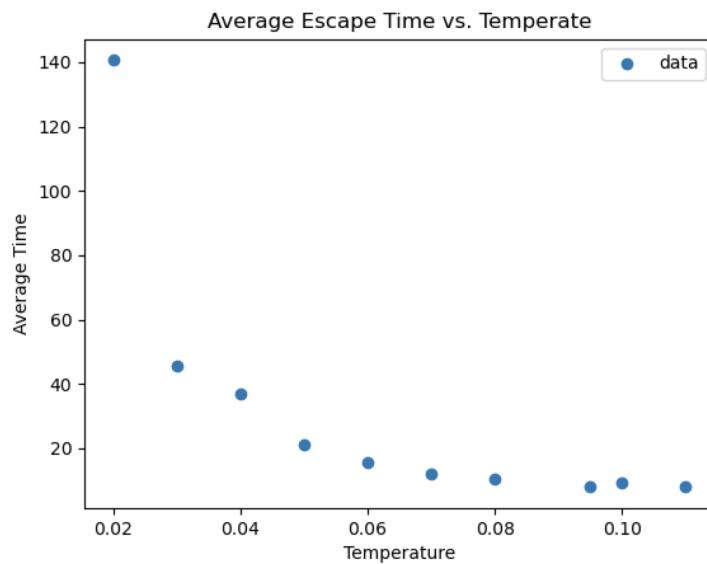


Figure 7: Plot of simulation data. Mean escape time vs. temperature. We see that as the temperature increases, the time it takes for the particle to escape from one well into the other decreases exponentially.

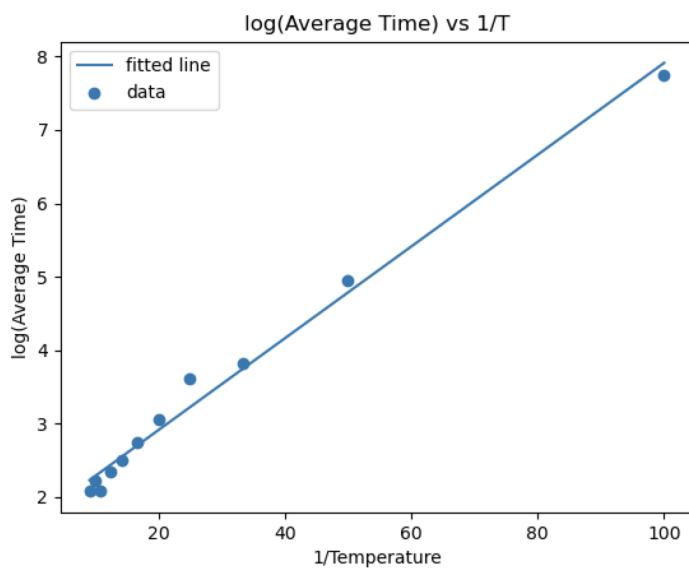


Figure 8: Plot of simulation data of  $\log(\text{average escape time})$  vs.  $1/T$ . Also shown is a fitted linear curve using linregress from `scipy.stats`. Slope of curve is calculated to be 0.06251 and the y-intercept is 1.663.

## Applications

The activation energy and rate of a chemical reaction are closely related, because the higher the activation energy, the slower the chemical reaction will be. The reason for this is that molecules can only complete the reaction once they have reached the top of the activation energy barrier. Yet, the higher the barrier is, the fewer molecules will have enough energy to make it over this energy barrier at any given time. Given a specific average temperature, single molecules in a sample will have a fairly wide range of kinetic energies, meaning that only a tiny fraction of molecules will make it over a high activation energy barrier, and the reaction will be very slow, therefore it will take a lot of time to get any measurable amounts of molecules (Nelson et al., 2008). Quite a large number of reactions have very high activation energies and will need an input of external energy to react. A great example is the combustion of a fuel like propane. It releases energy, yet the rate of reaction is effectively zero at room temperature. This is in our interest since canisters would spontaneously combust in the sales room or a storage place. A spark will then provide energy to get some molecules over the activation energy barrier, which in turn complete the reaction, releasing energy. This released energy helps other fuel molecules get over the energy barrier, leading to the reaction we seek in combustion engines (Engel et al., 2010). There are also chemical catalysts instead of sparks to facilitate propane combustion such as Lanthanum-Manganese Oxide (La-MnO<sub>x</sub>) (Xie et al., 2016).

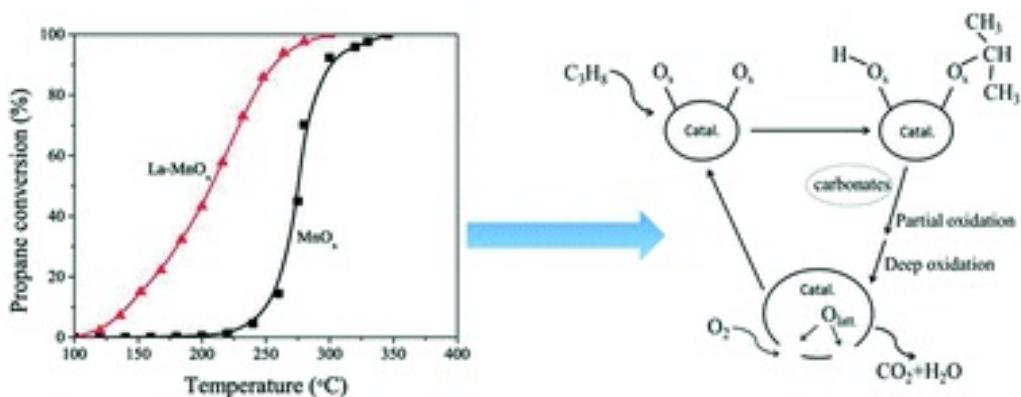


Figure 9: Propane conversion reaction showing how La-MnO<sub>x</sub> activated at a much lower temperature. The reaction pathway is shown on the right (Xie et al., 2016).

A majority of chemical reactions in cells are like this propane combustion example. When the activation energy is too high for the reactions to proceed significantly at ambient temperature, therefore the activation energy of the reaction must be lowered to increase the reaction rate. This process is known as catalysis, where the factor that is added to lower the activation energy is called a catalyst, and in biological systems and reactions catalysts are known as enzymes (Engel et al., 2010).

Another good example is the controlled breakdown of glycogen and release of glucose. This controlled breakdown increases the amount of glucose that is available to us between meals and serves as a buffer for blood-glucose levels maintenance. Glucose is the fuel used by our brain, except during prolonged starvation, but glucose also is readily mobilized for sudden strenuous activity. Glucose can provide energy in the absence of oxygen, supplying energy for anaerobic activity in our muscles. The two major sites of glycogen storage are the liver and skeletal muscle.

Glycogen synthesis requires an activated form of glucose, uridine diphosphate glucose (UDP glucose), which is formed by the reaction of uridine triphosphate and glucose 1-phosphate. UDP-glucose is added to the nonreducing end of glycogen molecules. Glycogen degradation requires the glycogen molecule to be remodeled for continued synthesis, and the regulation of this is complex and requires several enzymes to control the activation energies. These enzymes function allosterically allowing the adjustment of enzyme activity to meet the needs of the cell in which the enzymes are expressed. Furthermore, glycogen metabolism is regulated by hormonally stimulated cascades, leading to reversible phosphorylation of enzymes, and altering their kinetic properties, and this allows glycogen metabolism to adjust to the needs of the entire organism.

Glycogen degradation is integrated with glycogen synthesis by both these mechanisms (Berg et al., 2002).

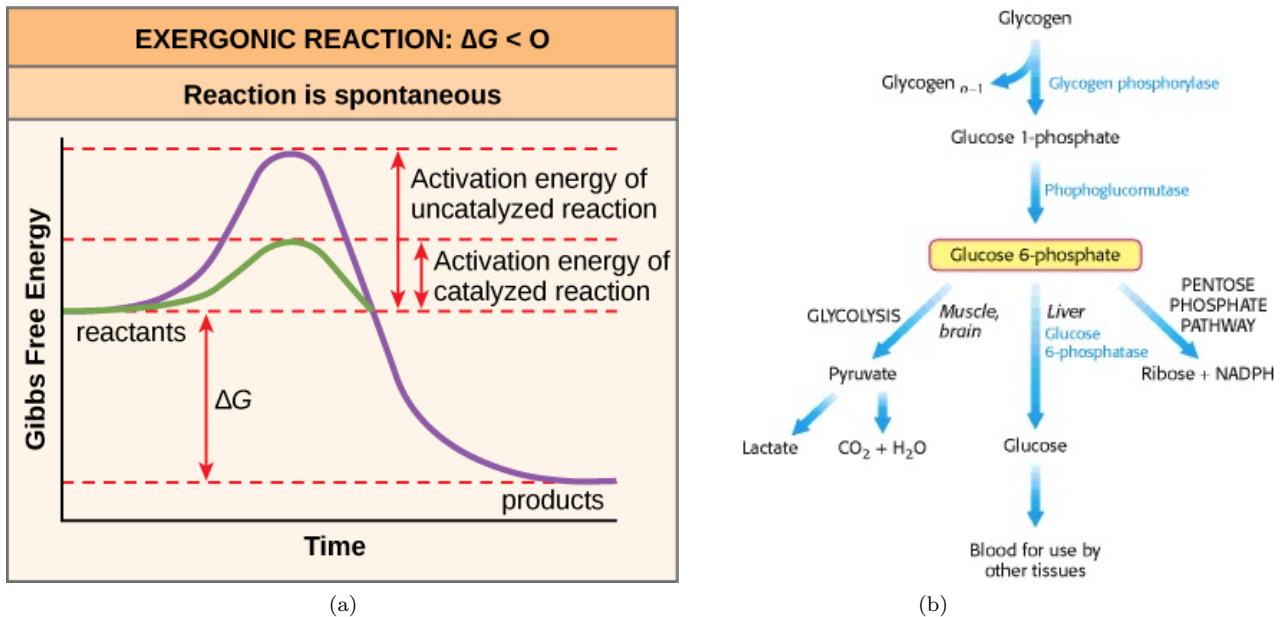


Figure 10: (a) An exergonic reaction showing the activation energy (Gibbs Free Energy) with and without using a catalyst. (b) The conversion of glycogen and different pathways are shown, as well as the catalytic enzymes (blue font) (Berg et al., 2002).

If we consider our model of a barrier, the activation energy is the height of that barrier and what is necessary for a molecule or particle to get over it. Catalysts such as enzymes lower this activation energy in form of a height of a barrier, and the possibility of more molecules getting over it is higher in a specific time and at a specific average temperature, also known as increasing or raising the rate of the reaction. Now it is also possible to raise the barrier with such reactions using inhibitors, which is also quite common in biological systems. Inhibitors will decrease or lower the rate of the according reaction.

Molecular motors function as cellular machines by converting chemical energy into mechanical work, and play key roles in many intracellular biological processes, including signaling and the transport of cargo. Myosin V, one class of molecular motors, performs these functions by binding to actin filaments and generating energy through ATP hydrolysis. It is a dimeric transport motor and is composed of two polymer chains joined at a pivot. Each chain has an actin-binding head and walks forward along the actin in a hand-over-hand stepping motion through alternating head detachment. The free head performs a diffusive search for actin binding sites during each step. This unidirectional motility requires a coordination called gating mechanism, regulated by the strain on the lever arms while the heads are bound to actin. Myosin V moves toward the plus end of the actin using two changes in the lever arm orientation, of which one is known as the power stroke, executed by an actin-bound head. It swings the lever arm forward, while the recovery stroke executed during diffusion, returns the lever to its original orientation (Hathcock et al., 2020). According to Astumian and Bier, 1996, These motions of Myosin are each using ATP hydrolysis to catalyze the activation energy and pass the barrier so that the reaction and hence the movement can proceed faster.

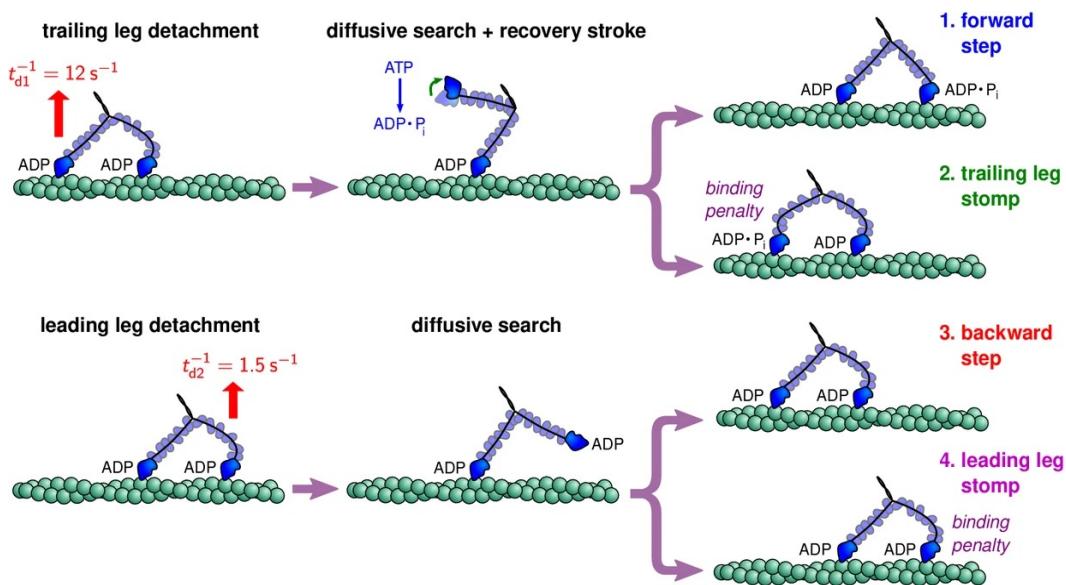


Figure 11: Myosin V's range of behaviors include steps and stomps involving each polymer leg. Myosin V starts in a waiting state with both legs strongly bound to the actin where each head has an associated ADP molecule. Each leg is in the post-power stroke orientation toward the plus end of actin. Leading leg is bent backward under tension. Myosin V then goes through one of many kinetic pathways, using enzymes for catalysis to lower the activation energy, hence lowering the barrier, and to speed up the reaction (Hathcock et al., 2020).

## 2 Modeling Bacterial Chemotaxis

### Intro

Bacterial chemotaxis is the movement of an organism in response to a chemical stimulus. Many cells, bacteria, and multicellular organisms direct their movements according to certain chemical profiles in their environment. Here we focus on bacteria which move in response to chemical gradients using flagella that rotate through the solution to push them in a certain direction. By comparing the present chemical concentration to the concentration at the cell's previous location, a cell can map out the chemical gradient in their environment at a basic level in order to direct its motion toward environmental conditions it deems attractive and/or away from surroundings it finds repellent (Webre et al., 2003). In simulating flagellated bacteria chemotaxis we consider the case of a chemical profile which the bacteria is attracted to (i.e. a food or energy source). The bacteria in this case wants to move itself to regions of highest chemical concentration.

The movement of flagellated bacteria is described by two kinds of behavior: smooth-swimming runs and random tumbles. The bacteria move using their flagella in more or less a straight line and then occasionally tumble to reorient themselves so that they can travel in a straight line in a new direction. An impressive feat of bacterial chemotaxis is how fast they are able to sense and compare the concentration gradient to then randomly reorient themselves. Another important aspect of chemotaxis is the bacteria's ability to change how often they tumble. In short, if the chemical concentration is increasing, the bacteria will tumble less often, if it is decreasing, they will tumble more often. This way if the bacteria is moving in the wrong direction aka areas of lower concentration, it can quickly tumble and try a random new direction to move in. If it is heading towards or is in areas of higher concentration there is less of a need to tumble and switch its direction of movement. So for this case of a bacteria searching for food, when it is moving up a gradient of chemical attractant, it tends to continue its course; an optimal aspect of this interestingly intelligent chemotaxis strategy. This idea is visualized in figure 1.

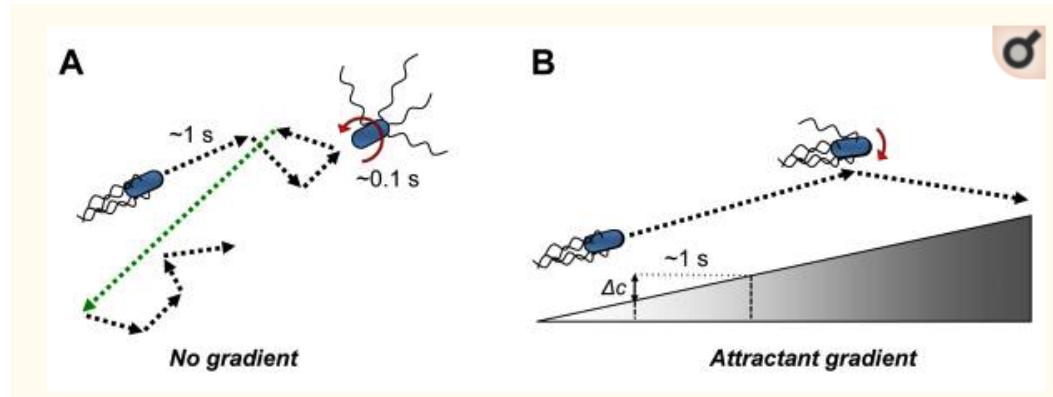


Figure 12:

(a) Chemotaxis of flagellated bacteria such as *E. coli* swimming in the absence of a chemical gradient. Movement consists of straight runs that are interrupted by a short tumble that reorients the bacteria in a random direction. Tumbles are caused by the clockwise rotation of one or several flagella. These tumble-run movements are a form of a random walk or Brownian motion and ensure effective exploration of the bacteria in its environment.

(b) Chemotaxis strategy of bacteria swimming up an attractive chemical gradient. Moving up the chemical gradient is ideal for the cell and hence it prolongs its runs and tumbles less often. The degree of cell body reorientation during these tumbles is smaller as well (Sourjik Wingreen, 2013).

The code, `chemotaxis.py`, lets us explore and simulate chemotaxis of flagellated bacteria. We can explore different concentration profiles and visualize the resulting distribution of the bacteria's position in a histogram.

## 1. Concentration Profiles

We can start by running this code for two distinct concentration profiles. The linear concentration profile is defined as  $c(r) = gr$ , where  $g$  is a gradient constant, and the quadratic concentration profile, which serves like a potential well, can be described as  $c(r) = kr^2/2$  where  $k$  is a constant.

Throughout the code, the concentration is being defined with a negative sign so that it agrees qualitatively with the notion of a potential. For the case of an attracting chemical, the smaller the concentration (potential), the more attractive it is to the bacteria. And since we model the concentration with a negative sign, let us think of this as the bacteria want to move to areas of high attractant concentration aka peaks in the concentration profile (when defined as  $-c(r)$ ). The bacteria aims to decrease its potential.

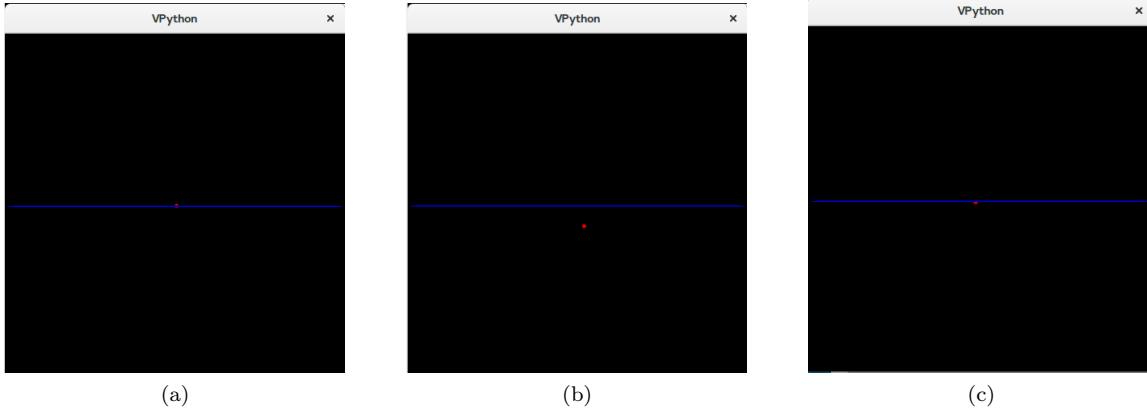


Figure 13: (a) Displays the graphic output for both the linear and quadratic concentration profiles when the iteration range (simulation run-time) is 1k. (b) Displays the starting graphic output for the quadratic concentration profile when the iteration range is 10k. For a few seconds, the red particle will move erratically. (c) Displays the ending graphic output for the quadratic concentration profile when the iteration range is 10k.

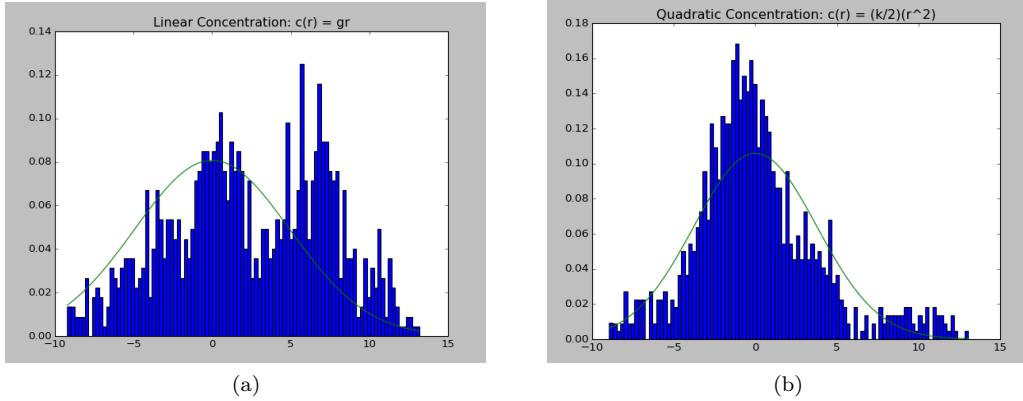


Figure 14: (a) Displays the histogram of the bacteria position and its comparison to a Gaussian for a linear concentration profile with an iteration range of 1k. The sigma value is 4.94. (b) Displays the histogram of the position and its comparison to a Gaussian for a quadratic concentration profile with an iteration range of 1k. The sigma value is 3.77.

For a linear concentration we note that  $c(x) \propto x$  has no equilibrium state, so we do not expect the bacteria to find an equilibrium state. If we do not bound the simulated spacial region and give the particle an infinite amount of time, we expect it to continue up the chemical gradient forever or at least try to since the gradient is always increasing with increasing  $r$ .

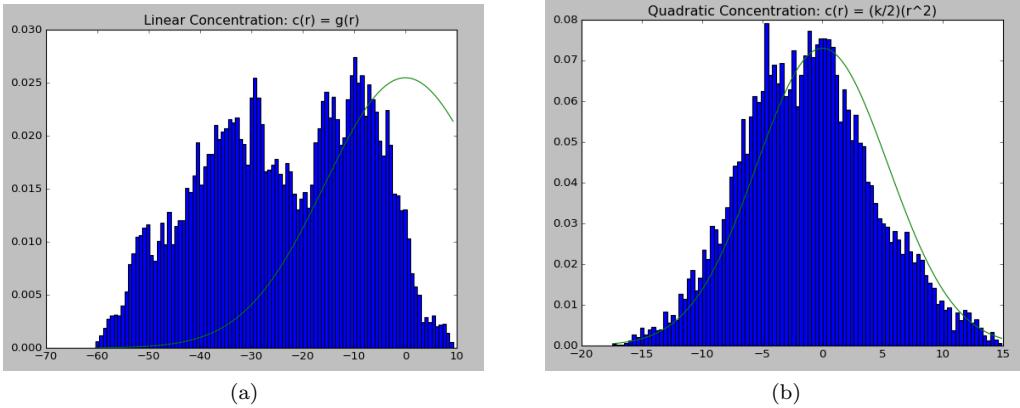


Figure 15: (a) Displays the histogram of the bacteria position and its comparison to a Gaussian for a linear concentration profile with an iteration range of 10k. The sigma value is 15.66. (b) Displays the histogram of the position and its comparison to a Gaussian for a quadratic concentration profile with an iteration range of 10k. The sigma value is 5.46.

For the quadratic concentration profile, we see from the histogram that the particle finds itself in some places much more than others as it should. We see that it indeed finds itself near the area of highest concentration (the origin) more than areas that extend farther away from here and that are of lower concentration.

### (2,3) Comparison to Brownian Particle and Bacterial Density Model

In coming up with a model for the bacterial density that works for different concentration profiles it is important to understand how this model compares to that of a Brownian particle in a potential well. As discussed in the previous assignment (diffusion in a double well potential), the equilibrium solution to the Langevin equation is an ideal model for the distribution of the particles position over time. This is the Boltzmann Gibbs distribution described by  $p(x) \propto \exp(-\frac{V(x)}{k_B T})$ . Although this distribution does not model bacterial chemotaxis well, there are similarities which should be noted in order to see what type of probability distribution can model chemotaxis. In both cases (diffusing particle in a potential, and bacterial chemotaxis), we know that the particles are localized around regions of high concentration (low potential). However what is different about chemotaxis is that the rate at which the particle decides to tumble (how often they randomize their direction) is dependent on the attractant concentration. For a diffusing particle, the rate of the random walk is constant.

We note that a bacterium tumbling randomly does look diffusive for long times but over short distance it moves in straight lines as discussed. So the fact that diffusion is much different than bacteria tumbling for short time scales means we can't model chemotaxis with only a Gibbs distribution.

What we find by comparing a Gibbs distribution such that:

$$p(x) \propto \exp(-Ac(x)) \quad (13)$$

where  $A$  is a constant and  $c(x)$  is the attractant profile, with the histogram of the particles position from the chemotaxis simulation, is that the distribution of particles position is much more spread out than the Gibbs distribution model. Meaning that compared to Brownian motion in a potential, the particle in chemotaxis spends more of its time in areas of lower concentration (equivalent to higher potential) if it is in that area. This make sense, as for a particle undergoing brownian motion in a potential, once it is in an area of low potential it has a tendency to stay there and it is harder for it to escape to areas of higher potential.

Analytically we can say that for brownian motion in a potential, the probability distribution is able to follow  $V(x)$  such that if  $(-V(x))$  is very sharply peaked,  $p(x)$  can follow that and also be very sharply peaked.

However this can't happen with the tumbling mechanism in chemotaxis if we are to relate the probability as in equation 13. But the particle is still found more often in regions of higher concentration,  $(-c(x))$ . So to model this probability distribution, equation 13 should be of some use.

To best fit the distribution of the particles position in chemotaxis, let us consider utilizing a convolution that will smear out  $p(x)$  described by equation 13. A convolution is essentially a way of combining two signals to form a new signal. It can be thought of as a way to filter an input signal and produce the resulting filtered signal. In our case, we have an input distribution  $p(x)$  which we want to smear out or filter with a specific function  $f(x)$ . The resulting convolution is the integral of the product of the two functions:

$$(p * f)(x) = \int_0^x p(x-s)f(s)ds. \quad (14)$$

We can also define a convolution in terms of the fourier transforms of  $p(x)$  and  $f(x)$ , such that:

$$F[p(x) * f(x)] = F[p(x)]F[f(x)] \rightarrow p(x) * f(x) = \text{realIF}[F[p(x)]F[f(x)]]. \quad (15)$$

This method can be a simple way to numerically calculate a convolution. We can also use a module such as `numpy.convolve()` so calculate the convolution.

So what type of function  $f(x)$  should we use to smear out  $p(x)$  as to make that probability distribution wider? It would make sense to convolute  $p(x)$  with either a gaussian or poisson distribution. Both should smear out  $p(x)$  as desired.

We know that the length scale for the straight line motion in chemotaxis is related to the tumbling time scale. For a gaussian fit we consider convoluting with something that dies off as  $\exp(-(x/L)^2)$  where  $L$  is a length scale constant. For a poisson fit we consider convoluting with something that dies off as  $\exp(-(|x|/L))$ .

The results of these convolutions and their normalized fit against the distribution data is shown below. A variety of concentration profiles,  $c(x)$ , were tested. It is important to only consider cases where there is a well defined probability as  $t \rightarrow \infty$  otherwise stated  $c(x)$  has a defined equilibrium state. These include concentration profiles such as a single well, double well, negative gaussian, etc.

Figure 16 shows the particles distribution and the best fit model determined by  $\text{convolve}(p(x), f(x)) = \text{convolve}(\exp(-Ac(x)), \exp(-(x/L)^2))$ . The concentration profile chosen was that equivalent to a single well potential,  $c(x) = kx^2/2$ , where  $k = 10$ . We see that the resulting normalized convolution is indeed a great fit to the gaussian-like profile of the data. Best fit parameter values were found to be  $A = 0.3$ ,  $L = 2.6$ . If this were data from brownian motion in the same potential profile, we would find that it follows the orange curve. For this chemotaxis system we see that the particles position is better fit by a convolution with a gaussian profile,  $f(x) = \exp(-(x/L)^2)$ , which acts to smear  $p(x)$ .

Figure 17 shows the histogram position data plotted against the same convolution fit except now the concentration profile has been switched to that which resembles a double well potential,  $c(x) = k((x-3)(x+3))^2$ . The same best fit parameters were used. Again, we see that the convolution fits the data quite nice. We also test the fitting model with a triple well concentration profile as seen in figure 18. Although the convolution fit does highlight the three corresponding peaks in the density data, the parameters used to construct the fit were not the best fit values; used was  $A = 0.3$ ,  $L = 2.1$ . The best fitting module `leastsq()` from `scipy.optimize` was giving me problems at this stage in the simulation progress so I was not able to extract the best fit parameter values for  $A$  and  $L$ .

It's important to note that these profiles for  $c(x)$  are unbounded. For these unbounded profiles it was apparent that a gaussian kernel,  $f(x) = \exp(-(x/L)^2)$ , in the convolution makes for a better fit than a poisson kernel,  $f(x) = \exp(-|x|/L)$ . However for an attractant profile that is bounded such as  $c(x) = -\exp(x^2)$  ( $c(x)$  goes to zero rapidly), we might expect a poisson kernel in the convolution to result in a better fit. A bounded attractant profile can be described as a localized food source with no food outside this area. Ultimately if I had more time to explore this chemotaxis simulation I would have liked to successfully fit some bounded concentration profiles. There were a handful of problems I ran into in trying to fit to the density data from a bounded concentration profile. The resulting fit was not a great fit and I was unable to numerically find best fit parameters for  $A$  and  $L$ .

All-in-all this was a fun problem to tackle and the specific convolution model showed to be a good fit to the density data for some of the unbounded attractant profiles simulated. There are lots of potential ways to smear out  $p(x)$  defined by equation 13 (equilibrium solution to Langevin equation) in order to best fit the bacterial density. For example,  $\exp(-A\text{convolution}(c(x), \exp(-|x|/L)))$  or  $\exp(-A\text{convolution}(c(x), \exp(-(x^2/L))))$  might also fit to the data quite well. Fitting with these additional convolutions is something I would have liked to do if I had more time to explore this problem.

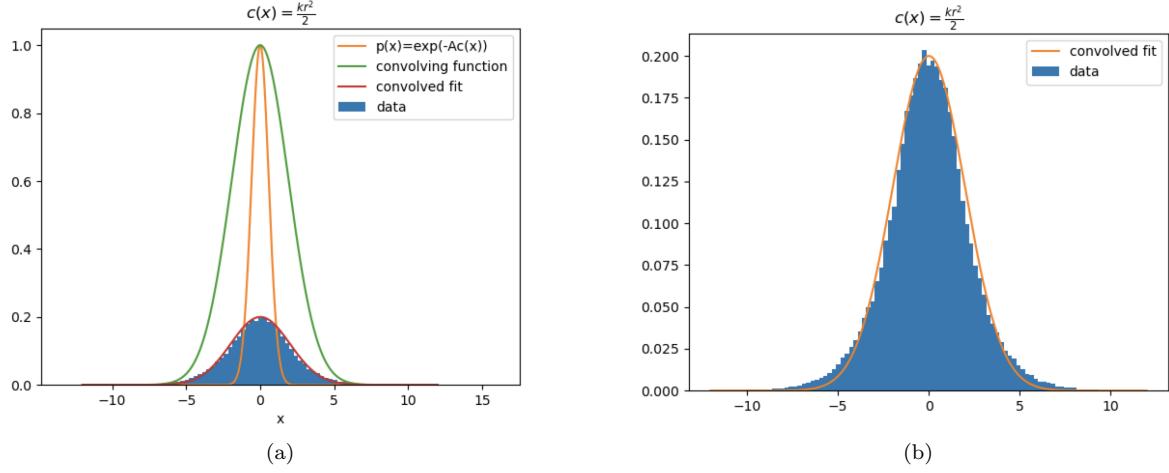


Figure 16: (a) Histogram of particles position from chemotaxis simulation. Concentration profile,  $c(x) = kx^2/2$  is equivalent to a single well potential. Particle finds itself the most at  $x=0$  where the attractant concentration is the highest. Plotted against the data is the fitted model which is a normalized convolution of  $p(x) = \exp(-Ac(x))$  with a gaussian kernal,  $f(x) = \exp(-(x/L)^2)$ . Best fit parameters were determined to be  $A = 0.3$ ,  $L = 2.6$ . (b) Same histogram as in left figure. Plotted is just the data and the best fit model.

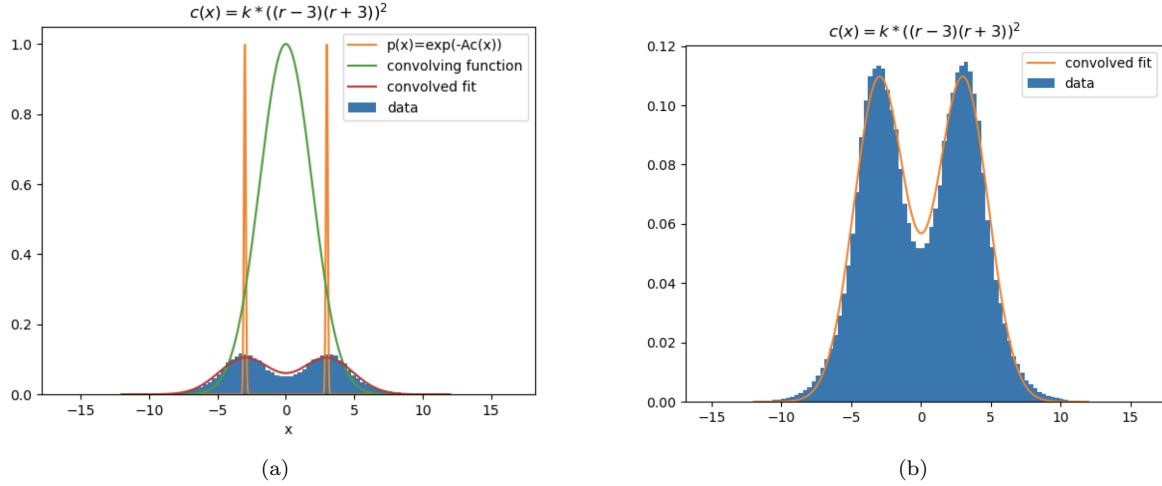


Figure 17: (a) Histogram of particles position from chemotaxis simulation. Concentration profile,  $c(x) = k((x - 3)(x + 3))^2$  is equivalent to a double well potential. Particle finds itself the most at  $x = \pm 3$  where the attractant concentration is the highest. Plotted against the data is the fitted model which is a convolution of  $p(x) = \exp(-Ac(x))$  with a gaussian kernal,  $f(x) = \exp(-(x/L)^2)$ . Best fit parameters were determined to be  $A = 0.3$ ,  $L = 2.6$ . (b) Same histogram as in left figure. Plotted is just the data and the best fit model.

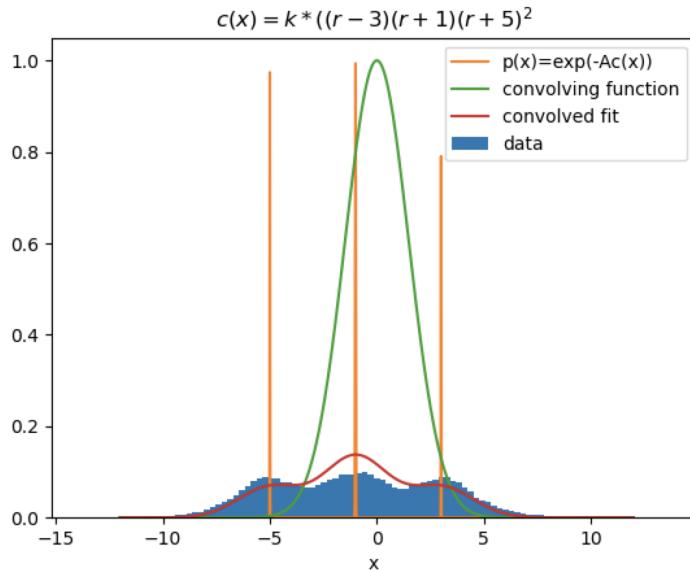


Figure 18: Histogram of particles position from chemotaxis simulation. Concentration profile,  $c(x) = k((x - 3)(x + 1)(x + 5))^2$  is equivalent to a triple well potential. Particle finds itself the most at  $x = 3, -1, -5$  where the attractant concentration is the highest. Plotted against the data is the fitted model which is a convolution of  $p(x) = \exp(-Ac(x))$  with a gaussian kernal,  $f(x) = \exp(-(x/L)^2)$ . Best fit parameters were determined to be  $A = 0.3$ ,  $L = 2.1$ . Note that the fit is not as good as for the single and double well concentration profiles because the best fitting parameter output was not cooperating as well.

## (4,5,6) Biological Discussion and Examples

For prokaryotic cells a minimum of about 250 to 450 essential genes are required for viability. The minimal viable cell diameter is expected to lie in the range of 250 to 300 nm, and the number of ribosomes required for adequate genome expression is a significant constraint on minimal cell size (nap.edu, 2021). Generally cell size is limited by the ratio of surface area to volume. The larger the cell, the smaller the ratio, hence more energy is needed to be motile and to fuel receptors. Therefore a small cell size with many receptors is the optimal design to move quickly and with minimum energy expenditure. Mycoplasmas are known as the smallest free-living organisms and therefore considered to be the simplest of any bacteria, with the smallest genome. Belonging to the bacterial class of Mollicutes, they are distinguished by their lack of a cell wall and a plasma-like form. The Pasteur Institute isolated the first strains of mycoplasma in 1898, and there are roughly 190 known to date. Due to their extremely basic genomes, they function as parasites in order to meet their energy and biosynthesis demands, exploiting host cells to survive. At approximately 100 nm, they are undetectable by the naked eye or even by optical microscopy. Because of their lack of a cell wall, they are furthermore resistant to many common antibiotics. Hundreds of mycoplasmas can attach to a single eukaryotic cell, invading the host by fusion with the cell membrane. Upon entry into the cell, they multiply, eventually outnumbering host cells by the thousands, and they can circumvent host defenses to survive (Invivogen.com, 2021). Ubiquitous in nature, they are widespread as parasites of all mammals, reptiles, fish, arthropods, and plants. They usually exhibiting organ and tissue specificity. The mycoplasma attaches to the outside cellular membrane, which will result in interaction between certain bacterial proteins, such as lipoproteins, lipopeptides or specific attachment organelles, and with specific cellular receptors on the surface of the target cells (Benedetti et al., 2020). Receptors in the lipoproteins of this organism find the host cells, but they are not selective in which cell they proliferate, but are usually first found in lung tissue. As for which cell size and gradient sensing have enough certainty, yeast cells are not motile, which suggest they use a spatial sensing mechanism, despite being smaller, at approximately 4m in diameter, than most eukaryotic cells. Yeast cells are capable of sensing linear gradients as shallow as 0.1 nM/m, and all information on the extracellular pheromone gradient comes from receptors on the cell's surface. Therefore, these receptors set the ultimate limits on gradient sensing (Lakhani and Elston, 2017).

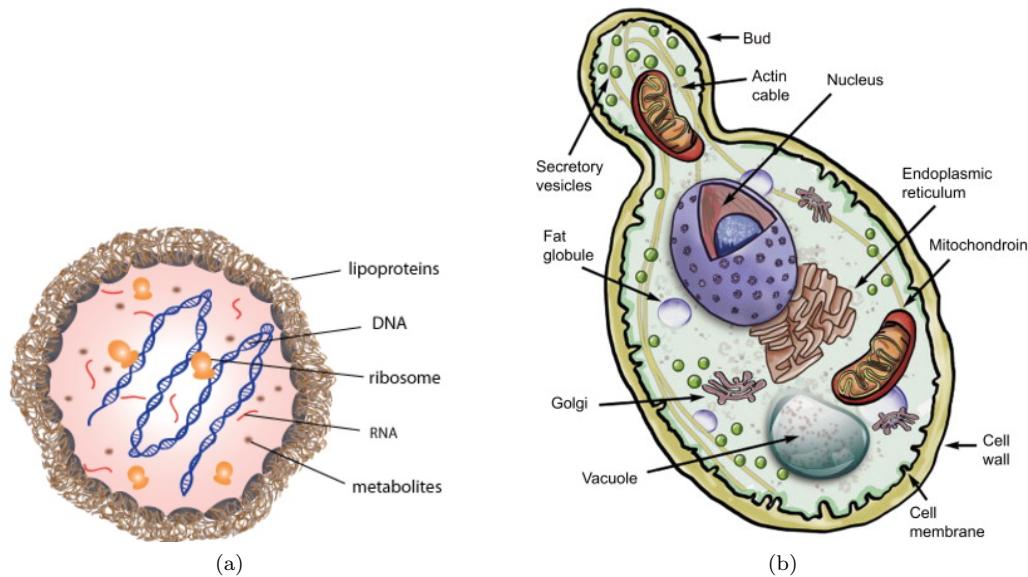


Figure 19: (a)The smallest known free-living organisms, the mycoplasma, shown are its cellular components (Invivogen.com, 2021). (b) Budding yeast shown with the bud (top) and organelles as well as DNA. (Speers and Forbes, 2015).

In the case of a static chemical gradient, infinite time, and many receptors, a cell would definitely be able to find a concentration gradient, but the question arises whether it can expend enough energy for that infinite

timeline. Receptors can be many, but they can be for a diversity of different chemicals, some repellent, some attractive, because both are needed to keep the cell thriving and from danger. They could also interfere with one another, due to the close proximity. The chances of finding a concentration gradient have higher probability when these receptors are a bit farther apart, yet time being infinite is not of any use. The general function of the cell will require energy within a boundary of time so that it doesn't starve. In the case of lots of food and lots of time, quite the opposite will occur. The cell will not evolve and create new sensitivity in its receptors or even create new receptors when dividing, but rather would develop less, since it has all that it needs and won't diversify. That being said, infinite time and infinite food are not good for survival mechanisms. Organisms and cells develop survival skills by the opposite scenario usually, prompted by little resources, diversity in gradients, and competition.

What we see as described by Berg and Purcell is that “the number of receptors a cell can usefully employ is not much larger than the ratio of cell diameter to patch diameter” (1977). A cell utilizing many more receptors than this ratio does not significantly increase its ability to detect a specific chemical. Therefore a cell of large enough size can theoretically have many, even hundreds, of such specific receptor systems consisting of multiple spread out receptor patches. And a cell can utilize this almost as effectively as if the entire surface of the cell was dedicated to detecting a specific chemical. From this, we can deduce that for cells of large enough size it would be optimal to have many receptors spread around the cell’s surface rather than an array of receptors at its head. As Berg and Purcell show, the probable error in a detected chemical concentration will be the same as that for a single receptor, given a specific total number of new molecules that have occupied the receptor patch during the observation period (1977). Therefore, in theory, having a uniform and maximal distribution of receptors over the cell’s surface would increase the amount of substance measured over time and hence allow the cell to find food faster. However, if the cell is too small and the dispersed receptors are too close together, random fluctuations in the number of molecules that are registered by the cell will not be clearly distinguishable as the difference between the signals is smaller than that for a larger cell with receptor patches more widely dispersed.

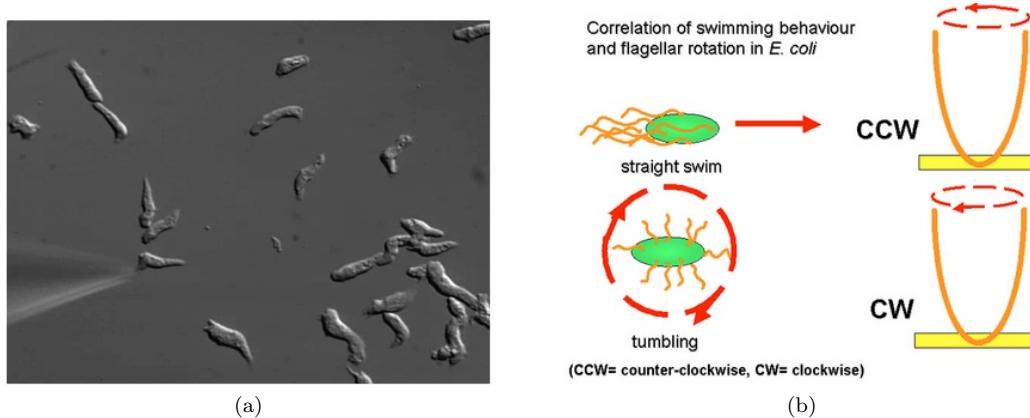


Figure 20: (a) Dictyostelium cells are chemotaxing towards a micropipette (bottom left) releasing the chemo-attractant cAMP. (Janetopoulos and Firtel, 2008). (b) In *E. coli* counter-clockwise rotation will align the flagella into a single rotating bundle, which will let the bacterium swim in a straight line. Clockwise rotation separates the flagella and each flagellum can point in a different direction, with the effect that the bacterium can tumble and rotate (Wikipedia, 2021).

Another strategy that could theoretically increase the cells intake of chemical molecules is local stirring. It has been suggested that this might be a function of the cells flagella (Berg and Purcell, 1977) along with producing the tumbling mechanism. The objective here is to transfer distant fluid from a region of relatively high chemical concentration to a region closer to the cell. This increases the concentration gradient near the absorbing surface of the cell and hence could possibly replace or accompany random tumbling which the cell goes through in order to move closer to higher chemical densities in the solution. However, there are prominent trade-offs with this strategy in that energy expenditure is high if a cell is to significantly increase

its chemical intake. Berg and Purcell found that despite the exact details of the stirring mechanism, “local stirring for the purpose of increasing intake changes from a hopelessly futile to a possibly useful activity somewhere in the range of cell size between a few microns and a few tens of microns” (1977).

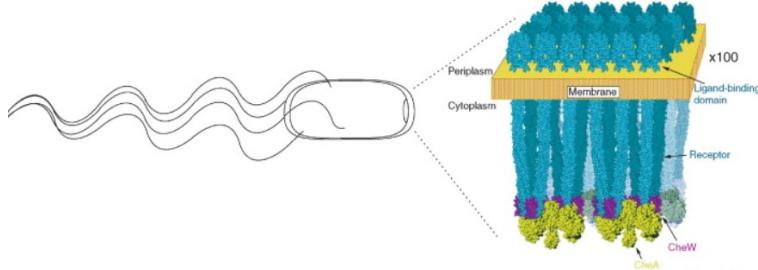


Figure 21: The bacterial "nanobrain" in examples such as *E. coli*, showing the chemotaxis sensory array which is located at the cell pole. A small segment of this network of proteins is shown in the expanded view. (Cell.com, 2021).

In nonuniform mediums a motile cell can seek out more favorable regions by detecting the changes in concentration and comparing the concentrations at different places or times. Consider cells that continuously measure the fraction of receptors bound to a chemical over time. If the fraction of bound receptors is increasing this means the cell is moving in the right direction. Berg and Purcell note that the random walk can be biased most effectively if the measurements are made in a short time interval compared to the mean run length (1977). In other words, to accomplish effective chemotaxis, the time available for chemical gradient determination is small and is controlled by the length of a cell's run. Experimental results show that bacteria such as *E. coli* and *S. typhimurium* can effectively bias their random walks/tumbles because they only take around one second to determine chemical gradients and make temporal comparisons of concentrations. Numerically, Berg and Purcell find that it would be possible rand also effective (for certain sized cells) if the cell's receptors were to be dispersed widely over the surface of the cell rather than solely at the base of each flagellum or at the tip of the cell (1977). An even more optimizing technique (assuming it doesn't drastically increase the time it takes the cell to detect chemical gradients) would be if the cell could recognize exactly which receptor/s have captured a molecule, to then move in the direction of that receptor relative to the rest of the receptors on the cell's surface which are not capturing or capturing less molecules.

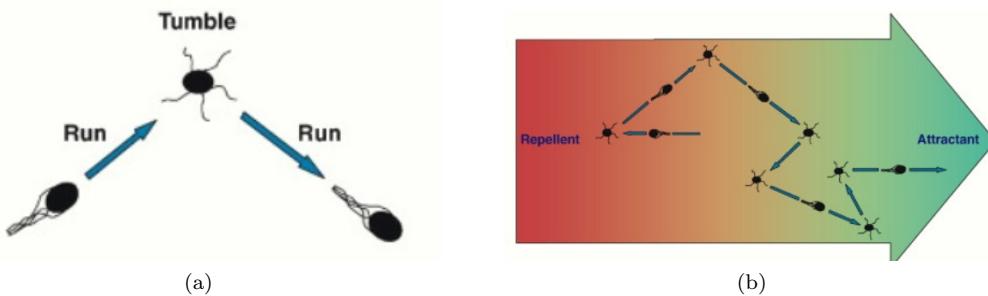


Figure 22: (a) Possible trajectory of tumble and run paths. (b) Tumble and run process away from repellents and towards attractants (Cell.com, 2021).

In developing an optimal chemotaxis method it is of interest whether or not a cell can employ a spatial mechanism in which it compares the chemical concentration at the front and the back or at the two sides. Berg and Purcell analyze this and find that a moving sphere that is the cell would find itself in an apparent spatial gradient with a specific decay length. Experiments have found that this decay length can correspond to an apparent gradient up to 600 times steeper than the real gradient. Ideally the apparent gradient as detected by the cell is as close as possible to the actual chemical gradient, so this is a problem we encounter

if we consider the cell to have receptors only at its front and back. Although it is possible for a bacterium in a spatial gradient to be able to compare the chemical concentration at each end of its cell, it's important to note that the concentration differences at the two ends would be too small to be effective for instantaneous comparison (Alder, 1975). This spatial mechanism is also analyzed in larger cells such as the slime mold *Dictyostelium discoideum* in which it is found to be much more effective. To sense the spatial gradient in the chemical by means of temporal comparisons, it would take about 17 seconds for a crawl or run velocity of  $0.2 \mu\text{m}/\text{s}$ . Ideally, the less time it takes to sense the chemical gradient the better. Although this sensing time is slightly long, it is still feasible and a temporal mechanism of sorts cannot be fully ruled out from an optimal chemotaxis mechanism.

An important aspect of chemotaxis includes the cellular mechanism of information processing and dealing with noisy measurements of chemical concentration. A cell has to accurately be able to detect concentration gradients in order to successfully survive and find food. Micali and Endres discuss how *E. coli* can correctly process information from their surroundings (2015). Mutual information is a measure of statistical dependence between two random variables whether it be linear or nonlinear. This theory says that if an input-output relationship has higher mutual information, it results in better signal inference. Enhancing and increasing this information gain is favorable for cells but comes at the expenditure of energy. Theory shows that energetic cost limits the accuracy of sensing. Micali and Endres also explain how there are three independent constraints to the sensing mechanism: "receptors and their averaging time, the number of downstream proteins, and fuel consumption" (2015). Although cells cannot adjust one of these constraints to alter the others, studies do show that *E. coli* optimally allocates resources to avoid unfavorable bottlenecks in the sensing pathway. To succeed in chemotaxis, these cells need to be able to detect and respond to shallow and rapidly changing external gradients. Although these cells are able to accomplish this in a relatively optimal fashion, there is still a trade-off in that it takes more energy to do so. It makes sense that an optimal chemotaxis strategy consists of a larger energy expenditure as long as the cell is not significantly depleting its available energy. Expending more energy to produce negative feedback during adaptation also reduces the adaptation error as well as variability in adaptation (Micali and Endres, 2015).

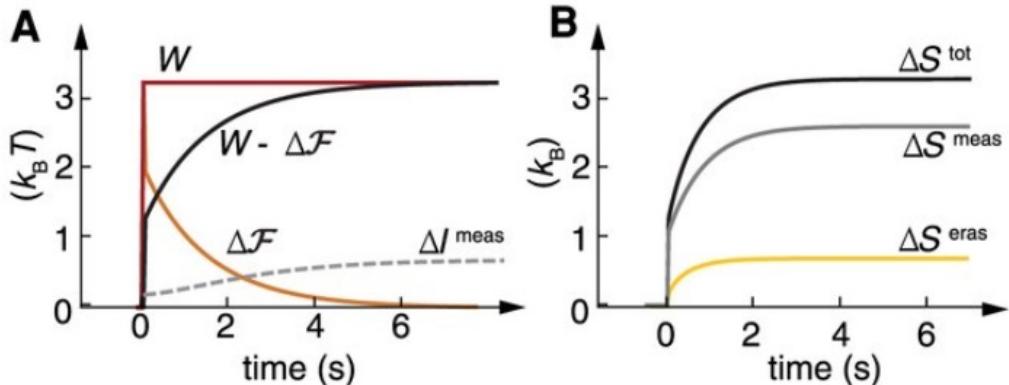


Figure 23: Thermodynamics of adaptation in an equilibrium sensory adaptation. (A) Energetic cost as a function of time given by the work provided by the environment (red), free energy change of the system (orange), and dissipated work (black), compared to the measured information (grey dashed), which gives the lower bound at every time. (B) Total entropic cost (black) and decomposition in measurement (gray) and erasure (yellow) (Figure and caption by Sartori et al., 2014).

Expanding on this idea and trade off between energy consumption and optimal chemotaxis, the ability of a cell to be able to adjust how much energy it is using or how fast it operates at different times in the chemotaxis process seems like an efficient characteristic. Although it might take more time to accomplish, if a cell is able to initially move in the correct direction of higher chemical concentration, it will be able to spend less energy undergoing the tumbling motion. This seems like an advantage which could result from multiple receptor sites spread around the cell's surface in that the cell can better detect chemical gradients allowing it to undergo less random motion. Indeed, this is already a trait of single cell chemotaxis as we

see that increased attractant concentration suppresses tumbling, leading to biasing of the random walk and allowing the cells to remain in these areas. These small cells are too short to detect a change in attractant gradient along their length and do not have receptor sites around their entire body. However they are able to accomplish optimal chemotaxis because they can communicate the detected chemical gradient at the membrane receptors to the flagellar motors (Tindall et al. 2008). In terms of communication, it is also of interest any ability these cells have to communicate with each other, which could result in a faster and more efficient process in locating areas of high chemical concentration. Once some of the cells have reached that area, a mechanism of signaling to other cells that food is in that area would be optimal for the cellular group as a whole.

It is important to note that the cells are able to adjust their rate of tumbling in response to both spatial and temporal gradients (Alder, 1975). The fact that a cell can distinguish between concentration gradients over time leads one to ponder the topic and extent of memory in cells, as one would consider this to be a beneficial and optimal trait for cellular chemotaxis. It is understood that temporal comparisons are a function of their receptors which sense chemicals and also remember the past via methylation status at specific glutamate residues (Sourjik and Wingreen, 2012).

In summary, the smaller the cell, the less effective are receptors spread on its surface, due to the interference of molecule sensing and input because the receptors are too close to each other. Time is of essence, because energy expenditure, especially during random tumbling will need to be replenished, therefore high sensitivity to certain attractive gradients as well as repellent gradients is important. Communication amongst organisms and cells is useful to shorten the search, meaning that if one cell has found nutrients, then direct "run" action can be taken toward the source, and less energy is spent randomly searching by tumbling. Receptors are better utilized by cells when distributed all over the surface as opposed to either or both head and tail end. Local stirring has advantage by moving fluid closer to the cell but could replace random tumbling and increase energy expenditure. To accomplish effective chemotaxis, the time available for chemical gradient determination is small and is controlled by the length of a cell's run, and a good technique would be if the cell could recognize exactly which receptor(s) have captured a molecule, to then move in the direction of that receptor relative to the rest of the receptors on the cell's surface which are not capturing or capturing less molecules. Expending more energy to produce negative feedback during adaptation also reduces the adaptation error as well as variability in adaptation. If a cell is able to initially move in the correct direction of higher chemical concentration, despite of needing more time, it will be able to spend less energy undergoing the tumbling motion.

### 3 HW6

#### Force on Ends of DNA Chain

Here we analyze how a force applied to the ends of a DNA chain effects its spatial evolution. Keeping the chain stiffness constant set to 0, and the chain length,  $n$  set to 32, we can vary the external force,  $f_{ext}$  so that we can analyze its relationship to the average end to end distance of the chain. Code was added to the script, dna.py, to record the difference in position between the chain's two endpoints after every time step in the simulation. Averaging these values over a long enough time gives a good estimate for the average end to end distance corresponding to a certain system external force value. This was repeated for a range of  $f_{ext}$  values and a plot showing the relationship between  $f_{ext}$  and the average end to end distance in the  $x$  direction,  $\langle x \rangle$ , was made. In terms of the distance between the chain's two endpoints, what we want is the vector average of the end to end distance vector. However, by symmetry, only the average along the direction of  $f_{ext}$  is non-zero. In this case  $f_{ext}$  acts in the  $x$  direction so we only need to consider the projection in the  $x$  direction. This is useful when comparing our data to theory, in this case the Marko-Siggia equation, because this projection along the  $x$  direction will give  $x = 0$  when  $f_{ext} = 0$  which agrees with the theory.

In terms of calculating the average distance between ends of the chain, what we want to consider is data after the system has equilibrated. As seen in figure 25, a plot of the end to end length in  $x$  vs.  $t$ , ( $x(t)$ ), can show us around what time the initial transients start to die off. These initial transients can be disregarded when calculate  $\langle x(t) \rangle$  to give a more accurate value. Ultimately we chose to disregard the first 1/5 of the  $x(t)$  data. As we can see in figure 24,  $x(t)$  starts to hover near its average value after this time has passed. As expected, for larger values of  $f_{ext}$ , the initial transients are much smaller than  $\langle x(t) \rangle$  so it is more important to disregard these initial transients for larger  $f_{ext}$ . Nevertheless we chose to disregard the first 1/5 of the  $x(t)$  data in calculating  $\langle x(t) \rangle$  for all of the  $f_{ext}$  values.

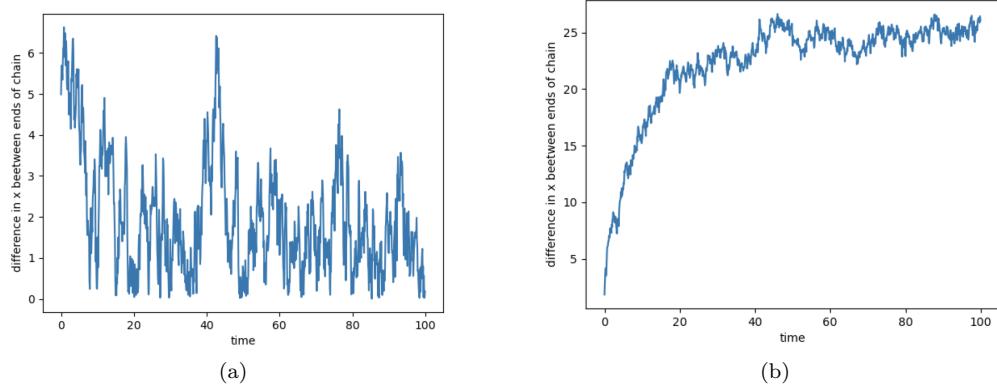


Figure 24: (a) Chains end to end distance vs.  $t$  for  $f_{ext} = 0.2$ . Notice there is a small amount of initial transients.(b) Chains end to end distance vs.  $t$  for  $f_{ext} = 4$ . Notice the initial transients are more prominent and therefore contribute to lessening the average end to end distance if not disregarded.

Below is a plot of  $f_{ext}$  vs.  $\langle x \rangle$  from our simulation data as well as the overlapped Marko-Siggia model. The Marko-Siggia model is described by:

$$F_{ext} = \left( \frac{k_B T}{L_p} \right) \left[ \frac{1}{4(1 - x/L_0)^2} - \frac{1}{4} + \frac{x}{L_0} \right] \quad (16)$$

where  $k_B$  is Boltzmann's constant,  $T$  is the absolute temperature,  $L_p$  is the polymers persistence length, and  $L_0$  is the chain length, which is  $n = 32$  total segments. For our simulation the first term,  $k_B T / L_p$  simplifies to 1. We are considering the chains spring coefficient to be zero, which makes  $L_p = 1$ . And we incorporate  $k_B$  into the temperature such that  $k_B T = 1$ . As seen in figure 13, this model fits the data well. We can

see the exponentially increasing relationship between  $f_{\text{ext}}$  and the average distance between the ends of the chain. This asymptotic behavior for large  $f_{\text{ext}}$  means that  $\langle x \rangle$  eventually stabilizes to a value close to the size of the chain length since the distance between the endpoints cannot exceed the length of the chain without breaking it. In other words, for increasingly large  $f_{\text{ext}}$  the chain will reach the max end to end distance quicker and then for the most part will stay in that configuration for the rest of the simulation.

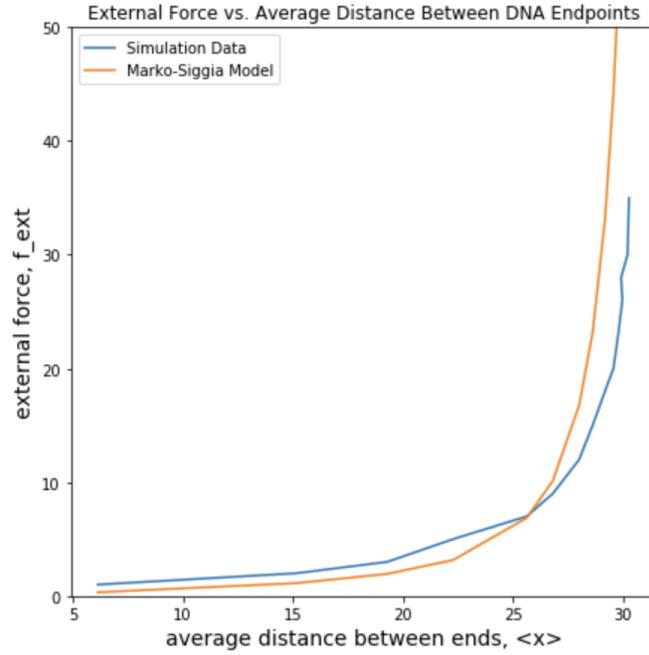


Figure 25: Plot of simulation data of  $f_{\text{ext}}$  vs. average distance between chain endpoints in  $x$  direction,  $\langle x \rangle$ . The Marko-Siggia model is also shown as described by equation 16.

## Brownian Motion Related to DNA Tethering

If we look at a particle's behavior in an optical trap, it can be estimated as having a linear restoring force towards the center of the trap. This causes the Brownian motion of the particle to become a roughly Gaussian distribution around the center of the optical trap (although the exact shape of the position distribution depends on how good the trap is). Similarly, we can model DNA as multiple connected particles undergoing Brownian motion. This means that the Brownian motion at all points on the DNA would make it inclined to head towards the center. However, the DNA also has self-interfering forces that would make it unable to completely crumple up in the center. This means that we expect the DNA to randomly wobble around the center of the trap with the Brownian motion roughly following the Gaussian distribution that is caused by the optical trap. When there is no external force that is applied on the dna molecule in the simulation, we are able to see this Brownian motion be less inhibited, and it can engage in a self-avoiding walk. As more external force is applied, we see the dna molecule become stretched out and rigid, with the dna molecule having a significant decrease in amount of Brownian motion.

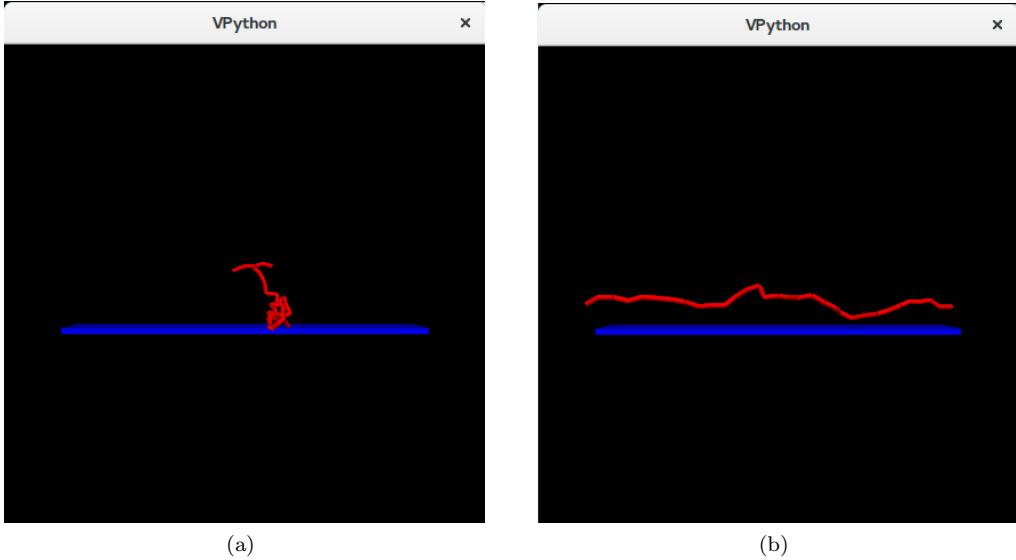


Figure 26: (a) The graphical output when the  $f_{ext}$  is 0.0. The dna molecule is shown to be concentration to be in the center as much as possible, though being limited by its self-interfering forces. (b) The graphical output when the  $f_{ext}$  is 4.0. The dna molecule is significantly stretched out with the ends showing little motion.

## Linear Biological Structures with Local Stiffness

Many tissue and cell functions use local stiffness to be motile and to protect cells and organs. Integrins, heterodimeric transmembrane proteins, have an important role in cellular mechano-sensing since they connect the cytoskeleton and the extracellular matrix in clusters. Most integrins that are known thus far can detect multiple ligands that share common binding motifs. Integrins, being bi-directional signaling receptors, where intracellular proteins bind to the tail region resulting in conformational changes in the head region increasing affinity for extracellular ligands. Ligand binding on the other hand can trigger conformational changes activating intracellular signaling cascades. Ligand binding can further promote integrin clustering, essential for cell spreading (Handorf et al, 2015).

The extracellular matrix (ECM), a complex protein mesh or net-like structure, forms the scaffold where cells adhere. It creates a mechanical support for cells and tissues. It further acts as a reservoir for growth factors, cytokines, as well as proteolytic enzymes. It comes in two broad classes, of which one is the basement membranes which are thin structures providing a 2D substrate where epithelial and endothelial cells adhere.

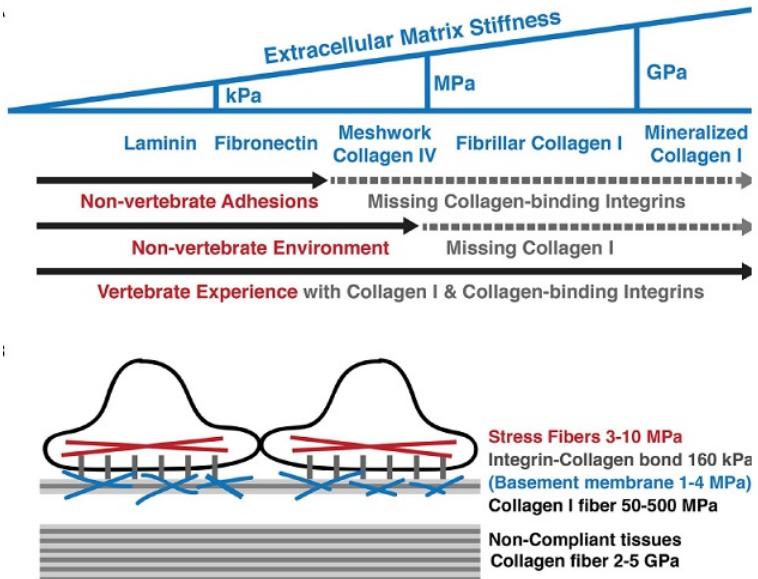


Figure 27: Extracellular Matrix Stiffness through evolutionarily conserved fibronectin, collagen IV, and laminin. These form meshwork-like matrices with stiffness variability in kilopascals (kPa), megapascals (MPa) and gigapascals (GPa) (Tang, 2020).

The second class is connective tissue which provides a 3D scaffold with components such as collagens, proteoglycans, and glycosaminoglycans. Collagen fibers are created specific to each tissue, such as stiff tissues like tendons for tensile strength or thin meshworks in the cornea for optical transparency (V. Tang, 2020). The cytoskeleton (CSK) is composed of protein filaments which enable maintenance of shape and mechanical strength in cells, enabling them to withstand external forces, yet remaining dynamic. In mammals the CSK is made of three different protein filaments, actin, microtubules and intermediate filaments. Actin and intermediate filaments are generally considered to provide the main source of cell stiffness, and microtubules, which are more rigid, can provide resistance to compression forces. Ventral stress fibers can span almost the entire cell length and are anchored at both ends to focal adhesions. The shorter dorsal stress fibers are connected to focal adhesions (FA) at one end. The degree of actin cross-linking and bundling will increase with an increase in substrate stiffness. Microtubules (MT) are well-known for their role in cell polarity and migration and in 2D substrates, they do not appear to influence the degree of cell spreading, yet they are crucial for cell spreading in 3D collagen networks. Microtubules may be required for delivery of a relaxing factor by kinesin motors, and crosslinking of growing microtubules to actin stress fibers is required to guide the MTs to FA sites (Jansen et al, 2015).

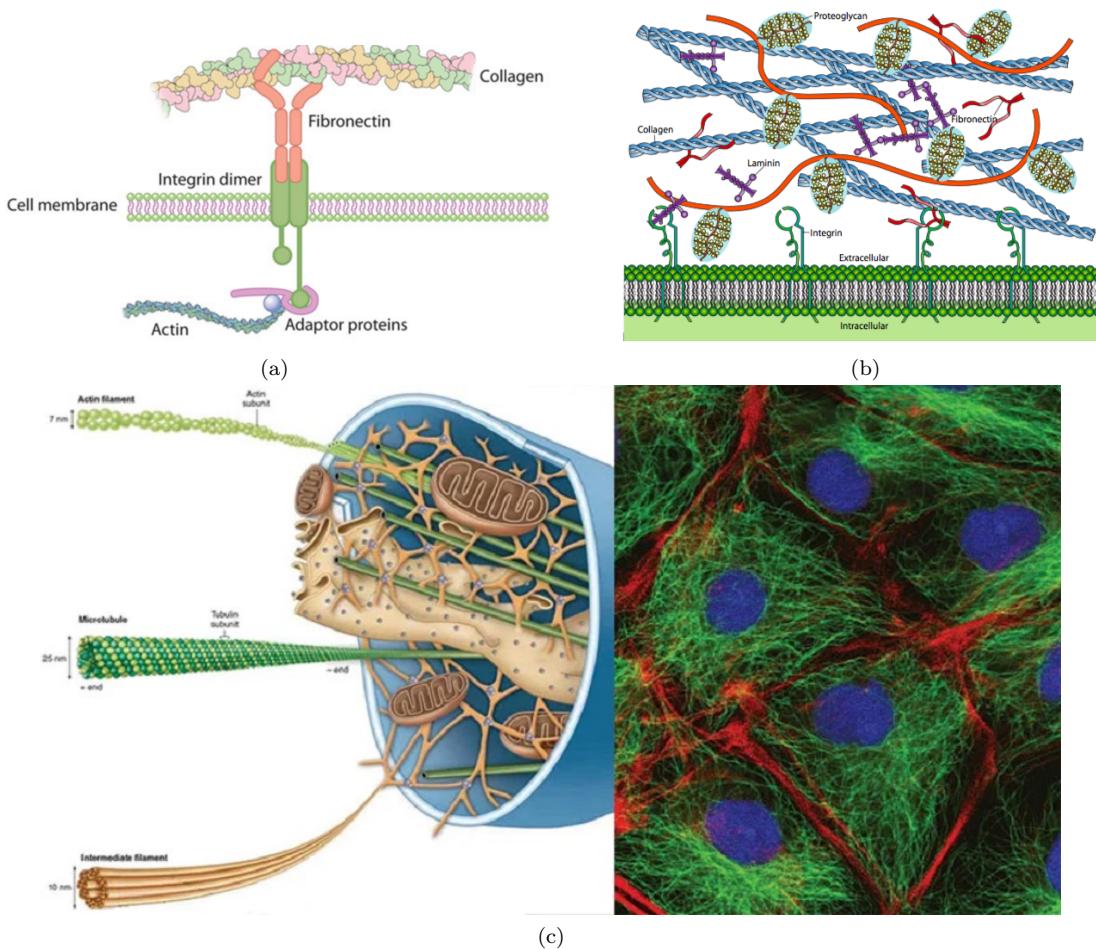


Figure 28: (a) Integrin, showing heterodimers, acting inside the cell as well as collagen on the outside bound by fibronectin (Openwetware.org, Dague 2021). (b) Extracellular Matrix showing adhesion of collagen and proteoglycan (Libretexts. org, Wong 2021). (c) Microtubules, Actin and Intermediate Filaments shown in the cellular Cytoskeleton (microbenotes.com 2021.)

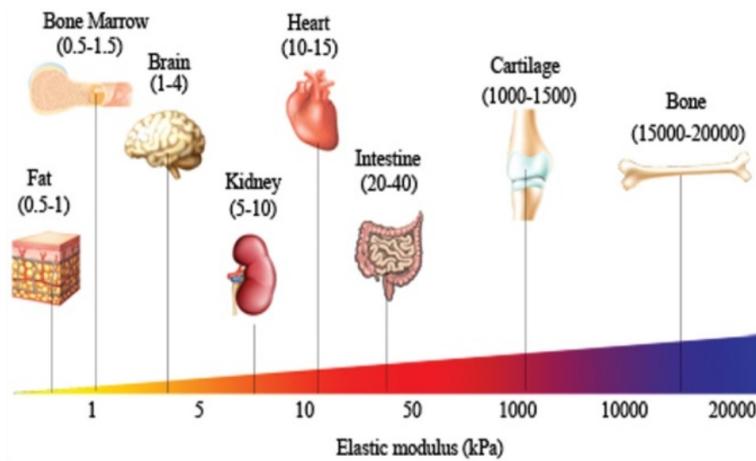


Figure 29: Elastic Modulus in human tissues showing very tissue-specific stiffness which is correlated according to tissue functions (Handorf et al. 2015).

## Creating Chain Stiffness in Molecules

Microtubules are made of individual strings of  $\alpha$  and  $\beta$  tubulin heterodimers or protofilaments. These are arranged parallel to each other and form a hollow tube of 25 nm total diameter. Neighboring protofilaments will shift relative to each other, hence creating helicity. The rise per turn and the number of protofilaments can vary, yet when observed *in vivo*, 3 monomers per turn and 13 constituent filaments are the most common configuration. This protofilament architecture makes microtubules a model system for the generalized theory of wormlike bundles (Taute et al, 2008). All known microtubules can adopt skewed protofilament geometries, yet the 13.3 configuration mentioned above is expected and most common *in vivo* (Donhauser et al, 2010). Tubulin has two heterodimers which contain two GTP-binding sites, where the GTP bound to the  $\alpha$ -tubulin site is stable acting as a structural function. GTP bound to the  $\beta$ -tubulin site may be hydrolyzed to GDP shortly after assembly. GTP bound at the site makes tubulin dimer more prone to polymerization, yet tubulin dimer with GDP bound tends to depolymerize. This GTP cycle is thus essential for the dynamic instability of the microtubule which move through the cell by polymerization on one end and de-polymerization on the other.

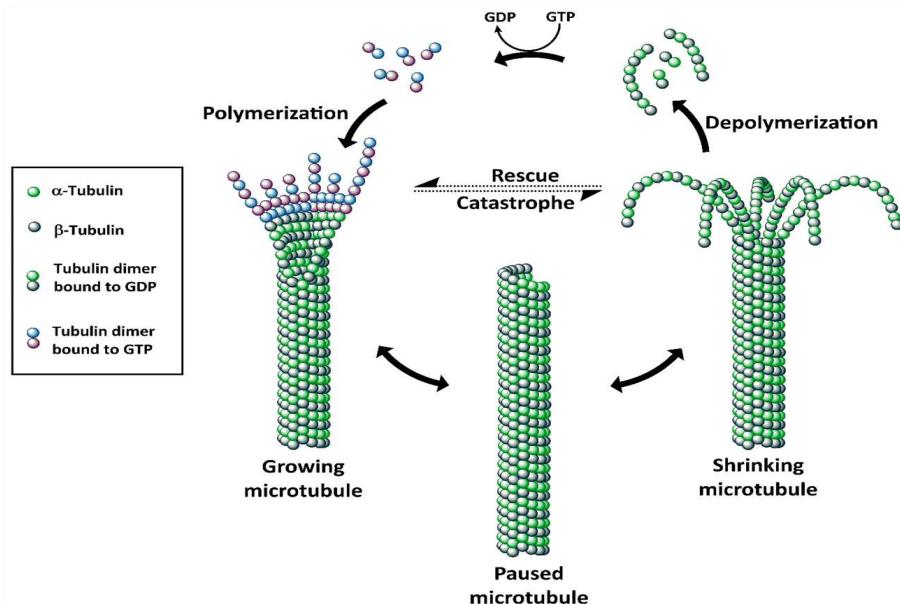


Figure 30: On the left is the polymerizing microtubule with GTP bound, on the right is a depolymerizing microtubule bound to GDP. In the center we see the stage between the two waiting either  $\alpha$  or  $\beta$  tubulin and GDP or GTP interaction defining in which direction it will react (creative-biolabs.com, 2021).

Actin has over 95% conservation in its primary structure and is one of the most highly-conserved proteins. The filament polymerization goes through three phases: nucleation, elongation and steady state, usually comprised of three actin monomers in complex. During elongation, monomers are quickly added to the filament at the barbed end, often facilitated by additional elongation factors like formin. For this process to occur, plus-end of the filament needs to be exposed, requiring removal of capping proteins (mechanobio.info, 2021). Actin molecules are globular proteins comprised of 375 amino acids, where each actin monomer has tight binding sites. These sites mediate head-to-tail interactions with two other actin monomers, polymerizing to form filaments. Each monomer is rotated by 166 degrees in the filaments, and this gives the appearance of a double-stranded helix. This helical shape gives the actin its strength, and since the filaments are all in the same direction, they have a distinct polar end. The incorporation rate into filaments of actin monomers is proportional to their concentration. A critical concentration of actin monomers exists at which the polymerization into filaments equals the dissociation, therefore monomers and filaments are in apparent equilibrium (Cooper, 2020). Further examples are primary protein structures which are a sequence chain of amino acids, themselves not in stiff conformations, but the secondary structures such as  $\beta$ -sheets, which are structures

made from extended  $\beta$ -strand polypeptide chains give slightly more stability and stiffness. The strands are linked to their neighbors by hydrogen bonds, and this extended backbone conformation gives resistance to stretching, where  $\beta$ -sheets in proteins possibly carry out low-frequency accordion-like motion. Besides  $\beta$ -sheets,  $\alpha$ -helices are the other secondary structure that will give more stability to amino acid chains. In these structures, intramolecular interactions, such as hydrogen bonding between backbone amine groups and carbonyl functional groups are crucial in maintaining the 3D shape. Each helical turn in an  $\alpha$ -helix has 3.6 amino acid residues. The R groups, which are the variant groups, of the polypeptide will protrude out from the  $\alpha$ -helix chain. Tertiary and quarternary protein structures are folded upon themselves giving 3D structures or consist of more than one protein. These are the most stable protein or molecule structures in organisms and need special catalysts or interactions to change their shape and functions. Cellulose is one of the most abundant carbohydrate while serving as a structural biopolymer. It is made up of  $\beta$ -1,4 linkages between glucose molecules, and creates a linear structure where each glucose monomer is flipped upside down compared with the one before it. When fibers from one cellulose strand are then aligned with another fiber they will usually form strong hydrogen bond interactions that build strong fiber networks (Wou.edu, 2021). An Industry example of strong polymer stiffness would be poly(*p*-phenylene terephthalamide) or PPTA. It has very stiff polymer chains due bonding of rigid phenylene rings in the para position (certain locations on the phenylene ring). Advantages of PPTA are the presence of amide groups at regular intervals along the linear macro-molecular backbone, resulting in extensive hydrogen bonding in a lateral direction between adjacent chains. The structure is anisotropic, giving higher strength and modulus in the fiber longitudinal direction than in the radial direction. Furthermore, weak hydrogen bonds in the transverse direction result in low shear moduli, poor transverse properties, and low compressive strength. The typical properties of commercial aramid fibers can be found for instance in Kevlar 49, which is the dominant form used today in structural composites because of its high modulus (M. Ertekin, 2017).

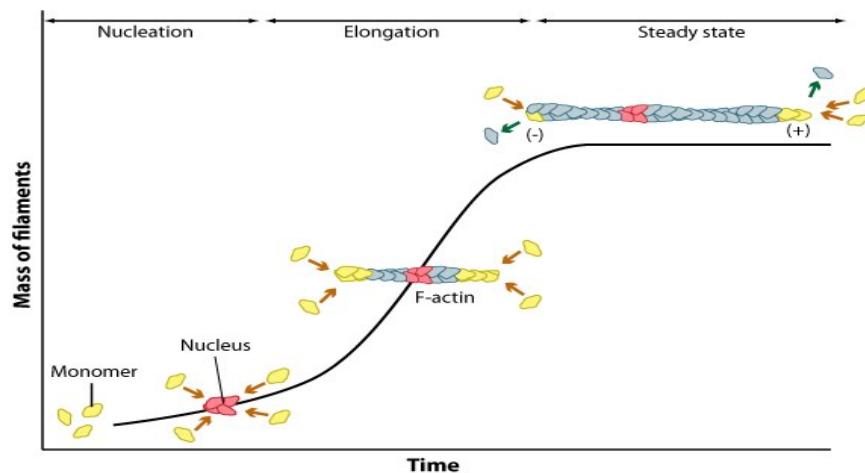


Figure 31: Actin filament assembly showing nucleation, elongation and steady state phases (mechanobio.info, 2021).

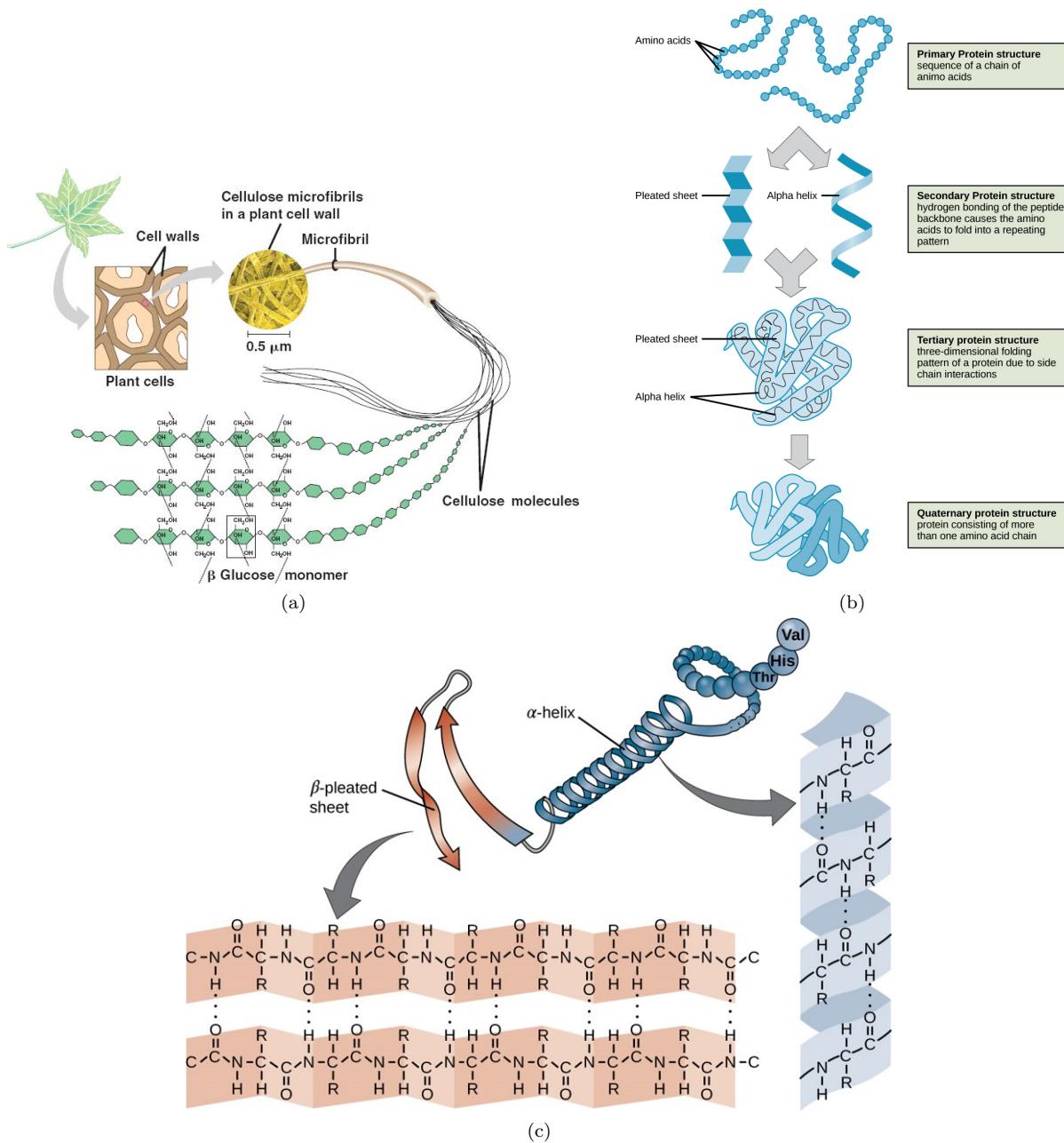


Figure 32: (a) Glucose subunits where glucose monomers are flipped relative to their neighbor giving a linear structure forming hydrogen bonds with neighboring strands resulting in strong fibrous material. (b) The 4 levels of possible protein structures. (c)  $\alpha$ -helices and  $\beta$ -sheet are structural motifs found in proteins, and held together by hydrogen bonding between the amine and the carbonyl oxygen within the amino acid backbone. (Wou.edu, 2021)

## Biological Importance of Force vs. Extension

Cells perform many integrated functions such as metabolism, control, sensing, communication, growth, remodeling, reproduction and cell death or apoptosis. They further perform an immense variety of functions such as synthesis, sorting, storage as well as transport of biomolecules, expression of genetic information, recognition, transmission, and transduction of signals, as well as conversion between different forms of energy. The majority of these processes involve mechanical force, or deformation, at the cellular, sub-cellular and molecular levels. Biomolecular motors convert chemical energy into mechanical work in performing a great range of functions. Contractile forces are generated within the cell during cell migration in order for the cell to be motile, but they also enable cells to sense the stiffness of the surrounding substrate and respond accordingly. Cell conditions are dependent upon or regulated by their mechanical environment, yet muscle and cochlear outer hair cells perform their mechanical function either by converting an electrical or chemical stimulus into mechanical motion. Ability of cells to sense mechanical force or deformation and transduce these mechanical signals into a biological response is of utmost importance, such as in endothelial cells which can recognize the magnitude, steady or pulsing mode, laminar or turbulent type, as well as duration of applied shear flow. The response will be accordingly, hence maintaining healthy endothelium. We find Vascular smooth muscle cells in the arterial wall, which will remodel themselves when subjected to pressure induced wall stress. Fibroblast cells crawl by pulling the cell body forward using contractile forces, and bone alters its structure so it can adapt to changes in mechanical environment such as in spaceflight. Stem cells can sense the elasticity of surrounding substrates and then differentiate into different phenotypes (Bao et al., 2010)

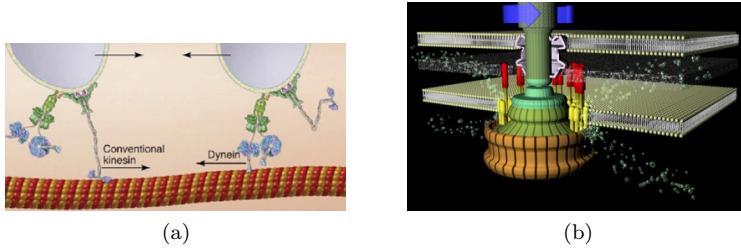


Figure 33: (a) Linear motors kinesin and dynein, which walk along microtubules in opposite directions, carrying vesicles during intracellular transport. (b) Rotary motor depiction of a bacterial flagellum (Bao et al., 2018).

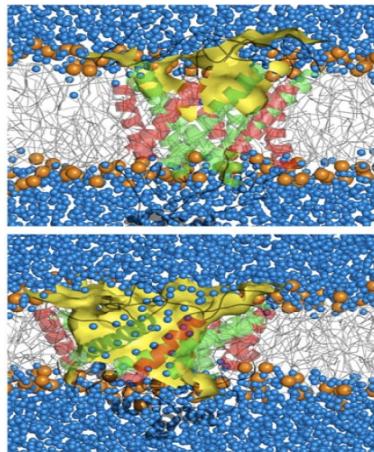


Figure 34: A mechanosensitive ion channel with small (top) and large (bottom) opening under the action of membrane tension. (Bao et al., 2018).

Gastrulation is a key embryonic event during which the 3D framework of an animal is established. In most animals, it relies on the complex interplay of multiple force-generating engines such as convergent extension and convergent thickening, which will spearhead how the body, all organs, and extremities will be formed (Shinuo et al., 2018).

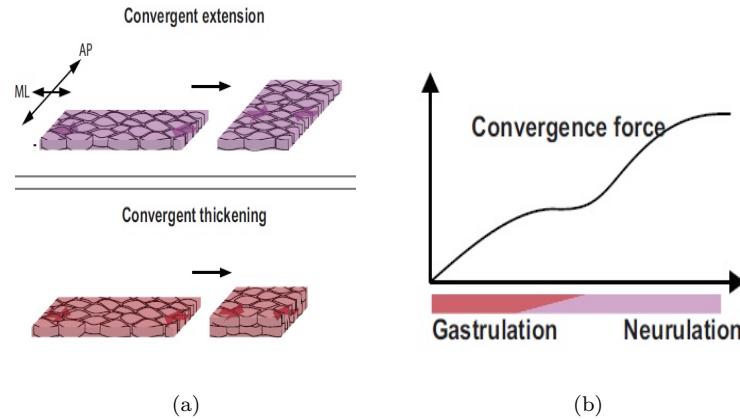


Figure 35: (a) The process by which tissue elongates, called convergent extension. Process called convergent thickening where the tissue becomes thicker in the direction at right angles to the convergent extension. (b) Convergence force (y-axis) increases through gastrulation, hits a plateau before increasing again in late neurulation. (Shinuo et al., 2018)

The most important example of force vs. extension one should mention is the tension and extension needed for the separation in DNA in all living organisms we know of. For the replication process to work properly, DNA needs to be separated and unwound in order to go through the whole replication complex. This is done by the helicase, several enzymes, and the ribosomal units, since the DNA strand is not readily able or willing to uncoil and separate (Jansen et al., 2015).

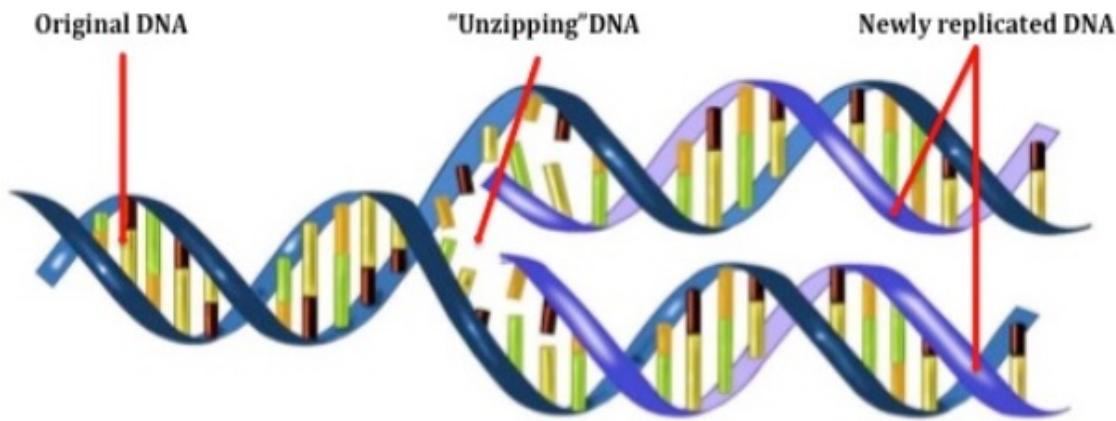


Figure 36: DNA original strand being separated by enzymatic forces and mechanisms showing the unzipped strand and how replication of DNA continues (northwestern.edu, 2021).

## 4 Sources and References

Adler, J. (1977). Chemotaxis in Bacteria. 18.

Astumian, R. Dean, and Martin Bier. "Mechanochemical coupling of the motion of molecular motors to ATP hydrolysis." Biophysical journal 70.2 (1996): 637-653.

Bao, Gang, et al. "Molecular biomechanics: the molecular basis of how forces regulate cellular function." Cellular and Molecular Bioengineering 3.2 (2010): 91-105

Benedetti, Francesca et al. "Mycoplasmas-Host Interaction: Mechanisms of Inflammation and Association with Cellular Transformation." Microorganisms vol. 8,9 1351. 4 Sep. 2020, doi:10.3390/microorganisms8091351

Berg, Howard C., Purcell, Edward M. "Physics of Chemoreception". Biophysical Journal, V 20, (1977).

CavendishLecturesOnDiffusion.pdf, Physics 180 UCSC Spring 2021 Lecture Material.

Charrier, Elisabeth E, and Paul A Janmey. "Mechanical Properties of Intermediate Filament Proteins." Methods in enzymology vol. 568 (2016): 35-57. doi:10.1016/bs.mie.2015.09.009

[https://www.cell.com/current-biology/comments/S0960-9822\(02\)01424-0](https://www.cell.com/current-biology/comments/S0960-9822(02)01424-0)

<https://cnx.org/contents/GFy-h8cu@9.85:FrKEod5j@6/Potential-Kinetic-Free-and-Activation-Energy>

Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. Structure and Organization of Actin Filaments. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK9908>

<https://www.creative-biolabs.com/drug-discovery/therapeutics/microtubule-aggregation.htm>

<https://ct-stem.northwestern.edu/curriculum/preview/13/page/3>

<https://www.degruyter.com/document/doi/10.1515/jnet-2018-0008/html>

Donhauser, Zachary J et al. "Mechanics of microtubules: effects of protofilament orientation." Biophysical journal vol. 99,5 (2010): 1668-75. doi:10.1016/j.bpj.2010.06.065

Engel, Thomas, and Philip J. Reid. Thermodynamics, statistical thermodynamics, and kinetics. Upper Saddle River, NJ, USA: Prentice Hall, 2010.

<https://en.wikipedia.org/wiki/Chemotaxis>

Ertekin, Mustafa. "Aramid fibers." Fiber Technology for Fiber-Reinforced Composites. Woodhead Publishing, 2017. 153-167.

Andrew M Handorf, Yaxian Zhou, Matthew A Halanski Wan-Ju Li (2015) Tissue Stiffness Dictates Development, Homeostasis, and Disease Progression, Organogenesis, 11:1, 1-15, DOI: 10.1080/15476278.2015.1019687

Hathcock, David, et al. "Myosin V executes steps of variable length via structurally constrained diffusion." Elife 9 (2020): e51569

Introduction to Extracellular Matrix and Cell Adhesion. Axolotl Academica Publishing, 3 Jan. 2021, <https://bio.libretexts.org/@go/page/16170>.

<https://www.invivogen.com/review-mycoplasma>

Janetopoulos, Christopher, and Richard A. Firtel. "Directional sensing during chemotaxis." FEBS letters 582.14 (2008): 2075-2085.

Jansen, Karin A., et al. "A guide to mechanobiology: where biology and physics meet." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1853.11 (2015): 3043-3052.

Lakhani V, Elston TC (2017) Testing the limits of gradient sensing. PLOS Computational Biology 13(2): e1005386. <https://doi.org/10.1371/journal.pcbi.1005386>

Micali, G., Endres, R. G. (2015). Bacterial chemotaxis: Information processing, thermodynamics, and behavior. ArXiv:1512.02542 [q-Bio]. <http://arxiv.org/abs/1512.02542>

<https://www.mechanobio.info/cytoskeleton-dynamics/what-is-the-cytoskeleton/what-are-actin-filaments/how-do-actin-filaments-grow>

<https://microbenotes.com/cytoskeleton>

<https://www.nap.edu/download/9638>

Nelson, David L., Albert L. Lehninger, and Michael M. Cox. Lehninger principles of biochemistry. Macmillan, 2008.

[https://openwetware.org/wiki/Cell\\_Adhesion\\_and\\_Integrins\\_by\\_Bry\\_Dague](https://openwetware.org/wiki/Cell_Adhesion_and_Integrins_by_Bry_Dague)

Sartori, Pablo, et al. "Thermodynamic costs of information processing in sensory adaptation." PLoS Comput Biol 10.12 (2014): e1003974.

Shinuo, Weng, and John B. Wallingford. "May the force be with you." eLife 7 (2018).

Sourjik, V., Wingreen, N. S. (2012). Responding to chemical gradients: Bacterial chemotaxis. Current Opinion in Cell Biology, 24(2), 262-268. <https://doi.org/10.1016/j.ceb.2011.11.008>

Speers, A., and Jessica Forbes. "Yeast: an overview." Brewing Microbiology. Woodhead Publishing, 2015. 3-9.

Stryer, L., J. L. Tymoczko, and J. M. Berg. "Biochemistry 5th ed Freeman." WH and Company 41 (2002).

Tang, Vivian W. "Collagen, stiffness, and adhesion: the evolutionary basis of vertebrate mechanobiology." Molecular biology of the cell 31.17 (2020): 1823-1834.

Tindall, M. J., Maini, P. K., Porter, S. L., Armitage, J. P. (2008). Overview of Mathematical Approaches Used to Model Bacterial Chemotaxis II: Bacterial Populations. Bulletin of Mathematical Biology, 70(6), 1570-1607. <https://doi.org/10.1007/s11538-008-9322-5>

Wang, Y., Yu, Y., Li, GB. et al. Mechanism of microtubule stabilization by taccalonolide AJ. Nat Commun 8, 15787 (2017). <https://doi.org/10.1038/ncomms15787>

<https://web.mst.edu/~gbert/LGH/hepler3.html>

<https://wou.edu/chemistry/courses/online-chemistry-textbooks/ch103-allied-health-chemistry/ch103-chapter-6-natural-products-and-organic-chemistry>

Xie, Yujie, et al. "A highly-efficient La-MnO<sub>x</sub> catalyst for propane combustion: the promotional role of La and the effect of the preparation method." *Catalysis Science Technology* 6.23 (2016): 8222-8233.